Visualising discourse coherence in non-linear document generation

Conference or Workshop Item

Visualising Discourse Coherence in Non-Linear Document Generation

Clara Mancini
Centre for Research in Computing
The Open University
Milton Keynes, MK7 6AA, UK
+44(0)1908.652165
C.Mancini@open.ac.uk

Donia Scott
Centre for Research in Computing
The Open University
Milton Keynes, MK7 6AA, UK
+44(0)1908.659032
D.Scott@open.ac.uk

ABSTRACT
In order to produce coherent linear documents, Natural Language Generation (NLG) systems have traditionally exploited the structuring role played by textual discourse markers such as relational and referential phrases. However, these coherence markers of the traditional notion of text do not work as efficiently in non-linear informative documents, where ambiguity and indeterminacy are usually undesirable. A new set of devices, not only textual but graphic, is therefore needed together with formation rules to govern their usage, supported by sound theoretical frameworks. While, in linear documents, graphic devices such as layout and formatting play a complementary role to textual devices in the expression of discourse coherence, we contend that in non-linear documents their role is much more important and needs to be adequately studied. In this paper, we present our work in progress, which explores new possibilities for achieving coherence in the automatic generation of hypertext documents.

Categories and Subject Descriptors
H.1. MODELS AND PRINCIPLES: H.1.1.Systems and Information Theory. H.5. INFORMATION INTERFACES AND PRESENTATION: H.5.1. Multimedia Information Systems; H.5.4. Hypertext/ Hypermedia

General Terms
Design, Languages, Theory

Keywords
Hypertext, Discourse Coherence, Cognitive Coherence Relations, Document Structure, Visual Meta-Discourse

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).

With hypertext, however, 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 (he thinks that) the text is conveying (Sanders and Spooren, 2001). To do so the reader needs to be able to identify the conceptual relations (he thinks to be) holding between the set of discourse elements (whether these are 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 cohesive formal 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 (Halliday and Hasan, 1976; Martin, 1992; Knott and Dale, 1994) or theory driven (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 plaid 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 semicolon. Likewise, in the sentence "*I had a busy morning: I had a work meeting, I went for 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 for shopping*
- *I picked up the children*

Elsewhere (Power et al., 2003) we propose that to account for the varying formulations of a text a separate descriptive layer is required, which we term *abstract document structure*. As we show in previous work (Piwek et al., 2005), 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 adverbials 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 non-motivated: 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 for shopping, I picked up the children.*", in which the clauses that follow the semicolon 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 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 TEXT TO HYPERTEXT ABSTRACT DOCUMENT STRUCTURE

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

At present, most hypertexts (especially on the web) make no use of dynamic graphics to signal rhetorical relations between nodes, 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 think that the non-linear medium could be used in a far more expressive and articulated way, if dynamic graphics was exploited to represent abstract document structure and support discourse coherence. Our work aims 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 distribute, 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 b determined by its relation to the node that precedes, first, and to the nodes 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 designed and begun testing a series of prototype visual patterns expressing coherence relations in non-linear discourse.

6. VISUALIZING RHETORICAL PATTERNS

Based on cognitive parameterisations of coherence relations (Sanders et al., 1993; Pander Maat, 1999; Louwerse, 2001), we selected a set of relations for experimental rendering and evaluation. The set included: CAUSALITY, CONDITIONALITY, SIMILARITY, CONTRAST, CONJUNCTION, DISJUNCTION, ELABORATION and BACKGROUND. For the criteria of selection and for the discussion of all the renderings, see (Mancini, 2005). Here we report on two examples: DISJUNCTION and CAUSALITY.

The graphical renderings of the relations were designed based on their parametrical description. In our descriptions of reference the bipolar parameters defining DISJUNCTION and CAUSALITY were: *basic operation*, according to which a relation can be *causal* or *additive*, and *polarity*, according to which a relation can be *positive* or *negative*. The values of each cognitive parameter defining the relations (Table 1) 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. The representation of DISJUNCTION was defined by the features rendering the values *additive* and *negative*. The representation of CAUSALITY was defined by the features rendering the values *causal* and *positive*. The renderings of the values are described in Table 2.

Relations	Basic Operation	Polarity
DISJUNCTION	additive	negative
CAUSALITY	causal	positive

Table 1. Parametrical description of Disjunction and Causality (Sanders et al., 1993).

To reify the relation renderings, examples of argumentative passages were taken from a history of science text, whose conceptual complexity and literary style were very accessible. Out of all the material provided by the book, a particular subject (theories about the orbiting of planets in the solar system) was selected, so that all the relations would be reified in the text within the same conceptual context. Short passages of text were then isolated, each passage consisting of a pair or a triple of sentences. The sentences of each pair or group held with each other one of the eight selected relations, all signalled by appropriate connectives. Finally, each pair or triple of related sentences was represented on screen respectively within a pair or triple of related text windows, and those windows were attributed certain graphical properties expressing the relation holding between the content of one sentence and the content of the other. On screen, all connective were removed from the text within the windows, and the connective function between the text spans was entirely delegated to the windows' graphical properties.

Parameter	Value	Rendering of each parametrical value
Basic Operation	additive	Windows aligned along horizontal axis. Same value throughout or at initial stage. Second window appearing next to the first or overlapped on one side.
	causal	Windows aligned along vertical axis. Gradual value intensification from one stage to the other. Second and third windows in turns slide down from behind respectively behind first and second.
Polarity	positive	Value intensification or stability, from appearance of one window to appearance of the other.
	negative	Value of the window appearing first in the visual field changes to contrast the value of the object appearing second.

Table 2. Description of the features used to design the parametrical values defining: conjunction, disjunction and causality.

In order to be as differentiated as possible, each representation had to be kept as minimal as possible, making use of no more formal elements than strictly necessary. A small number of graphical variables (Koch, 2001) were used following specific rules of graphics (detailed discussion of the design process for all relational renderings in Mancini, 2005). Below is the description of DISJUNCTION and CAUSALITY.

DISJUNCTION – In this relation two entities or phenomena do not coexist in a space-temporal interval, but are *alternative* or *opposed* to one another. The reasons of their opposition may remain unspecified, but the fact that they represent alternatives expresses the equivalence of their role in the given context. The text spans selected to reify the disjunctive relation 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.*

The two respective text windows were given the same appearance as those used to represent the additive relation, with the difference that as the second window appeared on the right 2 seconds after, the window on the left had the value of its background changed to a very light grey, which made it difficult to read the text. The concept of reciprocal exclusion of the two situations, was rendered through the fact that, as the second span of text appeared, the first one would become unreadable (Figure 1).

CAUSALITY - This is the strongest cognitive relation. It implies conjunction (the connected elements are part of the same context), sequence (one element necessarily follows the other) and the first element directly produces the second. The text spans, three this time, selected to reify causality were:

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.*

The three windows respectively containing the three text spans were arranged one under the other, the second sliding down from behind the first as soon as the first had appeared, and the third sliding down from behind the second as soon as it had reached its position. They all shared the same width, while the height of each was determined by the quantity of text contained in each window. The value of the windows' background became increasingly darker from the first to the third, and the ratio of increment was the same from the first to the second and from the second to the third, that is, they were equidistant, as far as the value was concerned. In this configuration, the order of the events was rendered by the arrangement of the text windows, while the fact that the second and the third windows appeared by sliding down from the previous one rendered the fact that the second and the third events followed, and were brought about, respectively by the first and the second event. At the same time, the darkening of the background rendered the idea of progression in the forging of a logical chain. Finally, the cohesion between the three events was reinforced by the fact that the three windows had the same width (Figure 2).

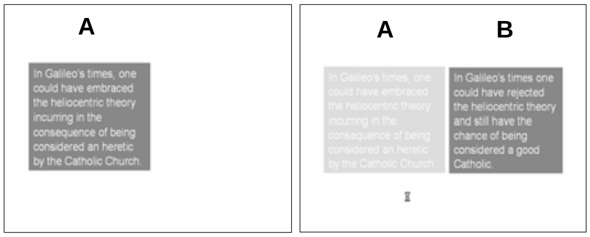

Figure 1. Two screen shots from the animated graphic rendering of disjunction (the letters above the text boxes are for illustration purposes only).

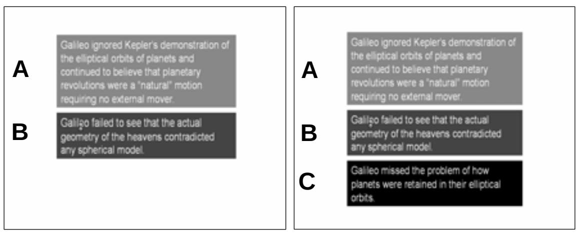

Figure 2. Two screen shots from the animated graphic rendering of causality (the letters beside the text boxes are for illustration purposes only).

The whole set of relations was rendered with the purpose of testing the renderings and their impact on users. In particular we wanted to find out whether the concurrent and consistent use of visual features according to certain perceptual principles and design criteria would determine the expressiveness of the configurations designed to represent the selected sub-set of discourse relations and whether people would discriminate the relational expressiveness of different visual configurations.

As a first form of verification, we designed and conducted an empirical study with a group of **24** participants. We asked them to choose from three different representations the one that in their judgement best expressed each relational concept. For each relation, three different representations were presented to the participants: the one that had been designed to represent that particular relation, plus two alternative representations originally designed to express different relations.

One at the time, the participants were given the original text that had been used to reify each relation, as well as an abstract definition of the relation in question, then were shown the three animations associated with it, from which they had to choose what they thought to be its most expressive representation. They were asked to go through a second round, in which they were allowed to modify, one way or the other, the choices made in the first round.

For each given relation, the great majority of participants converged on the same option, which in fact corresponded to the animated pattern that had been specifically designed to render that particular relation. For 6 of the relations - CONJUNCTION, CAUSALITY, SIMILARITY, CONTRAST, ELABORATION, BACKGROUND - the results were statistically significant (see Table 3).

Relation	1st round votes	Probability of significance	2nd round votes	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 3. Results of the experiment conducted with 24 participants, showing the renderings designed to respectively express each relation. Summarisation of chi squared results for all tested relations (calculated on the first and second round results).

In brief, albeit not conclusive, the results of this first study suggest that people did recognize a particular expressiveness in the options that had been designed to render the subset of discourse coherence relations. In other words, there is positive evidence that the concurrent and consistent use of graphical elements, according to certain perceptual principles and design criteria, can support the visual expression of relational concepts.

The fact that for two of the relations - CONDITIONALITY and DISJUNCTION - the renderings did not obtain the same consensus obtained by the others could be explained with the fact that both conditionality and disjunction are characterized by a greater degree of cognitive complexity. From a cognitive point of view, CAUSALITY, CONJUNCTION, SIMILARITY, CONTRAST, ELABORATION and BACKGROUND hold within a space-temporal continuity, or along one possible line of events. However, conditionality and disjunction hold across two possible lines of events. That is, they implicate the cognitive projection into an alternative space-temporal dimension (or narrative axis), before the conditioned or disjuncted situations can be presented. Such an abstraction is easy to express in natural language, but it is not as easy to express in visual languages.

Evidently, this work is still in progress and we are still exploring ways of presenting hypertext which employ the graphical features of the medium in a systematic and principled way. We have not implemented a system yet, but that is our goal, and the experimental results that we have obtained so far are encouraging.

7. APPLYING VISUAL RHETORICAL PATTERNS TO HYPERTEXT

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

[1] Some animals are 'nice' to each (>2) other, especially those who live life on the edge (>4).

[2] Certain animals show a behaviour known as 'reciprocal altruism' (>5), which simply means that they lend each other favours (>6) in the expectation that the favours will be repaid some time in the future (>3).

[6] In the game 'Prisoner's Dilemma', 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 (>5) with each other by keeping quiet then the police have no evidence and they will eventually be both released (>2).

[4] Vampire bats have been shown to share meals (>5). 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 (>6) that is starving.

[5] 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 (>4) 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.

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 (>6), with every bat benefiting from the arrangement (>3).

[3] Situations in which reciprocal altruism (>2) is necessary face 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 (>1).

Figure 3. Hypertext version of the linear text passage presented above.

"Some animals are 'nice' to each other, especially those who live on the edge.

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.

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.

[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.

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.

[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.

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."[1]

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 (Figure 3).

In the hypertext version, the underlined words or clauses constitute links and the numbers in brackets next to them indicate their target node (nodes are numbered for illustration purposes). Each node has at least two 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, no matter from where it is accessed. If connectives and referential phrases are not used to express the rhetorical relations holding between nodes, however, these relations could be expressed through graphic features. Following the rules of graphics visual attributes could 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.

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,

[1] Adapted from British Broadcasting Corporation (BBC) Learning site: http://www.bbc.co.uk/nature/animals/mammals/explore/altruism.shtml

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 4). 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 4).

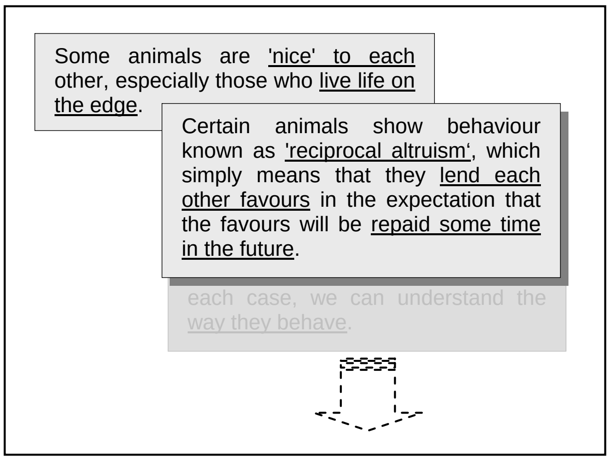

Figure 4. Hypertext transition in progress.

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 5). 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 6). 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 6) to position itself right next to node 4 (Figure 7). As it gets into place, the background colour 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 8). And, again, as node 5 gets into place, its original background colour changes to match the background colour 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.

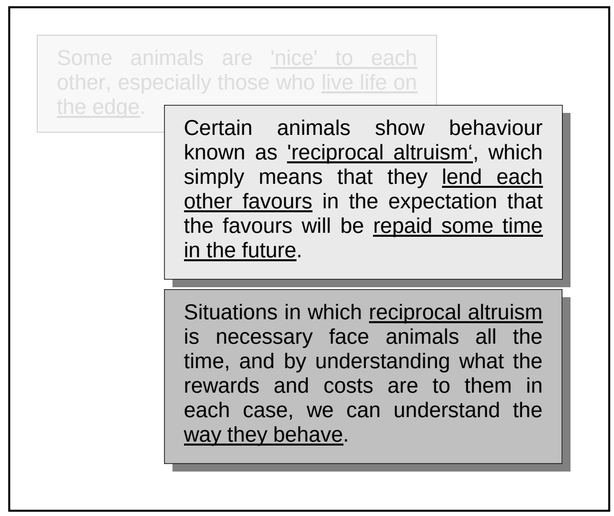

Figure 5. Hypertext transition completed.

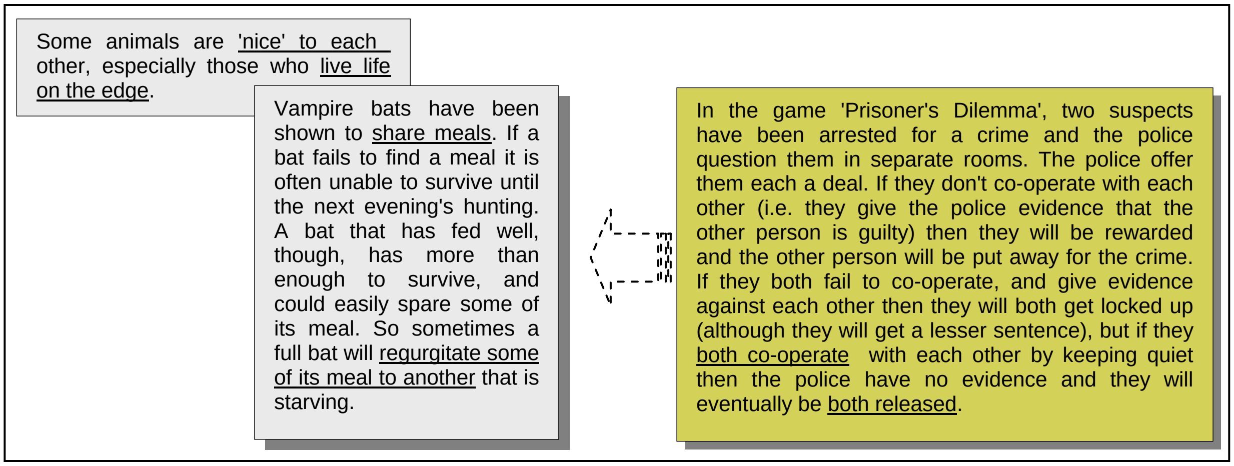

Figure 6. Hypertext transition in progress.

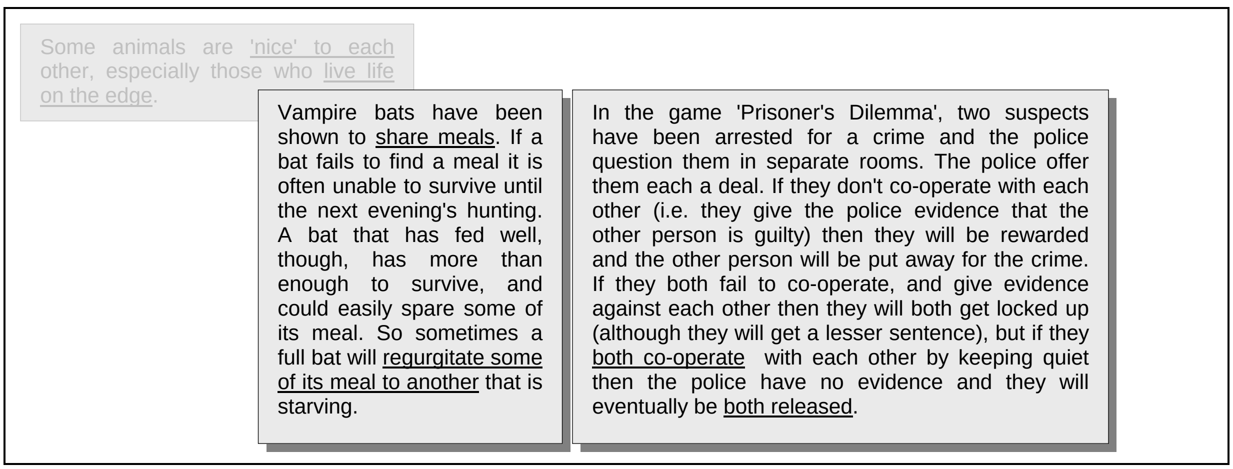

Figure 7. Hypertext transition completed.

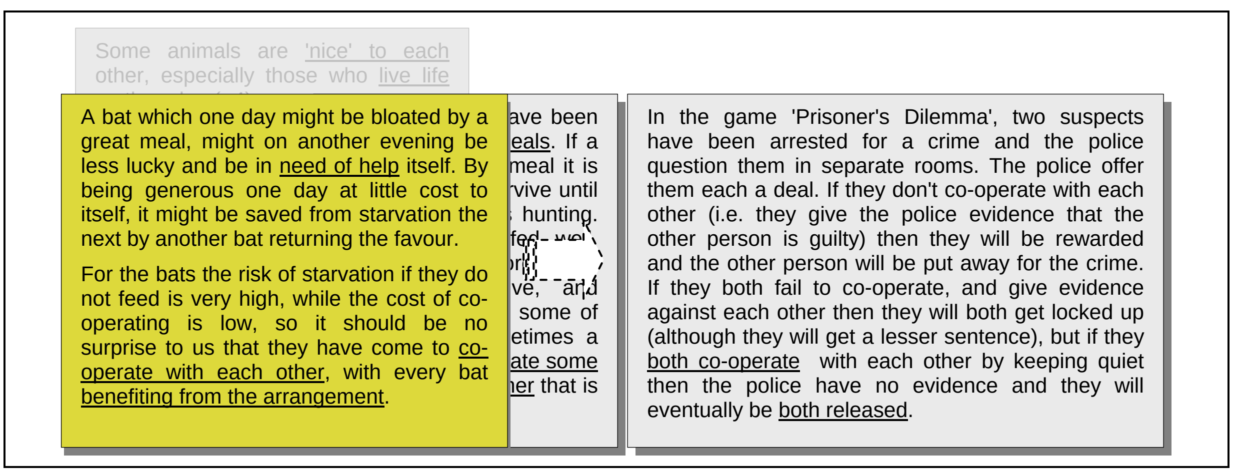

Figure 8. Hypertext transition in progress.

8. CONCLUSIONS

If a reader is to understand a text, their mental representation of its content has to (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 efficiently within a linear structure and they are not as 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. In traditional text, these visual elements work within the bi-dimensional space of the page. However, in hypertext they have to work in a three-dimensional space as well as in time, which pushes 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 is theoretical ground and some preliminary evidence that they can express at least the most basic relational concepts (for instance, CAUSALITY, CONJUNCTION, SIMILARITY). The condition for that is the consistent and concurrent use of the properties of the image according to specific rules, in order to establish a linguistic context in which different configurations become recognisable as having different meaning. 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, 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. Further tests will also be designed.

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).

9. REFERENCES

Bernstein, M. (1998). Patterns of Hypertext. In *Proceedings of ACM Hypertext'98:* Pittsburgh, PA, ACM Press, New York, pp.21-29

Bernstein, M, Greco, D. (2002). Card Shark and Thepsis: Exotic Tools for Hypertext Narrative. In Wardrip-Fruin, N., Harrigan, P. (Eds.) *First Person*, MIT Press, Cambridge, MA

Bertin, J. (1967). *Semiologie Graphique.* English translation (1983). *Semiology of Graphics: Diagrams, Networks, Maps.* University of Wisconsin Press, Madison

Bolter, J.D. (1991). *Writing Space: The Computer, Hypertext, and the History of Writing*, Eastgate Systems, Cambridge MA

Brown, G. and Yule, G. (1983). Discourse Analysis. Cambridge University Press, New York

Campbell, K.S. (1995). *Coherence, Continuity, and Cohesion. Theoretical Foundations for Document Design*. Lawrence Erlbaum Associates Publishers, Hillsdale (NJ), Hove (UK)

Carter, L.M. (2000). Arguments in Hypertext: A Rhetorical Approach. In *Proceedings of ACM Hypertext '00*, ACM Press, New York, pp.87-91

Dijk, van T.A. (1977). Explorations in the Semantics and Pragmatics of Discourse. Longman, London – NY

Douglas, J. Y. (1991). Understanding he Acto of Reading: the WOE Beginner's Guide Dissection. In *Writing on the Edge*, 2(2), pp.112-125

Grimes, J.E. (1975). The Thread of Discourse. Mouton Publishers, Berlin, New York, Amsterdam

Halliday, M.A.K., Hasan, R. (1976). *Cohesion in English*. Longman

Hobbs, J.R. (1985). *On the Coherence and Structure of Discourse*. Technical Report CSLI-85-37, Stanford

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kamp, H. and Ryle, U. (1993). *From Discourse to Logic.* Dordrecht: Kluwer.

Knott, A., Dale, R. (1994). Using Linguistic Phenomena to Motivate a Set of Coherence Relations. In *Discourse Processes*, 18(1), pp.35-62

Knott, A., Mellish, C. (1996). A feature-based account of the relations signalled by sentence and clause connectives. In *Language and Speech*, 39(2/3), pp.142-183

Koch, W.G. (2000/2001). Jaques Bertin's Theory of Graphics and its Development and Influence on Multimedia Cartography. In *Information Design Journal*, 10(1), pp.37-43

Kolb, D. (1997). Scholarly Hypertext: Self-Represented Complexity. In *Proceedings of ACM Hypertext '97*, ACM Press, New York, pp.29-37

Kress, G. and van Leeuwen T. (2001). Multimodal discourse: the modes and media of contemporary communication. London: Arnold.

Landow, G.P. (1991). The Rhetoric of Hypermedia: Some Rules for Authors. In Delany, P., Landow, G.P. (Eds.) *Hypermedia and Literary Studies*, MIT Press, Cambridge, MA, pp.81-104

Louwerse, M. (2001). An Analytic and Cognitive Parametrization of Coherence Relations. In *Cognitive Linguistics*, 12 (3), pp. 291-315

Luesebrink, M. (1998). The Moment in Hypertext. In *Proceedings of ACM Hypertext '98*, ACM Press, pp.106-112

Mancini, C. (2005). Cinematic hypertext. Investigating a new paradigm. Amsterdam: IOS Press.

Mancini, C., Buckingham Shum, S. (2004). Towards Cinematic Hypertext. In *Proceeding of ACM Hypertext '04*, Santa Cruz, CA, USA, Aug 9-13, ACM Press, New York, pp.115-124

Mann, W.C., Thompson, S.A. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organisation. In *Text*, 8 (3), pp.243-281

Martin, J.R. (1992). *English Text. System and Structure*. John Benjamins Publishing Co., Amsterdam

Nunberg, G. (1990). The Linguistics of Ponctuation. CSLI, Stanford, USA

Pander Maat, H. (1999). The Differential Linguistic Realisation of Comparative and Additive Coherence Relations. *Cognitive Linguistics, 10*, 147-184

Piwek, P., Power, R., Scott, D., van Deemter, K. (2005) Generating multimedia presentations: from plain text to screenplay Intelligent Multimodal Information Presentation, Text Speech and Language Processing, vol. 27 pp. 203-226 O. Stock and M. Zancanaro (ed.) Kluwer: Dordrecht

Power, R., Scott, D., Bouayad-Agha, N. (2003). Document Structure. *Computational Linguistics*, vol. 29 issue 4 pp. 211-260

Riley, K., Parker, F. (1998). Parallels between visual and textual processing. *IEEE Transactions on Professional. Communication, 41*, 175-185.

Sanders, T.J.M., Spooren, W.P.M., Noordman, L.G.M. (1993). Coherence Relations in a Cognitive Theory of Discourse Representation. In *Cognitive Linguistics*, 4(2), pp.93-133

Saussure, F. (1922). *Cours de Linguistique Générale*. Editions Payot, Paris.

Schiffrin, D. (1987). Discourse Markers. Cambridge University Press, New York

Walker, J. (1999). Piecing Together and Tearing Apart: Finding the Story in afternoon. In *Proceedings of ACM Hypertext '99*, ACM Press, New York, pp.111-118.