

---

# Visualising discourse coherence in non-linear documents

Clara Mancini, Donia Scott and Simon Buckingham Shum

Centre for Research in Computing  
The Open University  
Walton Hall, Milton Keynes  
MK7 6AA, United Kingdom

[C.Mancini@open.ac.uk](mailto:C.Mancini@open.ac.uk); [D.Scott@open.ac.uk](mailto:D.Scott@open.ac.uk); [S.Buckingham.Shum@open.ac.uk](mailto:S.Buckingham.Shum@open.ac.uk)

---

*ABSTRACT.* To produce coherent linear documents, natural language generation systems have traditionally exploited the structuring role of textual discourse markers such as relational and referential phrases. These coherence markers of the traditional notion of text, however, do not work in non-linear documents: a new set of graphical devices is needed together with formation rules to govern their usage, supported by sound theoretical frameworks. If in linear documents graphical devices such as layout and formatting complement textual devices in the expression of discourse coherence, in non-linear documents they play a more important role. In this paper, we present our theoretical and empirical work in progress, which explores new possibilities for expressing coherence in the generation of hypertext documents.

*RÉSUMÉ.* Dans la production de documents linéaires, les systèmes de génération automatique du langage ont traditionnellement exploité le rôle structurel des marques textuelles du discours, comme les phrases relationnelles et référentielles. Pourtant, ces marques de la cohérence textuelle ne fonctionnent pas dans des documents non linéaires: des nouveaux dispositifs graphiques sont nécessaires avec des règles d'utilisation guidées par des structures théoriques. Si des dispositifs graphiques, comme composition et formatage, contribuent à l'expression de la cohérence dans des documents linéaires, leur rôle est beaucoup plus important dans des documents non-linéaires. Dans cet article nous explorons des nouvelles possibilités pour représenter la cohérence dans des documents hypertextuels.

*KEY WORDS:* hypertext, cognitive coherence, document structure, visual discourse patterns

*MOTS-CLÉS:* hypertexte, cohérence cognitive, structure du document, motifs visuels du discours

---

## **1. Introduction**

There is a long and well-established literature on textual devices that signal the coherence structure of a discourse to the reader, within both theoretical (e.g., van Dijk, 1977; Halliday and Hasan, 1976; Grimes, 1975; Brown and Yule, 1983) and computational linguistics (e.g., Hobbs, 1985; Mann and Thompson, 1988; Schiffrin, 1987; Knott and Mellish, 1996). Most of the work so far addresses the traditional conceptualisation of text as a two dimensional array on a physical page, traversed in a set pattern (e.g., left to right, top to bottom in the Western tradition).

Hypertext, however, is different: it is read on a computer screen and is non-linear, with several reading paths available through the document. The reader moves from node to node by mouse-clicking on links. A node can be the equivalent of a traditional text page or can contain just a few sentences. A link can be a word in the text or a graphical element in the node. As nodes are connected via multiple outgoing and incoming links, the author can only partially control the order in which the reader will access them. In other words, with hypertext a new conceptualisation of text emerges as a three-dimensional array on a computer screen, which can be traversed in any number of ways (one can virtually move across the screen's surface in two dimensions as well as in depth into a third dimension).

The coherence markers of the traditional notion of text do not work as efficiently for this medium, therefore a new set of devices, not only textual but graphic, is needed together with formation rules to govern their usage, supported by sound theoretical frameworks. Being concerned with the presentation of medical information to patients and doctors in hypertext form, we explore new possibilities for achieving coherence in non-linear documents. Because in non-linear documents discourse is organized as a network of self-standing units rather than as a hierarchy of interdependent segments, our analysis of discourse coherence departs from the tradition whereby text is described as a hierarchical structure (e.g., Mann & Thompson, 1988). Instead, we take a cognitive approach according to which coherence is a characteristic of the mental representation that the reader constructs during the process of text interpretation (e.g., Johnson-Laird, 1983).

## **2. Coherence representation in linear text**

Text comprehension depends on the reader's ability to construct a coherent representation of what (she thinks that) the text is conveying. To do so the reader needs to be able to identify the conceptual relations (she thinks to be) holding between the set of discourse elements (whether sentences, paragraphs or entire text sections). Conceptual relations are primarily identified on the basis of the content of the related discourse elements, but in linear text their identification is facilitated by a number of formal cohesive elements.

Over the years, the study of text coherence has concentrated on two types of cohesive element: those which function at the level of discourse structure and those which function at the level of document structure. A lot of work has focussed on discourse structure. Whether data-driven (e.g., Halliday and Hasan, 1976; Martin, 1992; Knott and Dale, 1994) or theory-driven (e.g., Hobbs, 1985; Kamp and Ryle, 1993; Mann and Thompson, 1988; Sanders et al., 1993), this work has mainly studied the use of discourse markers and referring expressions. For instance, in the sentence “*Lucia arrived at work late because she had missed her train*” the two clauses are related through the connective *because* and through the pronoun *she*, whose semantic content facilitates the interpretive work of the reader.

Other work, on the other hand, has highlighted the role played by graphical features such as punctuation and layout in text organisation. In particular, Nunberg (1990) distinguishes text structure from syntactic structure. For Nunberg, text structure is characterised by *abstract* (semantic) *features* which can be realised by different *concrete* (syntactic) *features* such as punctuation and other graphical marks (parentheses, dashes, etc.), layout and formatting in general (section titles, emphasis, etc.). For instance, in the sentence “*Lucia arrived at work late: she had missed the train*” the same causal relation previously expressed by the connective “*because*” is now expressed by a colon. Likewise, in the sentence “*I had a busy morning: I had a work meeting, I went shopping, I picked up the children.*” the conjunctive relation between the second, third and fourth clause is expressed by a comma and the connective “and”, but it could be otherwise expressed by a bulleted list:

- I had a busy morning:*
- *I had a work meeting*
  - *I went shopping*
  - *I picked up the children*

Elsewhere, we propose that to account for the varying formulations of a text a separate descriptive layer is required, which we term *abstract document structure* (Power et al., 2003). As we show in previous work (Piwek et al., 2005; Power et al., 2003; Bouayad-Agha et al., 2000), the abstract document structure is an intrinsic part of Nunberg’s text structure (closely analogous to semantics) and can be conveyed by a range of concrete visualisations (the syntax). We explore the role of dynamic graphics as a concrete representation of abstract document structure – along with layout (e.g. use of indentation), punctuation (e.g., use of full stops) and cue phrases (e.g., use of adverbials such as ‘although’).

### 3. Abstract discourse structure: visual vs. textual

Different concrete features have different semiotic characteristics, in that whereas devices like cue phrases and punctuation are textual, devices like layout and formatting are visual. In written (alphabetical) text, the minimal linguistic unit is the character, a non-signifying differential element, whose combination generates

words, successively articulated to produce phrases, clauses, sentences, etc. (Saussure, 1922). As the character is a symbolic element, in written text the association between signifier and signified is not-motivated; rather, the correspondence between them is conventional. Because of this, in written text abstract concepts can be explicitly expressed. For instance, in the sentence “*I was late for the meeting because I had missed the bus*”, the relation of CAUSALITY holding between the segments is made explicit by the connective “because”.

Its symbolic nature also implies that text can deploy along a single line, which can be articulated using punctuation, dashes, parentheses and the like. These are purely graphical symbols, which signal different types of textual articulation and inflection, and whose use is also regulated by strict conventions. For instance, a period marks the end of a *text-sentence*, while a semicolon marks the end of a *text-clause*.

Substantially different from adverbials, punctuation and the like, layout and formatting in general transform the line of text into a visual configuration capable of conveying discourse structure on the space of the page. In visual configurations the association between a sign and its meaning is characterised by a degree of isomorphism, which makes this association partially motivated. For instance, consider again the sentence “*I had a busy morning: I had a work meeting, I went shopping, I picked up the children.*”, in which the clauses that follow the colon play an equivalent rhetorical role (Pander Maat, 1999). In the bulleted list version, this equivalence is expressed by the fact that the clauses are given the same visual rendering: each one starts on a new line with a bullet. Likewise, the title of the sections in a text will be visually more prominent than the title of the subsections in order to render the hierarchy of the text structure, just as emphasis is visually expressed through a format that stands out.

Unlike textual representations, visual representations tend to be regulated by conventions that are less strict and more dependent on the context of use. For instance, our list of clauses could be indented or not, bulleted, numbered or scored, but whatever the chosen configuration, it is important that all clauses are rendered in the same way (i.e., with parallel syntax) and occupy the same horizontal position under the first (introductory) clause. Even though they respond to flexible conventions, however, visual features can express discourse connections so effectively that the use of cue phrases or punctuation becomes redundant. So, in our bulleted list the use of connectives, commas and full stop is superfluous, as the conventions at work in the visual configuration of the list override the conventions that regulate the use of discourse connectives and punctuation.

#### **4. Coherence representation in non-linear text**

The devices described above constitute cohesive elements that can be used to express discourse coherence in linear text, either on paper or in electronic

documents that maintain linearity. However, discourse markers such as relational and referential connectives can only be effectively used when discourse units are arranged in a predefined sequence, where they are accessed in a univocal order. But because hypertext is a network of interconnected nodes, the order in which discourse parts will be accessed can only be partly controlled. Order can be established locally (a node can be linked to another node), but establishing global order and coherence through extended structures requires the imposition of constraints (e.g., restricted navigational paths – Bernstein, 1998) or the use of other expedients (e.g., transitional nodes – Bernstein and Greco, 2002). But both solutions in principle contradict the non-linearity of hypertext.

As it is a fundamental characteristic of hypertext that each node be accessible in more than one way, the use of relational and referential connectives to signal the discourse relation between nodes is problematic, especially for certain discourse genres. If, for instance, in literary hypertext a degree of ambiguity and indeterminacy is part of the ‘game’ (Douglas, 1991; Walker, 1999), in informative hypertext clarity and determinacy are important instead. Consequently, hypertext nodes tend to be written as self-standing units of text. A hypertext node typically will not use pronouns or referential phrases to refer to the content of another node, instead any information contained in the latter that would need to be referred to in the former has to be repeated. In fact, text sentences or paragraphs that are strongly related (for instance, by CAUSALITY) will normally be kept within the same node: since they constitute strongly inter-dependent discourse parts, the writer is reluctant to put them in different nodes, because the reader might miss one or the other. However, it is less problematic to separate into different nodes discourse parts that are less strongly related (for instance, by ELABORATION or BACKGROUND) and therefore less inter-dependent, so they can more easily be put into different nodes, their connection being expressed paratactically via a link (Mancini and Buckingham Shum, 2004). Finally, the same limitations that apply to discourse connectives also apply to punctuation and the like, which usually only work within nodes and do not facilitate the transition between link words and their target nodes.

If the non-linearity of hypertext does not lend itself to the use of textual features such as relational and referential connectives, or punctuation to signal the connection between nodes, however, things are different for visual features, because they work in space. Because of its technical characteristics, hypertext is a spatial medium, and indeed numerous proposals that tackle the issue of non-linearity seek to compensate for the lack of control on discourse order by exploiting the spatial nature of hypertext. Some have proposed spatial metaphors as a way of describing discourse structure (Landow, 1991; Bolter, 1991; Kolb, 1997); others propose the use of maps, schemas, outlines (Carter, 2000) or navigational patterns (Bernstein, 1998) to return to the author’s hands as much control as possible on the way in which discourse takes shape before the reader’s eyes and coheres in their mind. But hypertext is also a temporal medium, in which spatial structures have a temporal dimension and realisation (Luesebrink, 1998). So, both space and time can be exploited to express discourse coherence and - we contend - in hypertext the notion

of abstract document structure consists of both spatial and temporal configurations working in a three-dimensional space.

### **5. From abstract to concrete hypertext document structure**

If coherence is a cognitive phenomenon, then it should be possible to express coherence relations not only through textual markers, but also through visual patterns. And if this can be done by using spatial abstract features in linear documents, then it should also be possible using spatial and temporal abstract features in non-linear documents. In particular, we propose that graphics and animation could be used to express discourse coherence in hypertext (Mancini and Buckingham Shum, 2004).

At present, most hypertexts (especially on the web) make no use of graphical features to signal rhetorical relations between nodes (Miles-Board et al. 2002), and nodes often consist of long text pages with a few links targeting other pages, from where the source page can no longer be seen. However, we envision that the non-linear medium could be used in a far more expressive and articulated way, if graphic features were exploited as discourse markers to support coherence. Our work aims specifically at identifying visual devices that can play the role of discourse markers in the non-linear, three dimensional space of hypertext.

One of these devices could consist of creating much smaller hypertext nodes and using the screen as a visual field across which they can configure themselves, as links are clicked and new nodes appear, composing meaningful patterns. The appearance and distribution of the nodes should signify the rhetorical role that their content plays within the immediate context in which the reader comes across them. Therefore, each node should have as many renderings as the relations it holds with other nodes and, on each reading path, its appearance should be determined by its relation to the node that precedes, first, and to the node that follows, then. To achieve that, rhetorical relations could be used as document structuring principles during discourse construction to define hypertext links. These could then be dynamically rendered during navigation through the consistent and concurrent use of the medium's spatial and temporal graphic features.

In this respect, having established a parallel between textual and visual processing, based on the correspondence between fundamental principles of textual and visual cognition (Riley and Parker, 1998), some have derived from Gestalt theory useful guidelines for document design (Campbell, 1995). In particular, *similarity*, *proximity*, *size* and *symmetry* define cohesion in visual space-temporal configurations. For instance, the more similar and closer the elements of a configuration, the more likely they are to be perceived as a unit; the more equivalent in size and symmetrical two configuration, the more likely they are to be perceived as related (whether by SIMILARITY or CONTRAST); etc. Furthermore, a number of

representational rules for visually expressing discourse relations between hypertext nodes could be derived from the semiology of graphics, according to which graphic variables can be employed to express conceptual relationships of *similarity*, *difference*, *order* and *proportion* exploiting the properties of the visual image in a three-dimensional dynamic space (Koch, 2001). Following Gestalt principles and graphic rules (see Mancini, 2005, for a detailed discussion), we have designed and begun testing a series of prototype visual patterns expressing coherence relations in non-linear discourse.

## 6. Visualising discourse patterns

The preparation of the graphical renderings of the relations involved three steps. Firstly, we selected a set of relations for experimental rendering and evaluation. Secondly, we selected a subset of static and dynamic graphical variables to be used according to Gestalt principles and graphic norms to visually render the cognitive coherence relations of the selected subset. Finally, for each selected cognitive relation, we implemented a small animation of text boxes, in which the connection holding between the text chunks contained in the boxes was rendered through a dynamic visual pattern.

### 6.1. *Selecting an experimental set of relations*

While it needed to be representative of the most frequent discourse relations, the relational set also had to be small enough to ensure that the respective renderings could be clearly differentiated, thus minimising confusability. For the same reason, the relations also needed to be close to their primitive form, to facilitate the rendering process. Finally, it was desirable that the set be based on those coherence relations that are better understood and more solidly established in the study of discourse coherence. Given this, we selected the following set, based on established cognitive parameterisations of coherence relations (see Sanders et al., 1993; Pander Maat, 1999; Louwerse, 2001).

**CAUSALITY:** holding between the propositional content A of a discourse part and the propositional content B of another discourse part, when A is presented as causing B.

**CONDITIONALITY:** the hypothetical form of CAUSALITY, holding between A and B, when A is presented as causing B, but only if A holds in the first place.

**CONJUNCTION:** holding between the propositional content A of a discourse segment and the propositional content B of another discourse segment, when A is presented as simply coexisting with B.

DISJUNCTION: the negative of CONJUNCTION, holding between A and B, when A is presented as being alternative to B.

SIMILARITY: holding between the propositional content A of a discourse segment and the propositional content B of another discourse segment, when A is presented as being similar or equivalent to B in some relevant respect.

CONTRAST: the negative of SIMILARITY, holding between A and B, when A is presented as being opposed or unequal to B in some relevant respect.

ELABORATION: holding between the propositional content A of a discourse segment and the propositional content B of another discourse segment, when B is presented as expanding, explaining A.

BACKGROUND: holding between the propositional content A of a discourse part and the propositional content B of another discourse part, when B is presented as the explanatory context in which A exists or occurs.

The graphical renderings of the relations were designed based on their cognitive parameterisation (see Sanders et al., 1993; Louwse, 2001). In our set, the bipolar cognitive parameters defining the relations were: *basic operation*, according to which a relation can be *causal* (when an implication relation can be deduced between two text segments) or *additive* (when a conjunctive relation holds between the segments); *source of coherence*, according to which a relation can be *semantic* (when the two discourse segments are related on the basis of their propositional content) or *pragmatic* (when the two segments are related on the basis of their argumentative or rhetorical function); *polarity*, according to which a relation can be *positive* (when the content of the two related segments consistently express the same basic operation) or *negative* (when the content of one of the two segments defies the rule of the basic operation expressed by the other segment); *hypotheticality*, according to which a causal relation can be *actual* (when the causal connection holds in actuality) or *hypothetical* (when the causal connection holds under certain conditions); *comparativeness*, according to which an additive relation can be *conjunctive* (when two situations are related on the basis of their joint relevance with respect to a whole) or *comparative* (when two situations are related on the basis of their similarity or contrast: see Pander Maat, 1999). Table 1 provides the parameterised description of each relation in the set.

To produce the graphical renderings of the relations, the values of each cognitive parameter defining them were rendered through graphical features. As a result, each relation was visually defined by the sum of the graphical features rendering the cognitive values that define it. So, for instance, the graphic representation of CAUSALITY was defined by the features rendering the values *causal* and *positive*; the representation of DISJUNCTION was defined by the features rendering the values *additive* and *negative*; the representation of SIMILARITY was defined by the features rendering the values *additive*, *positive* and *comparative*; etc. The renderings of the values are later described in Table 3.



Relation	Source of coherence	Basic operation	Polarity	Hypotheticality	Comparativeness
CONJUNCTION	semantic	additive	positive	actual	conjunctive
SIMILARITY	semantic	additive	positive	actual	comparative
DISJUNCTION	semantic	additive	negative	actual	conjunctive
CONTRAST	semantic	additive	negative	actual	comparative
CAUSALITY	semantic	causal	positive	actual	conjunctive
CONDITIONALITY	semantic	causal	positive	hypothetical	conjunctive
ELABORATION	pragmatic	additive	positive	actual	conjunctive
BACKGROUND	pragmatic	causal	positive	actual	conjunctive

**Table 1.** *Parameterised description of the relations: each parametrical value was attributed graphical features in the relational renderings*

## 6.2. Selecting a set of graphical variables

Variable	Variable's properties	Use in the renderings
<i>Colour value</i>	Expresses order in depth	Gradual change of value was used to express directionality and progress in a chain of events (CAUSALITY, CONDITIONALITY) or in discourse progress (ELABORATION, BACKGROUND). Radical change of value was not used to express order, but to achieve an effect of SIMILARITY (when the colour value of a text window changes to that of the other) or CONTRAST (when the colour value of a text window changes to the opposite of that of the other) in comparative relations.
<i>Distribution</i>	Expresses relationships	The final positioning along a vertical axis was used to express the order of events in the world being described (CAUSALITY, CONDITIONALITY, partially ELABORATION), and the final positioning along a horizontal axis was used to express order in discourse (CONJUNCTION, DISJUNCTION, SIMILARITY, CONTRAST, partly ELABORATION, BACKGROUND).
<i>Trajectory</i>	Expresses provenance	The trajectory of objects placing themselves to position was used to express the provenance of events (CAUSALITY and CONDITIONALITY) or concepts (SIMILARITY and CONTRAST).
<i>Overlapping</i>	Expresses order and dependency	This was used to express the sense of enclosure of one object within one behind it (CONDITIONALITY and BACKGROUND), and to express a sense of difference in discourse level (ELABORATION and BACKGROUND).

**Table 2.** *Graphic variables that were used to design the relational renderings*

To maximise the difference between renderings, we made them as simple as possible, using a minimum number of graphical features to express the values of the parameters defining the relations, in a visually consistent fashion. Of the full set of graphical variables that we could have possibly used (see Koch, 2001), we selected: *distribution* (which can express relationships, emphasising either similarity or difference), *colour value* (which can express order in space), *overlapping* (which can express order in space, and also interaction or dependency) and *trajectory* (which can express provenance). More easily than others, these variables can be used concurrently to produce effects of visual cohesion, establishing visual relations of similarity, difference and order between objects. We excluded variables such as

*texture, colour, orientation, shape*, etc., because they are less basic and their use may easily interfere with the effect produced by the more basic ones, causing representational inconsistencies, whereas we wanted to obtain as simple and visually consistent configurations as possible. For example, if the same two objects have different colour value, this difference suggests an order in depth, which in turn produces an effect of subordination of one object with respect to the other. However, if the objects also have different colours, comparing the difference between their colour value and recognising the subordination of one to the other becomes difficult, because different colours have different intensities. So, for all renderings, we chose to use the same colour (grey) while using colour value (on a grey scale) to express order, subordination or even contrast (for a more detailed discussion on the selection of variables, see Mancini, 2005). Table 2 shows which variables were used to design the renderings and with what function, based on their properties.

### 6.3. Designing the relational set

Parameter	Value	Rendering of parameter's value
Source of Coherence	Semantic	Positioning of the objects one next to the other. Equal length of the objects' sides that find themselves next to each other.
	Pragmatic	Overlapping of objects on one of the sides (ELABORATION) or completely (BACKGROUND).
Basic Operation	Additive	Alignment of objects along the horizontal axis (except in ELABORATION). Use of the same value throughout or at the initial stage (except in SIMILARITY, CONTRAST, ELABORATION). Appearance of the second object next to the first object (except in SIMILARITY and CONTRAST) or overlapped to it (ELABORATION).
	Causal	Alignment along the vertical axis (except in BACKGROUND). Gradual intensification of value from one state of events to the other. Sliding down of the second/third object from behind the first/second object (except in BACKGROUND).
Polarity	Positive	Intensification or stability of value, from the appearance of one object to the appearance of the other (except in ELABORATION).
	Negative	Change of the value of the object that was first in the visual field to a value that contrasts the value of the object that appears second.
Hypotheticality	Hypothetical	Complete or partial enclosure of an object (containing the text that refers to the consequence) within the object behind it (containing the text that refers to the pre-existing condition).
	Non-Hypoth.	(See rendering of CAUSALITY)
Comparativeness	Comparative	Radical change of value of the object that is already in the visual field to contrast or match the value of the object coming into the visual field second. Entering of the second object from the side of the visual field opposite to where the first object is; sliding of the second object towards the first and positioning next to it.
	Non-Comp.	(See renderings of CONJUNCTION and DISJUNCTION)

**Table 3.** Graphical rendering of the parametrical values defining our relational set

To produce the renderings, we used examples of argumentative passages from a text on the history of science, whose content and literary style were both accessible and uncontroversial, to facilitate the interpretation of the logical connections in the text. For contextual consistency, we selected text sections about a particular subject (theories of the orbiting of planets in the solar system). From the relevant sections, we then isolated short passages of text, each passage consisting of a pair or triple of sentences holding with each other one of the selected relations. We removed all cue phrases originally connecting the sentences and made the sentences syntactically and semantically independent. We distributed each pair or triple of related sentences respectively within a pair or triple of related text boxes, which were attributed graphical features and animated in order to visually render the relation originally holding between the text sentences. As previously mentioned each rendering was defined by the parametrical values defining each relation. Table 3 shows how each parametrical value was rendered.

While the text boxes were attributed graphical features, the text itself was not: font, size, style and colour of text were treated as constants, except in one case (in the rendering of CONTRAST, to maintain readability, the very light grey of the text used by default turned into a very dark grey, when the medium grey of the text window turned into a much lighter grey).

The detailed descriptions of each relation rendering are listed below.

**CONJUNCTION:** denotes the presence or appearance of two entities or phenomena at the same time in the same space. That is, whenever two entities or phenomena are recorded in such circumstances, they are connected by CONJUNCTION. The relation does not say anything about the reasons why two entities or phenomena coexist or the modalities of their co-presence. As far as the specific context of their occurrence is concerned, they play an equivalent and complementary role in constituting a whole. In this respect, they are also similar in terms of the role that they play, that is, of the importance that they have in the general picture.

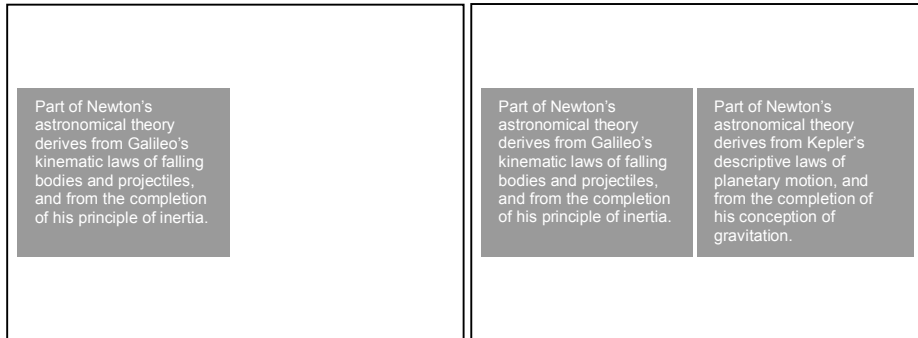
In this specific case, the additive relation was reified by the text spans:

**A.** *Part of Newton's astronomical theory derives from Galileo's kinematic laws of falling bodies and projectiles, and from the completion of his principle of inertia.*

**B.** *Part of Newton's astronomical theory derives from Kepler's descriptive laws of planetary motion, and from the completion of his conception of gravitation.*

They were rendered as shown in Figure 1. The two respective text windows have the same value and their vertical sides have the same length; they appear on the screen next to each other, one at a time, the window containing the first text span appearing on the left and the window containing the second text span appearing on the right after 2 seconds. Firstly, the concept of *addition* is rendered by the windows appearing next to each other, with the order of appearance following the direction of reading familiar in the Western world. Secondly, the concept of *equivalence*, and *similarity* is rendered by the value of the windows' areas, and is reinforced by the

fact that their vertical sides are of identical length, and they appear next to each other and not, say, one under the other. The way in which the windows position themselves is the simplest possible one, to render the fact that the two entities are simply related as complementary components of a whole.



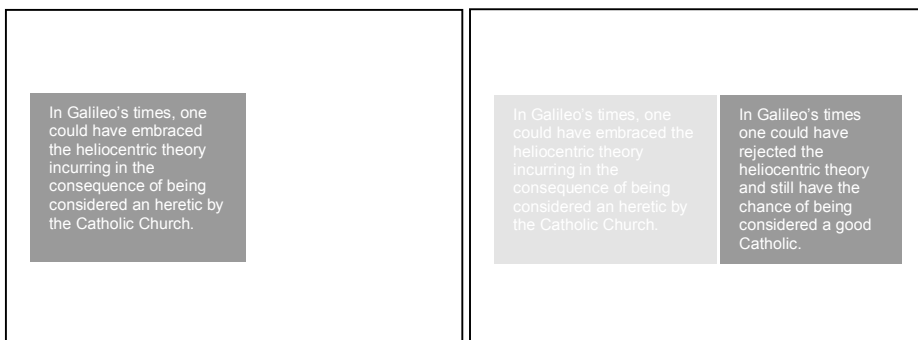
**Figure 1.** *Two screen shots from the animated graphic rendering of CONJUNCTION*

**DISJUNCTION:** the negative of a conjunctive relation is a relation that fails to meet the expectation of CONJUNCTION, or else defies the rule set by the basic operation. Two entities or phenomena do not coexist in a space-time interval, but are alternative to one another, that is, exclude each other. The relation obviously implies their actual existence, but it also implies that this can only be at different times, in different places, or in different circumstances.

The text spans selected to reify DISJUNCTION were:

**A.** *In Galileo's times, one could have embraced the heliocentric theory incurring the consequence of being considered a heretic by the Catholic Church.*

**B.** *In Galileo's times, one could have rejected the heliocentric theory and still have the chance of being considered a good Catholic.*



**Figure 2 -** *Two screen shots from the animated graphic rendering of DISJUNCTION*

They were rendered as shown in Figure 2. The text windows have the same appearance as those used to represent the additive relation, with the difference that

when the second window appears on the right, 2 seconds later, the value of the background of the window on the left changes to a very light grey, which makes it difficult to read the text. In other words, the concept of alternative, of reciprocal exclusion of the two situations, the defeat of the expectation of additiveness expressed by this negative relation, are rendered through the fact that, as the second span of text appears, the first one becomes unreadable.

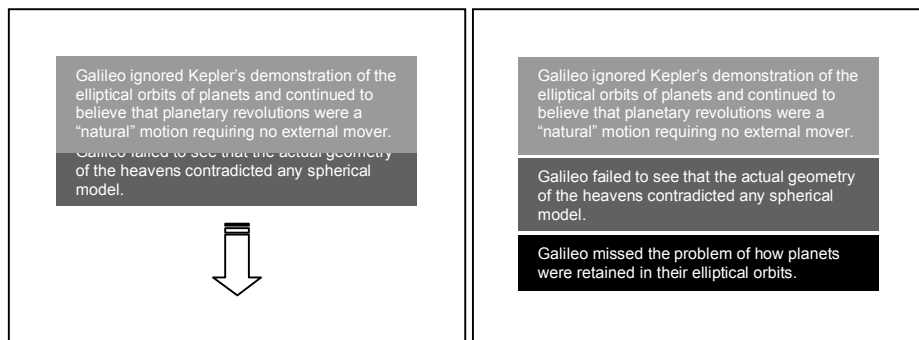
**CAUSALITY:** on the other side of the spectrum with respect to additiveness, it is the strongest logical relation between two entities or phenomena. The causal relation implies additiveness, in that the two entities or phenomena connected are part of the same picture, context, or situation. It implies sequentiality, that is, order, in that one entity or phenomenon necessarily follows the entity or phenomenon that has caused it. It implies **CONDITIONALITY**, in that the appearance of one entity or phenomenon necessarily conditions the appearance of the other; in fact, the bound is so relevant and the connection so specific, that the first entity or phenomenon is directly generating the second.

The text spans, three this time, selected to reify **CAUSALITY** are:

**A.** *Galileo ignored Kepler's demonstration of the elliptical orbits of planets and continued to believe that planetary revolutions were a "natural" motion requiring no external mover.*

**B.** *Galileo failed to see that the actual geometry of the heavens contradicted any spherical model.*

**C.** *Galileo missed the problem of how planets were retained in their elliptical orbits.*



**Figure 3** - Two screen shots from the animated graphic rendering of **CAUSALITY** (the white arrow in the left shot illustrates the movement of the boxes)

They were rendered as shown in Figure 3. The three windows respectively containing the three text spans are arranged one under the other, the second sliding down from behind the first as soon as the first has appeared, and the third sliding down from behind the second as soon as the second has reached its position. They all share the same width, while their height is determined by the quantity of text

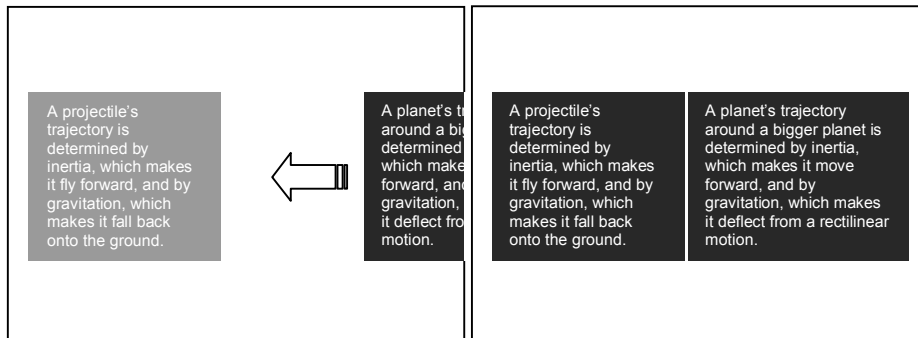
contained in each window. The windows' background becomes increasingly darker from the first to the third, and the ratio of increment is the same from the first to the second and from the second to the third, that is, they are equidistant, as far as the value is concerned. In this configuration, the order of the events is rendered by the arrangement of the text windows, while the fact that the second and the third windows appear by sliding down from the previous one renders the fact that the second and the third events follow, and are brought about, respectively by the first and the second event. At the same time, the darkening of the background renders the idea of ordered progression in a necessary process, from one stage to the other. Finally, the cohesion between the three events is reinforced by the fact that the three windows have the same width.

**SIMILARITY:** expresses a connection between two entities or phenomena that may belong to different semantic worlds and that may not have any logical connection with each other. However, the connection established between them often enlightens their nature, and often reveals certain aspects of one or the other that may have been not as evident before the comparison occurred. This relation does not imply chronological order, and between the connected entities or phenomena there is no hierarchy, but rather equivalence, that is, they are independent objects connected on the grounds of what they happen to have in common.

The text spans selected to reify SIMILARITY are:

**A.** *A projectile's trajectory is determined by inertia, which makes it fly forward, and by gravitation, which makes it fall back onto the ground.*

**B.** *A planet's trajectory around a bigger planet is determined by inertia, which makes it move forward, and by gravitation, which makes it deflect from a rectilinear motion.*



**Figure 4** - Two screen shots from the animated graphic rendering of SIMILARITY (the white arrow illustrates the movement of the box)

They were rendered as shown in Figure 4. The two corresponding windows are arranged to end up next to each other, with the left-hand one first appearing on the

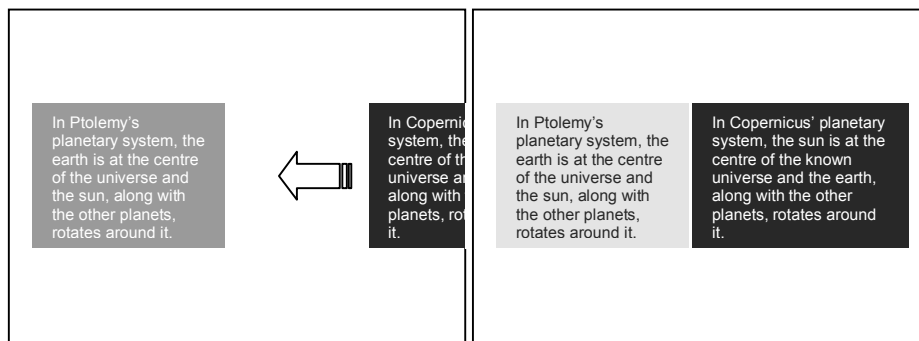
screen, followed by the second one entering the screen from the centre-right and sliding into place next to the first. They have the same height, while their width is determined by the quantity of text contained in each one. The window containing the left text span (the one appearing first) has the default grey background, whereas the window containing the second text span (the one sliding in) has a very dark grey background. However, as soon as the second window reaches the first one, the background of the first one turns into the same very dark grey. In this configuration, the assimilation of the phenomenon described in the first text span to the phenomenon described into the second text span is rendered by the change of background to which the first window is subject. The fact that the second window slides in refers to the “coming together” of different semantic worlds on the basis of a structural analogy between them; while the fact that the second window comes from a side and takes position next to the first window (and not below or above) refers to the fact that there is no subordination between them. Finally, the cohesion between the two objects is reinforced by their identical height.

**CONTRAST:** is the negative of the SIMILARITY relation, as it fails to meet the expectation of similarity, or else defies the rule set by the positive SIMILARITY relation. That is, it is a relation that connects two entities or phenomena presuming a possible similarity between them on the basis of certain elements or aspects is finally contradicted on the basis of certain other elements or aspects.

The text spans selected to reify CONTRAST are:

**A.** *In Ptolemy’s planetary system, the earth is at the centre of the universe and the sun, along with the other planets, rotates around it.*

**B.** *In Copernicus’ planetary system, the sun is at the centre of the known universe and the earth, along with the other planets, rotates around it.*



**Figure 5** - Two screen shots from the animated graphic rendering of CONTRAST (the white arrow illustrates the movement of the box)

They were rendered as shown in Figure 5. The windows containing the two spans of text above are shaped in the same way as they are in the SIMILARITY relation, with the difference that this time, when the second window reaches the first

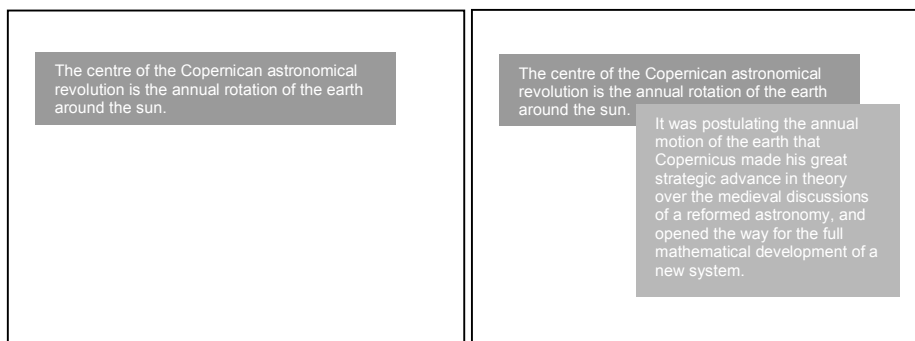
window already in place, the background of the first one turns into a very light grey, which visually produces a great contrast between the two objects (unlike in DISJUNCTION, in this case, and only in this case, the text colour changes, from the standard light grey to a dark grey, to still be readable). In this configuration, the concept of equivalence and potential comparability between the two phenomena described in the text spans is still rendered by the use of graphical variables, however the failure of the comparison is rendered by the contrast of the background values.

**ELABORATION:** connects an element of discourse or a concept to its expansion (in terms of explanation, clarification, or articulation), a deeper level of discourse with respect to the expanded element or concept. For some recipients, that expansion may be superfluous to the understanding of the discourse's structure and development, since they already have the knowledge that the elaboration is meant to provide, but for other recipients it may be useful or necessary. In some cases, especially in hypertext, ELABORATION may constitute the main connection through which an argument develops and explores its conceptual possibilities.

The spans of text selected to reify ELABORATION are:

**A.** *The centre of the Copernican astronomical revolution is the annual rotation of the earth around the sun.*

**B.** *It was in postulating the annual motion of the earth that Copernicus made his great strategic advance in theory over the medieval discussions of a reformed astronomy, and opened the way for the full mathematical development of a new system.*



**Figure 6** - Two screen shots from the animated graphic rendering of ELABORATION

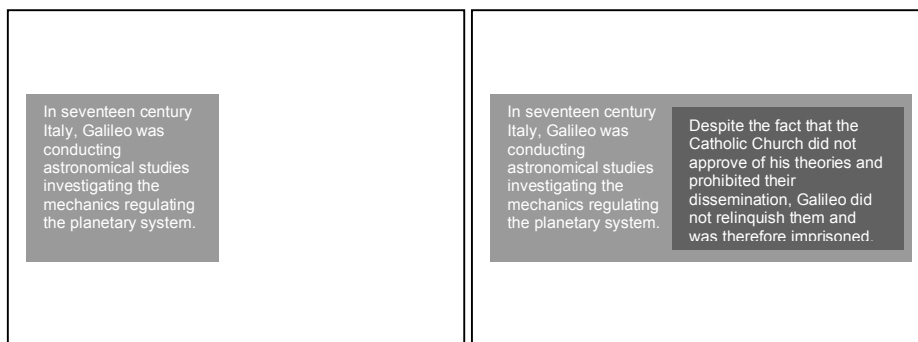
They were rendered as shown in Figure 6. The two windows containing the spans of text above are this time overlapping, the second one appearing over the first one slightly overlapped to its edges, in a way that the text of the first one can still be read, though. In addition, the window containing the first span of text is wider but lower, whereas the window containing the second span of text is about one third narrower and about two thirds taller. Also, none of the sides of the two windows are aligned, but the right edge of the second window is more to the right than the right



edge of the first window. The background of the second window is slightly lighter than that of the first one. In this configuration, the fact that the two discourse units do not belong to the same discourse level is rendered by the differences in background and lack of alignment. The slight overlapping of the two objects suggests the existence of different, although interconnected, layers in the visual field, reinforced by the difference in background. Finally, the distribution of the windows suggests that the second one constitutes an appendix to the first, but also a deviation from the main track.

**BACKGROUND:** relates an element of discourse or a concept to its context (in terms of justification for its occurrence or context defining its meaning). At the semantic level, one of the related entities or phenomena provides the context in which the other entity or phenomenon gains its meaning. On the pragmatic or speech-act level, the content of the first discourse part provides the information necessary to understand the content of the second discourse part. As with ELABORATION, the information provided as background may be unnecessary to some recipients, but very useful and even fundamental to others, depending on their knowledge about the entities or phenomena in question. Also this relation is frequently used in hypertext discourse construction.

The spans of text selected to reify BACKGROUND are: **A.** *In Seventeenth Century Italy, Galileo was conducting astronomical studies investigating the mechanics regulating the planetary system.* **B.** *Despite the fact that the Catholic Church did not approve of his theories and prohibited their dissemination, Galileo did not relinquish them and was therefore imprisoned.*



**Figure 7** - Two shots from the animated graphic rendering of BACKGROUND

They were rendered as shown in Figure 7. The window containing the first text span has a default grey background, but the window containing the second text span has a darker grey background. This second window appears over of the first one, or rather, over of an extension of the first one: as the second window appears to the right of the first, the first is extended so that the second ends up included within the first. This way the second window overlaps with the first while all the text of the first one remains readable. In this configuration, the concept of context is rendered by the visual inclusion of one window within the other, and the concept of

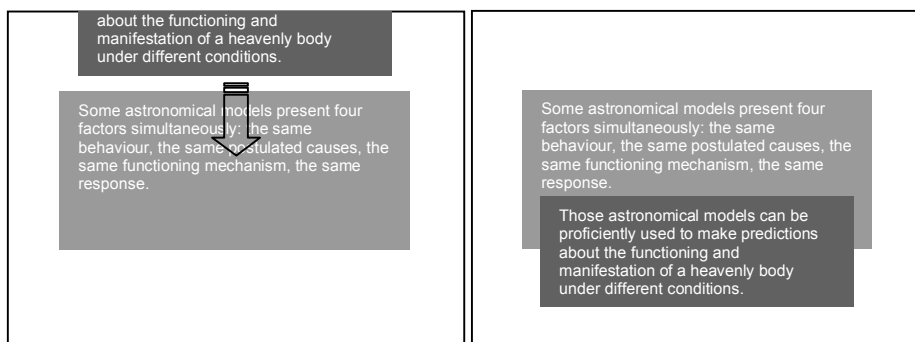
background is suggested by the layering effect of the overlap, reinforced by the change of value.

**CONDITIONALITY:** is sitting between pure CAUSALITY and BACKGROUND. It is similar to a CAUSALITY relation in that the appearance or occurrence of the second entity or phenomenon depends on the appearance or occurrence of the first entity or phenomenon, although the former does not necessarily cause the latter to appear or occur. CONDITIONALITY also shares something in common with the BACKGROUND relation in that the first entity or phenomenon sets the possibility, the context, in which the second entity or phenomena can exist or hold true. At the semantic level, one of the related entities or phenomena provides the context in which the other entity or phenomenon gains its meaning, while at the pragmatic or speech-act level the content of the first discourse part provides the information for the understanding of the content of the second discourse part. In any case, as in CAUSALITY and BACKGROUND, the two entities or phenomena are not equivalent to each other, and as far as the context in which they exist is concerned, their relation is hierarchical.

The text spans related by CONDITIONALITY in this case are:

*A. Some astronomical models present four factors simultaneously: the same behaviour, the same postulated causes, the same functioning mechanism, the same response.*

*B. Those astronomical models can be proficiently used to make predictions about the functioning and manifestation of a heavenly body under different conditions.*



**Figure 8** - Two screen shots from the animated graphic rendering of CONDITIONALITY (the arrow illustrates the movement of the box)

They were rendered as shown in Figure 8. The graphical representation of this relation is something between the visualisation of CAUSALITY and the visualisation of BACKGROUND. The two windows respectively containing the first and the second text span have different width and area: the first one is wider, taller and lighter, whereas the second one is narrower, shorter and darker. The second is sliding from above, but instead of sliding down behind, it slides down over the first one, and stops when still half overlapping it, remaining partly included in it (the text of the

first window still being fully readable). This configuration renders the concept of context through the partial overlapping and inclusion of the second window into the first one. But it also renders the idea of CAUSALITY through the sliding down of the second unit below the first one, and through the darker background of the second window, as a representation of transformation and development of a situation from one stage to the other.

We tested the relational renderings just described to see whether they could be recognised by people who were not aware of the design rationale set out above. Specifically, we wanted to find out if and to what extent the concurrent and consistent use of visual features according to certain perceptual principles and design criteria could produce visual configurations capable of expressing the selected set of relational concepts.

## **7. Evaluating visual discourse patterns: an empirical study**

The study described here constitutes a first verification of our main research hypothesis: that cognitive coherence relations between textual nodes in an argument can be rendered visually, using systematic graphical and animation cues, in such a way that viewers with no training are able to interpret them. While we are in the process of designing further studies to investigate the implications of this proposal for learning and comprehension, this first study focussed specifically on one question: are there visual stereotypes held by viewers that can be exploited to communicate conceptual relationships between textual nodes? If so, then these are prime candidates for rendering coherence relations, and if our relational renderings followed the perceptual principles and graphic criteria that define these stereotypes, then the viewers should be able to consistently identify them among several visual renderings.

### **7.1. *The experiment***

We designed and conducted an empirical study asking people to choose, from three different visual representations, the one that best expressed each relational concept. That is, for each relation, three different representations were presented: the one that had been designed to represent that particular relation, plus two alternative representations designed for the purposes of the experiment (obviously, to create the two alternative representations of each relation, we used the same textual content, giving it a different graphic and animation format). In choosing the alternative renderings to be presented for each relation we tried to associate graphical representations that were visually different enough from one another, in order to avoid dispersion of votes. For instance, associating the CAUSALITY pattern with the CONJUNCTION and DISJUNCTION patterns was intended to make participants'

sensitivity to the visual expression of relational concepts more evident. The associations are summarised in Table 4.

Tested representation	Associated representations	
CAUSALITY	CONJUNCTION	DISJUNCTION
CONDITIONALITY	ELABORATION	CONTRAST
CONJUNCTION	CONTRAST	DISJUNCTION
DISJUNCTION	BACKGROUND	SIMILARITY
SIMILARITY	CONTRAST	CAUSALITY
CONTRAST	CONJUNCTION	SIMILARITY
ELABORATION	SIMILARITY	CONTRAST
BACKGROUND	CONTRAST	SIMILARITY

**Table 4.** List of the relational renderings to be tested and the alternative options associated with them for the experiment

All representations were created in Microsoft PowerPoint, within a single file containing 24 animated slides, that is, 8 groups of three slides corresponding to the 8 relations examined. Before each triple of slides, a white slide only reporting the name of the relation represented in the three following slides was inserted. Each slide of every triple contained the animation of a different relational representation, whose order within the triple itself was random: the representation designed to render the particular relation could find itself in first, second or third position, so that the order of presentation of the renderings could not bias the experimental results. The slide display was controlled by an experiment conductor, ensuring that each pattern be displayed for the same length of time.

Additional material was prepared on which people could record their choices. It consisted of 8 forms, each one devoted to the analysis of a relation and bound to the others in the same order of presentation of the triples of slides in PowerPoint. All the forms were structured in exactly the same way, consisting of three sections (Figure 9). A section at the top of the page provided a brief abstract definition of the relation being examined, to give the participants a clear idea of the relational concept they were asked to focus on. In the section underneath, the abstract relational concept was expressed by the same example used in the animations, with the difference that the cue phrases indicating the relations between the sentences were still in place. Underneath, three pairs of thumbnails were provided, respectively referring to the three animations: for each pair, the thumbnail on the left showed the beginning stage of the corresponding animation, while the thumbnail on the right showed its final state. The order of the pairs from top to bottom followed the order of display of the animations. On the right side of the form, each pair was identified by a letter (A, B or C), below which there was a space to write notes.

Twenty four participants were recruited on a voluntary basis from The Open University. None were specialist hypertext researchers, but they were all computer literate and regular computer users. All participants were tested in the same room, under the same low lights, in front of the same quality screen, and at the same distance from it. They were asked to look at both the definition of the relation and the example on the first form, then watch the three animations presented one after

the other on the screen, and finally mark on the form the option that they preferred by circling the corresponding letter, optionally explaining in an adjacent note why that particular option was preferred over the other two. After seeing the representational options and choosing their favourite representation for the first relation, they did the same for the second relation, and so on through to the eighth. To mark their preferences and write their notes in all the forms, they were given a green pen.







CONDITIONALITY = A is condition for B (if A, then B)		
Two entities or phenomena are related by the fact that in order for the other to be there or happen the first one must be there.		
TEXT EXAMPLE		
If some astronomical models present four factors simultaneously: the same behaviour, the same postulated causes, the same functioning mechanism, the same response, Then those astronomical models can be proficiently used to make predictions about the functioning and manifestation of a heavenly body under different conditions.		
VISUAL REPRESENTATION		
		Do you prefer option A? (please, circle) Can you tell why? (please, write below)
		Do you prefer option B? (please, circle) Can you tell why? (please, write below)
		Do you prefer option C? (please, circle) Can you tell why? (please, write below)

Figure 9. One of the eight forms given to the participants to elicit their preferences

After the completion of this first round, participants were asked to repeat the entire process. This was done to give them the possibility to change decisions made during the first round. They could do so by simply circling a different letter and adding their comments in the appropriate space. To make it possible to track any changes afterwards, the green pens were replaced with red pens, to be used for any corrections or additional comments. We expected participants to want to make changes as we expected that they would need to go through all the relations and see all the representational options before being able to evaluate the relative expressiveness of each option and decide what relation each option expressed most effectively. In other words, we expected that people would attribute meaning to each animation not just in absolute terms, but also in relative terms: they would rate the best one for each relation within the context of the whole set of renderings. This reflects the fact that any language is a system and that therefore the meaning of each sign or pattern is determined contextually. By allowing participants to consider their choices a second time we accounted for this contextual dimension, with the

expectation that this would give us more favourable, but also more realistic, final results.

## 7.2. The experimental results

Table 5 shows the order in which the different rendering options were presented for each relation, Table 6 shows the votes gained and lost by each relation in the first and second round, and Table 7 shows the statistical significance for each result. As shown in Table 6, all the predicted options were by far preferred by the participants and, as shown in Table 7, these results are statistically significant for most of the relations. In other words, this first experiment seems to indicate that people found particularly expressive the visual configurations that had been specifically designed to represent the set of relations. This is corroborated by the fact that most people motivated their choices by expressing the rationale behind their selection or rejection of different options: if they could motivate their choices, it is unlikely that they chose randomly (although we cannot rule out the possibility of post hoc rationalisation as an experimental artefact).

Tested relation	Option A	Option B	Option C
CAUSALITY	conjunction	CAUSALITY	alternative
CONDITIONALITY	CONDITIONALITY	elaboration	contrast
CONJUNCTION	contrast	disjunction	CONJUNCTION
DISJUNCTION	DISJUNCTION	background	similarity
SIMILARITY	contrast	SIMILARITY	causality
CONTRAST	conjunction	similarity	CONTRAST
ELABORATION	similarity	ELABORATION	contrast
BACKGROUND	similarity	contrast	BACKGROUND

**Table 5.** List of the relations with the three representational options proposed for each of them. The capital bold highlights the rendering designed to express the relation being tested.

Table 7 shows that the renderings of CONDITIONALITY and DISJUNCTION did not obtain statistically significant results, but as we discuss below there are good reasons why this could have happened (and the fact that both of them obtained twice as many votes gathered by the alternative options should not be disregarded). For the most part, however, the concurrent and consistent use of graphical elements, to render our set of relational concepts according to certain perceptual principles and design criteria, appears to have produced a set of visual configurations that the participants were able to recognise as representing those relations.

For the six relational renderings whose results are significant, the experimental data suggests that they must be particularly intuitive, since they already gathered a significant number of votes in the first round, which increased in the second round (except for the rendering of CAUSALITY and BACKGROUND, which maintained the same votes, and the rendering of ELABORATION, which lost one vote). The increase of votes in the second round could be explained by the fact that, as we expected, by then people were better able to evaluate the different options provided for one

relation in comparison with the options provided for other relations. It would be reasonable to think that contextualisation played a role in the interpretation of the renderings, but at this stage this is only a hypothesis that should be investigated in further studies.

	Options	Votes in 1 <sup>st</sup> round	Vote changes in 2 <sup>nd</sup> round		Final results
			Gained by designed option and lost by other options	Lost by designed option and gained by other options	
CAUSALITY	A	4	-3	+1	2
	<b>B</b>	<b>19</b>	<b>+4</b>	<b>-1</b>	<b>22</b>
	C	1	-1	+0	0
CONDITION	<b>A</b>	<b>10</b>	<b>+3</b>	<b>-0</b>	<b>13</b>
	B	5	-1	+2	6
	C	9	-5	+1	5
CONJUNCTION	A	6	-3	+0	3
	B	0	-0	+0	0
	<b>C</b>	<b>18</b>	<b>+3</b>	<b>-0</b>	<b>21</b>
DISJUNCTION	<b>A</b>	<b>12</b>	<b>+1</b>	<b>-1</b>	<b>12</b>
	B	5	-1	+1	5
	C	7	-1	+1	7
SIMILARITY	A	2	-1	+0	1
	<b>B</b>	<b>16</b>	<b>+3</b>	<b>-1</b>	<b>18</b>
	C	6	-2	+1	5
CONTRAST	A	4	-0	+0	4
	B	0	-0	+0	0
	<b>C</b>	<b>20</b>	<b>+0</b>	<b>-0</b>	<b>20</b>
ELABORATION	A	1	-0	+2	3
	<b>B</b>	<b>21</b>	<b>+0</b>	<b>-1</b>	<b>20</b>
	C	3	-1	+0	1
BACKGROUND	A	0	-0	+0	0
	B	3	-0	+0	3
	<b>C</b>	<b>21</b>	<b>+0</b>	<b>-0</b>	<b>21</b>

**Table 6.** Numeric results from the experiment. The predicted options, and the votes gained or lost by them in both rounds, are shown in bold in the grey rows.

It is not surprising that the relational renderings that gathered a significant number of votes were more intuitive to design than DISJUNCTION and, especially, CONDITIONALITY. This could be explained with the fact that CAUSALITY and CONJUNCTION, SIMILARITY and CONTRAST, ELABORATION and BACKGROUND present different situations or processes as given, along a linear narrative axis. However, CONDITIONALITY and DISJUNCTION are cognitively more complex relations. Since the former presents a hypothetical situation and the second presents an alternative

situation, their interpretation requires the projection of two different space-time dimensions and narrative axes: one in which the hypothetical or alternative situation verifies and one in which the hypothetical or alternative situation does not verify. This non-linearity makes it difficult for visual languages based on the articulation of space-temporal units, such as graphics, to express non-linear relations, because they lack the power of abstraction that characterises natural language.<sup>1</sup> Further studies, exploring different representations of those relations and testing them on a larger number of participants, might produce more significant results.

Relation	1 <sup>st</sup> round results	Probability of significance	2 <sup>nd</sup> round results	Probability of significance
CAUSALITY	19	$\chi^2 = 23.25$ ( $p < 0.001$ )	22	$\chi^2 = 37$ ( $p < 0.001$ )
CONDITIONALITY	10	$\chi^2 = 1.75$ ( $p > 0.05$ )	13	$\chi^2 = 4.75$ ( $p > 0.05$ )
CONJUNCTION	18	$\chi^2 = 21$ ( $p < 0.001$ )	21	$\chi^2 = 32.25$ ( $p < 0.001$ )
DISJUNCTION	12	$\chi^2 = 3.25$ ( $p > 0.05$ )	12	$\chi^2 = 3.25$ ( $p > 0.05$ )
SIMILARITY	16	$\chi^2 = 13$ ( $p < 0.01$ )	18	$\chi^2 = 19.75$ ( $p < 0.001$ )
CONTRAST	20	$\chi^2 = 28$ ( $p < 0.001$ )	20	$\chi^2 = 28$ ( $p < 0.001$ )
ELABORATION	21	$\chi^2 = 31.75$ ( $p < 0.001$ )	20	$\chi^2 = 27.25$ ( $p < 0.001$ )
BACKGROUND	21	$\chi^2 = 32.25$ ( $p < 0.001$ )	21	$\chi^2 = 32.25$ ( $p < 0.001$ )

**Table 7.** Summarisation of chi squared results for all tested relations (calculated on the first and second round results)

In conclusion, the results so far obtained encourage us to think that discourse relations (at least the most basic ones) can indeed be signalled graphically, and that graphical configurations can indeed act as discourse markers, to support the representation of discourse structure and coherence when textual discourse markers are no longer as effective as they are in linear text. The patterns that we have designed and tested in this first study are not necessarily the best ones, but they appear to represent a good start.

## 8. Applying visual discourse patterns to hypertext: an example

Now let us illustrate an example of how in non-linear text the expression of coherence could be supported by visualising rhetorical patterns. Consider the following text passage:

“Some animals are 'nice' to each other, especially those who live on the edge. [i]

**For example**, vampire bats have been shown to share meals. If a bat fails to find a meal it is often unable to survive until the next evening's hunting. A bat that has fed well, though, has more than enough to survive, and could easily spare some of its meal. So sometimes a full bat will regurgitate some of its meal to another that is starving. [iii]

<sup>1</sup> For example, in cinematic language, also based on the *monstration* ('act of showing': Gardies, 1984) of space-temporal units, the representation of conditional or disjunctive relations requires expedients such as the use of parallel montage showing different alternatives in the development of an action (e.g., the film *Sliding Doors*, Peter Howitt, 1998).



**These animals** are showing behaviour known as 'reciprocal altruism', which simply means that they lend each other favours in the expectation that the favours will be repaid some time in the future. [iii]

[**For example**] A bat which one day might be bloated by a great meal, might on another evening be less lucky and be in need of help itself. By being generous one day at little cost to itself, it might be saved from starvation the next by another bat returning the favour. [iv]

**This process** can be explained with a game called 'Prisoner's Dilemma'. In the game, two suspects have been arrested for a crime and the police question them in separate rooms. The police offer them each a deal. If they don't co-operate with each other (i.e. they give the police evidence that the other person is guilty) then they will be rewarded and the other person will be put away for the crime. If they both fail to co-operate, and give evidence against each other then they will both get locked up (although they will get a lesser sentence), but if they both co-operate with each other by keeping quiet then the police have no evidence and they will eventually both be released. [v]

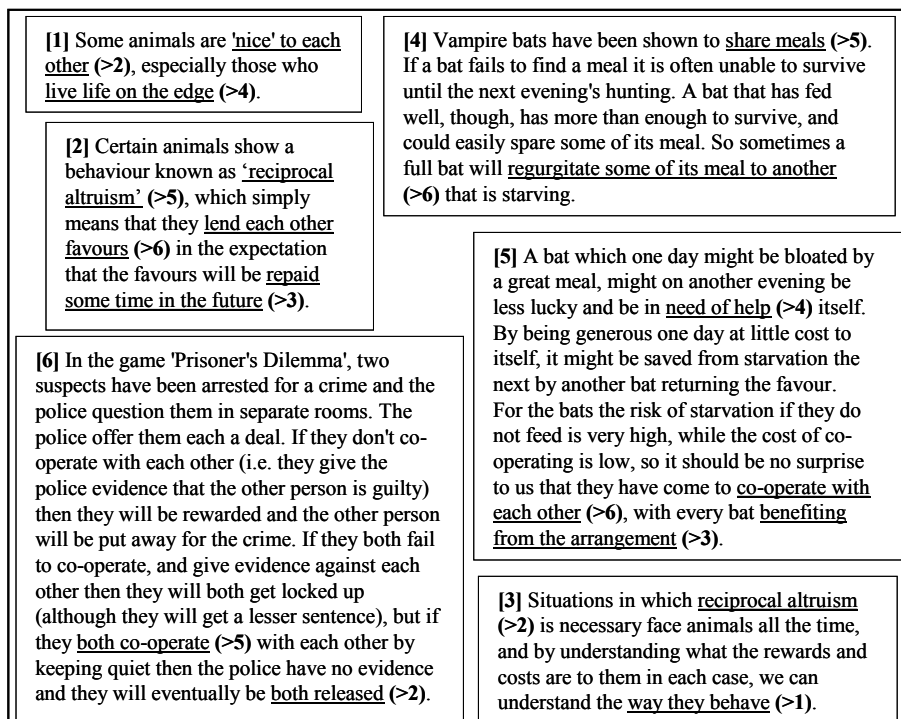
[**Going back to our example**] For the bats the risk of starvation if they do not feed is very high, while the cost of co-operating is low, so it should be no surprise to us that they have come to co-operate with each other, with every bat benefiting from the arrangement. [vi]

**This sort of situation** faces animals all the time, and by understanding what the rewards and costs are to them in each case, we can understand the way they behave. [vii]<sup>2</sup>

This is composed of four paragraphs, each of which is made up of two or three sentences. As far as the content is concerned, three different narrative levels – marked by the indentation of the layout - can be identified, whose relations are expressed by connective or referential phrases (in bold) or simply by paratactic juxtaposition (in bold and square brackets). The author explains an animal behaviour known as *reciprocal altruism*, at one level as an abstract concept, at another level with an example from the animal kingdom, and at yet another level with a metaphor from a game. Now let us consider the case in which the linear text passage is turned into a hypertext (Figure10).

The linear passage is composed of 7 paragraphs (numbered in square brackets). Since they constitute definite discourse units, in the hypertext version each paragraph has become a node (except for [iv] and [vi], which we merged into one node, as the latter constitutes a continuation of the former after the insertion of the comparative paragraph [v]). The nodes are connected via links. Each node (aside from node 1) has at least two incoming links, which means that each node can be accessed at least from two other nodes. Because of that, none of the nodes here contain connectives or referential phrases that relate to other nodes: each one is a self-standing fragment. For each node, each phrase constituting a link is chosen because it summarises the relevant part of the content in the target node (for instance, the link “nice to each other” summarises the content of the target node, which elaborates on the source node by describing the concept of reciprocal altruism). Finally, the rhetorical relations holding between nodes can be expressed through graphic features. Following the rules of graphics, visual attributes can be used consistently and concurrently to render relations of order between nodes in a three-dimensional space, marking the rhetorical relations holding between the discourse parts contained in the nodes.

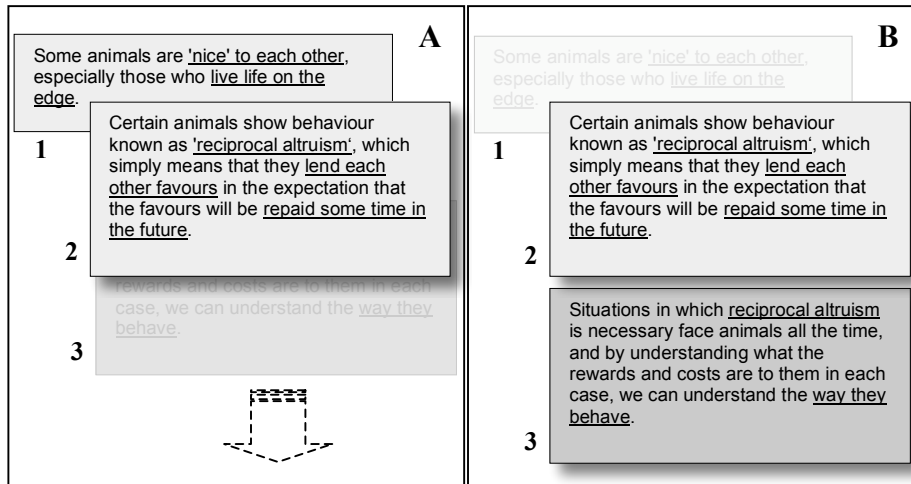
<sup>2</sup> Adapted from the BBC Learning site:  
<http://www.bbc.co.uk/nature/animals/mammals/explore/altruism.shtml>



**Figure 10.** Hypertext version of the linear text passage presented above. The underlined phrases are links. The numbers have illustrative purpose: those in square brackets identify the nodes, those in round brackets refer to the links' target nodes.

Let us hypothesise that one reader follows the path that leads from node 1, to node 2, to node 3, by following first the link 'nice' to each other in node 1 and then the link repaid some time in the future in node 2. Node 1, the starting point in the hypertext, expresses in a nutshell the concept of 'reciprocal altruism', which is the subject of the passage. Node 2 elaborates the concept and, on the basis of that ELABORATION, node 3 comes to a CONCLUSION. At first, node 1 is on the screen on its own, but, when the reader clicks on the link 'nice' to each other, node 2 appears (Figure 11, A).

The relation of ELABORATION holding between nodes 1 and 2 could be expressed as follows: node 2 overlaps on the lower edge of node 1, projecting a small shadow. That is, through the slight overlapping and projected shadow of node 2, this configuration aims to reflect the fact that the two units do not belong to the same discourse level: the first one, higher up and more in depth in the visual field, states the basic concept that the second one, lower and more to the forefront in the visual field, restates and expands. At this point, as the reader clicks on the link repaid some time in the future, node 3 slides down from behind node 2, greyed out at first (Figure 11, A). As it positions itself under node 2, node 3 becomes readable and node 1 greys out instead, leaving the other two both in evidence (Figure 11, B).



**Figure 11.** Hypertext transition in progress (A); hypertext transition completed (B).

The relation holding between the nodes - CONCLUSION - is a pragmatic form of CAUSALITY. This is expressed by the origin and trajectory of node 3, which physically descends from node 2 and by the fact that the background of node 3 has a darker value. Moreover, the fact that node 2 and 3 have the same width and are aligned closely one under the other aims to express the fact that they constitute the interconnected parts of a larger unit. Finally, by the greying out of node 1 the presentation underlines the unity of node 2 and 3.

Now let us hypothesise that another reader follows a different path, going from node 1, to node 6, to node 5, to node 3, by respectively following the links live life on the edge, regurgitate some of it's meal to another, both co-operate and benefiting from the arrangement. This second reading constitutes a different navigational experience, to which corresponds a different visual experience. At first, node 1 is on its own on the screen, but as soon as the reader clicks on the link live life on the edge, node 4 appears (Figure 12, C). The content of node 4 is an exemplification of the concept stated in node 1, and since exemplification is a form of conceptual ELABORATION, the visual relationship between node 1 and 4 is represented in the same way as the visual relationship between node 1 and 2 in the previous path. As the reader now clicks on the link regurgitate some of its meal to another, node 6 enters the screen from the right hand side (Figure 12, C) to position itself right next to node 4 (Figure 12, D). As it gets into place, the background colour value of node 6 turns the same as the background colour of node 4. This is how the conceptual similarity holding between the content of node 4 and the content of node 6 is rendered through a graphic similarity: node 6 moves in towards node 4, it has the same height as node 4, it positions itself next to it and it changes its original background colour (which signals a different domain from which the comparison is drawn) to match that of node 4. As the reader clicks on the link both co-operate, node 5 enters the screen from the left hand side to position itself where node 4 was

before, so that it gets into the same position as node 4 with respect to node 6 (Figure 12, E). And, again, as node 5 gets into place, its original background colour value changes to match the background colour value of node 6. This is a representation of the fact that the same conceptual similarity that holds between nodes 4 and 6 also holds between nodes 6 and 5. Consistently with that, node 5 has the same height as node 4 and ends up in the same position.

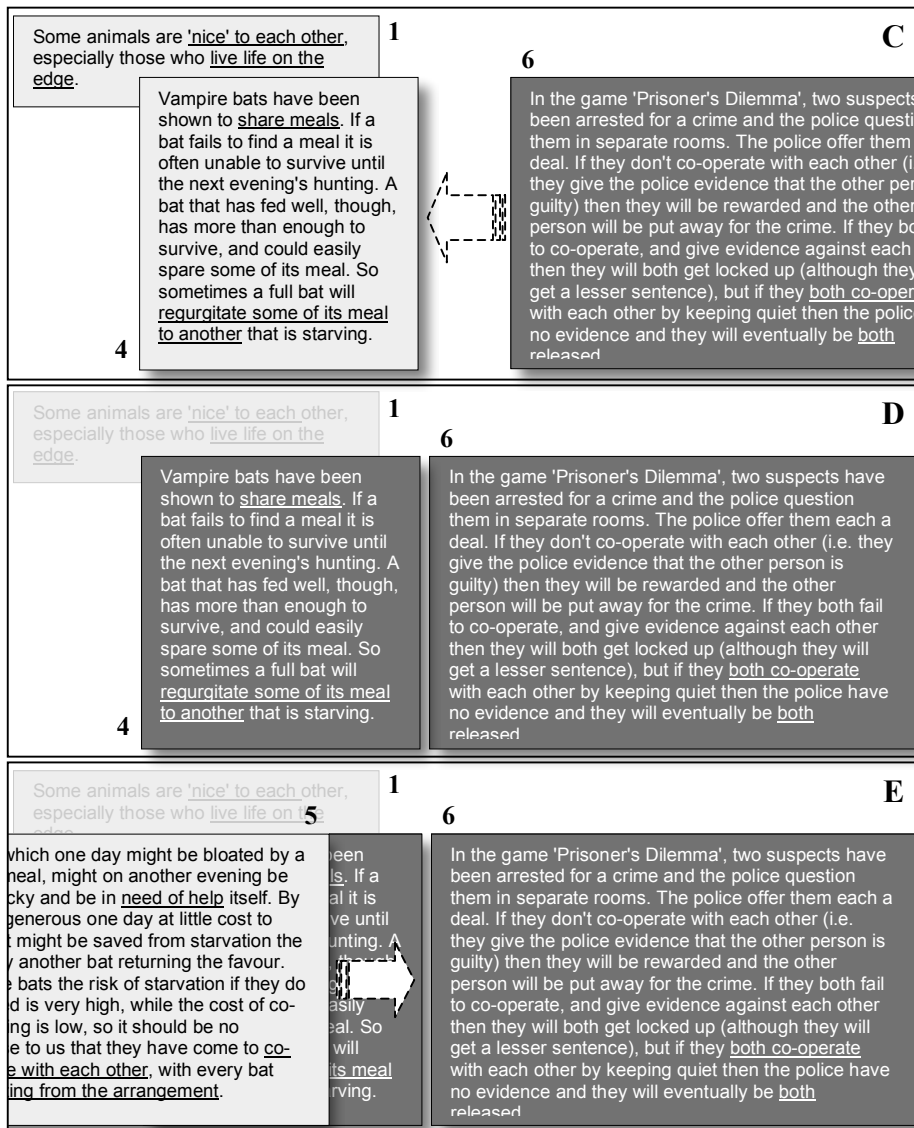


Figure 12. Hypertext transitions: in progress (C); completed (D); in progress (E).  
<http://mcs.open.ac.uk/nlg/vdend/>

## 9. Conclusions and discussion

If a reader is to understand a text, their mental representation of its content must (at least to some degree) reflect the coherence structure intended by the writer. In linear documents, a number of textual devices signalling the coherence structure of discourse facilitate this process of reconstruction. However, these devices only work within a linear structure and they are no longer helpful in the interpretation of non-linear documents. When it comes to non-linear media, such as hypertext, a different set of signalling devices is required, which are visual rather than textual. These visual elements constitute the abstract document structure in traditional text, where they work within the two-dimensional space of the page. However, in hypertext they have to work in a three-dimensional space as well as in time, which expands the boundaries of the notion of abstract document structure.

As we pointed out, there is a fundamental semiotic difference between visual configurations and textual expressions: since it is a symbolic code, text can express relational concepts with precision and subtlety. Although visual languages do not have the same semiotic capabilities of abstraction, there are theoretical grounds and - as reported here - preliminary evidence that they can express at least the most basic relational concepts (for instance, CAUSALITY, CONJUNCTION, SIMILARITY). The conditions under which this was demonstrated are the consistent use of visual properties, combined and animated according to specific rules in order to establish a linguistic context and a local language in which different configurations come to signal different meanings. As reported, we have evidence that within our subject sample, there were stereotypical preferences for particular visual renderings of coherence relations. Of course, the use of visual patterns to express coherence relations in hypertext could be associated with other devices (Kress and van Leeuwen, 2001). For instance, exploiting text generation capabilities, hybrid representational forms could be used, in which symbolic connectives are used in addition as soon as two nodes appear on the screen. However, our aim is to identify ways of presenting hypertext discourse which employ graphical features in a systematic and principled way, extending the notion of abstract document structure, so that it applies to hypertext as well as linear text, by making articulate use of the space-temporal dimensions of the electronic medium, thus more fully exploiting its expressive potential.

Still in its infancy, this work is at this stage more concerned with identifying the right questions than with presenting the right answers. We have not implemented a system yet, but that is our goal, and the experimental results obtained so far are encouraging. As a next step we will be carrying out further tests on the visual renderings of rhetorical relations. For example, we intend to test the same relational renderings with a larger number of participants from different backgrounds, carrying out a qualitative analysis of their responses. We have also started to construct hypertext mock-ups using our set of coherence relations to define the links between nodes and rendering the connections through their corresponding visual patterns.

These are to be tested with users: as they navigate, and visual patterns take shape on the screen, they will be asked to identify the relations holding between nodes, which will be indicated solely by the graphical clues. We also intend to carry out comprehension tests comparing the performance of two groups of users: one group having navigated a graphical hypertext mock-up and one group having navigated a hypertext that has the same content and structure but that does not make use of graphical devices to signal discourse coherence.

Our long-term goal is the application of this work to a larger effort in natural language generation, whereby the same semantic content is rendered differently for different readerships. In particular, we are generating paraphrases that vary not just along the traditional dimensions (discourse, syntax, lexicalisation) but also in terms of graphical presentation (e.g., as textual reports in different styles – including linear vs. non-linear – or as slides for a presentation).

#### 10. Acknowledgements

We are grateful to Marc Eisenstadt for his advice during the early stages of this work, particularly in the analysis of the experimental data.

#### 11. References

- Bernstein M., “Patterns of Hypertext”, *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia*: Pittsburgh, PA, ACM Press, New York, 1998, pp. 21-29.
- Bernstein M., Greco D., “Card Shark and Thepsis: Exotic Tools for Hypertext Narrative”, Wardrip-Fruin N., Harrigan P. (Eds.) *First Person*, MIT Press, Cambridge, MA, 2002.
- Bolter J.D., *Writing Space: The Computer, Hypertext, and the History of Writing*, Eastgate Systems, Cambridge, MA, 1991.
- Bouayad-Agha N., Power R. and Scott D., “Can text structure be incompatible with rhetorical structure?” *Proceedings of the First International Conference on Natural Language Generation (INLG'2000)*, Mitzpe Ramon, Israel, 2000, pp. 194-200.
- Brown G., Yule G., *Discourse Analysis*, Cambridge University Press, New York, 1983.
- Campbell K.S., *Coherence, Continuity, and Cohesion. Theoretical Foundations for Document Design*, Lawrence Erlbaum Associates Publishers, Hillsdale (NJ), Hove (UK), 1995.
- Carter L.M., “Arguments in Hypertext: A Rhetorical Approach”, *Proceedings of the Eleventh ACM Conference on Hypertext and Hypermedia*, ACM Press, New York, 2000, pp. 87-91.
- Dijk van T.A., *Explorations in the Semantics and Pragmatics of Discourse*, Longman, London – NY, 1977.

- Douglas J. Y., "Understanding the Act of Reading: the WOE Beginner's Guide Dissection", *Writing on the Edge*, 2(2), 1991, pp. 112-125.
- Gardies A., "Le Su et le Vu", *Hors Cadre*, 2, 1984, pp. 45-65.
- Grimes J.E., *The Thread of Discourse*, Mouton Publishers, Berlin-NY-Amsterdam, 1975.
- Halliday M.A.K., Hasan R., *Cohesion in English*, Longman, 1976.
- Hobbs J.R., *On the Coherence and Structure of Discourse*, Technical Report CSLI-85-37, Stanford, 1985.
- Johnson-Laird P. N., *Mental models: Towards a cognitive science of language, inference, and consciousness*, Harvard University Press, Cambridge, MA, 1983.
- Kamp H., Ryle U., *From Discourse to Logic*, Kluwer, Dordrecht, 1993.
- Knott A., Dale R., "Using Linguistic Phenomena to Motivate a Set of Coherence Relations", *Discourse Processes*, 18(1), 1994, pp. 35-62.
- Knott A., Mellish C., "A feature-based account of the relations signalled by sentence and clause connectives", *Language and Speech*, 39(2/3), 1996, pp. 142-183.
- Koch W.G., "Jaques Bertin's Theory of Graphics and its Development and Influence on Multimedia Cartography", *Information Design Journal*, 10(1), 2001, pp. 37-43.
- Kolb D., "Scholarly Hypertext: Self-Represented Complexity", *Proceedings of the Eighth ACM Conference on Hypertext and Hypermedia*, ACM Press, New York, 1997, pp. 29-37.
- Kress G., Leeuwen van T., *Multimodal discourse: the modes and media of contemporary communication*, Arnold, London, 2001.
- Landow G.P., "The Rhetoric of Hypermedia: Some Rules for Authors", Delany P., Landow G.P. (Eds) *Hypermedia and Literary Studies*, MIT Press, Cambridge, MA, 1991, pp. 81-104.
- Louwerse M., "An Analytic and Cognitive Parametrization of Coherence Relations", *Cognitive Linguistics*, 12 (3), 2001, pp. 291-315.
- Luesebrink M., "The Moment in Hypertext", *Proceedings of ACM Hypertext '98*, ACM Press, New York, 1998, pp. 106-112.
- Mancini C., *Cinematic Hypertext. Investigating a New Paradigm*, IOS Press, Amsterdam, 2005.
- Mancini C., Buckingham Shum S., "Towards 'Cinematic' Hypertext", *Proceedings of the Fifteenth ACM Conference on Hypertext and Hypermedia*, ACM Press, New York, 2004, pp. 115-124.
- Mann W.C., Thompson S.A., "Rhetorical Structure Theory: Toward a Functional Theory of Text Organisation", *Text*, 8 (3), 1988, pp. 243-281.
- Martin J.R., *English Text. System and Structure*, John Benjamins Publishing Co., Amsterdam, 1992.

- Miles-Board T., Carr L., Hall W., “Looking for Linking: Associative Links on the Web”, *Proceedings of the Thirteenth ACM Conference on Hypertext and Hypermedia*, ACM Press, New York, 2002, pp. 76-77.
- Nunberg G., *The Linguistics of Punctuation*, CSLI, Stanford, US, 1990.
- Pander Maat H., “The Differential Linguistic Realisation of Comparative and Additive Coherence Relations”, *Cognitive Linguistics*, 10, 1999, pp. 147-184.
- Piwiek P., Power R., Scott D., Deemter van K., “Generating multimedia presentations: from plain text to screenplay Intelligent Multimodal Information Presentation”, Stock O. and Zancanaro M. (Eds), *Text Speech and Language Processing*, vol. 27, Kluwer, Dordrecht, 2005, pp. 203-226.
- Power R., Scott D., Bouayad-Agha N., “Document Structure”, *Computational Linguistics*, 29 (4), 2003, pp. 211-260.
- Riley K., Parker F., “Parallels between visual and textual processing”, *IEEE Transactions on Professional Communication*, 41, 1998, pp. 175-185.
- Sanders T.J.M., Spooren W.P.M., Noordman L.G.M., “Coherence Relations in a Cognitive Theory of Discourse Representation”, *Cognitive Linguistics*, 4(2), 1993, pp. 93-133.
- Saussure F., *Cours de Linguistique Générale*, Editions Payot, Paris, 1922.
- Schiffrin D., *Discourse Markers*, Cambridge University Press, New York, 1987.
- Walker J., “Piecing Together and Tearing Apart: Finding the Story in afternoon”, *Proceedings of the Tenth ACM Conference on Hypertext and Hypermedia*, ACM Press, New York, 1999, pp. 111-118.