A Formal Model to Support Discourse Semantic Landscape Analysis

Isabelle Linden[a], Bruno Dumas[b] and Anne Wallemacq[c]
Namur Digital Institute (NADI), University of Namur, Belgium

Keywords: Discourse Analysis, Semantic Landscape, Text Formal Model.

Abstract: Discourse Analysis is a broadly spread methodology in human and social sciences. The Evoq Software (Clarinval et al., 2018) has been developed to offer advanced support for a deep semantic analysis of discourse by providing intermediary transposition tools that allow the exploration of the semantic landscape underlying a discourse from multiple angles. This paper presents the formal knowledge model to support these functionalities development and ensure a strong coherence between multiple visualisations seen as so many intermediary transpositions of the same object.

1 INTRODUCTION

Discourse analysis is an essential practice at the heart of the activities of many researchers in the humanities and social sciences. Unfortunately, few software packages offer advanced features for qualitative analysis. Most of the tools available to humanities researchers are limited to quantitative functions. The qualitative approach is often neglected or limited to text manipulation, analysis and dictionary functions (Lejeune, 2010; Lejeune, 2021).

The EFFaTA-MeM (Evocative Framework for Text-Analysis - Mediality Model) research project and the Evoq software have the essential ambition of offering advanced support for a deep semantic analysis of discourse by providing intermediary transposition tools that allow the exploration of the semantic landscape underlying a discourse from multiple angles (Linden et al., 2020).

To ensure the strong coherence of the intermediary transposition, a key question is addressed in this paper: *How to formalise the semantical elements of a discourse to support the extraction, the visualisation and the analysis of its semantic landscape?*

To answer this research question, the paper proceeds as follows. Section 1 presents the theorical background used to support our text analysis approach, namely the structural analysis and introduces the notion of semantic landscape. Then, Section 3 proposes a basic formal model integrating the key element of this theory. This model is extended in Section 4 so as to integrate the elements useful for intermediary transposition. Finally, Section 5 present the transposition integrated in the Evoq Software and explains how they are supported by the formal model.

A short analysis example drawn on text presenting the Transition Network (Transition Network, 2021) serves as running example all along the paper.

2 STRUCTURAL ANALYSIS AND SEMANTIC LANDSCAPE

One of the essential goal of the EFFaTA-MeM project is to conceive ways of interacting and visualising texts that foster new insights in their analysis. This means that we investigate (i) the mediality of texts and of graphical and pictorial representations (Kucher and Kerren, 2015; Gibson et al., 2013), (ii) the intermedial transposition from the textual semiotic system to a pictorial semiotic (Rajewsky, 2002; Rajewsky, 2005; Wolf, 1999; Elleström, 2010; Elleström, 2014), and (iii) the meaning enrichment opportunities offered by this transposition.

The theoretical background used to develop the mediality studied in this research is based on a deep reformulation of the post-structuralist principles. From a mediality point of view, post-structuralism is interesting because it fundamentally questions and reframes the linearity of the text. Post-structuralism opens a very interesting conception of language of which we present here some essential features for our work.

[a] https://orcid.org/0000-0001-8034-1857
[b] https://orcid.org/0000-0001-5302-4303
[c] https://orcid.org/0000-0001-9822-4966

Linden, I., Dumas, B. and Wallemacq, A.
A Formal Model to Support Discourse Semantic Landscape Analysis.
DOI: 10.5220/0011298300003269
In *Proceedings of the 11th International Conference on Data Science, Technology and Applications (DATA 2022)*, pages 118-126
ISBN: 978-989-758-583-8; ISSN: 2184-285X

2.1 A Synchronous View on the Text

While the text is commonly considered as a linear continuous deployment from a beginning to an end, post-structuralism considers it as a field of forces given synchronously. It means that the beginning and the end of the text are considered in the same way and with the same status, just as everything inbetween.

2.2 A Relational Semantics

In this global view on the text, rather than phrases or words, basic units are couples of words in opposition, such as White/Black.

Moreover, each word in the opposition is wrapped up in a set of associative relations or simple evocations which may be explicitly formulated in the text or implicitly supposed by the culture or the context. So for example, in the White/Black opposition White is associated pure and good while Black brings with it implicitly dirty and bad. The meaning of words is therefore larger than the simple denotation. White is far more than a simple colour. Opposed to black it conveys the semantic universe of purity, angel, paradise, untouched.

Let's now consider another text in which white appears in tension with red. Red introduces associations with life, warmth,... to love or anger. In tension with this red, white then becomes the bearer of death, cold, impassive,... radically different from its meaning in the previously considered opposition white/black. These considerations highlight the deeply relational nature of word semantics in the structuralist perspective, which is no longer intrinsic to the word under consideration but depends on its interweaving in the network of words in the text

2.3 The Semantic Fields

The concept of *s*emantic field has been proposed to cover this system of oppositions and associations. Semantic fields are not always explicitly revealed by the words in the text, they are commonly assumed, resulting from the culture or the specific context. In the structuralist approach, this distinction is referred to as the distinction between the level of language and the level of discourse. It implies that an author (writer or speaker) never has a complete mastery of the meaning of her words. Indeed, her discourse is received by the audience integrating these semantic fields that are collectively produced and taken for granted. The semantic field has thus to be considered as the semantic surrounding the text.

Going even further, Derrida (Derrida, 1967) suggests that these semantic fields are not a quiet equilibrium but always in power tension according to the idea that there are dominant relationships between competing semantic fields.

2.4 Structural Analysis

By studying the semantic fields which are underlying the explicit and conscious discourse, structural analysis aims to reveal the balance of power in defining the dominant worldview underlying a text.

3 BASIC CONCEPTUAL MODEL

In the post-structuralist approach as presented in the previous section 2, a text becomes much more than a linear sequence of words. It involves a domain knowledge, implicit representations and tensions which are captured by relations. This section presents the basic formal model developed to capture this approach and formalise a notion of *Analysis project* which extends the notion of text according to the principles of structural analysis.

This formal model plays two key roles in the overall framework of the research:

- first, it is itself a transposition on which artificial intelligence techniques can be applied.
- then, it can be used to ensure the coherence between the multiple proposed intermedial transpositions,

After the introduction of the underlying *Intuitions* subsection 3.1, subsection 3.2 introduces the notation of the basic model and subsection 3.3 formalises the model transformations reflecting the steps of an analysis process.

3.1 Intuitions

The synchronous nature of the approach adopted allows the consideration of the analysis as a single object without any temporal aspect. As the main focus is on relationships, a key element of our model consists of a set of relationships between objects that, for sake of simplicity we call *words*. Actually, there are three different types of words depending of the object of interest for the analyst. The first ones are single literal expression (as e.g. a brand name or ”Transition” in the context of our illustration). The second ones are equivalence classes on the natural language dictionary for which a root represent itself and all its derivative.

And, the last ones are concepts represented by expression involving several words (as e.g. "collective intelligence"). In our model, we denote by *Word* the set of all these possible values.

Relations between these words can be of two different natures: associations denoting a proximity (whatever its nature), and oppositions that materialise the tensions of the semantic field. The description of the approach does not suggest any orientation of the relations, consequently they are modelled by symmetrical relations.

3.2 Notations

The basic elements of our model are build from three sets: *Text* which denotes the texts that can be analysed, *Word* introduced above, and *BasicRel* the set of relations, formalised as $(Word \times Word \times Boolean)$. With these notations, the basic model of an analysis is a triple

$$(t, wl, rl) \in Text \times P(Word) \times P(BasicRel)$$

where

- t is the analysed text,
- wl is the set of the words of interest,
- rl is the set of the relations,
- all the word appearing in rl are in wl[1]
- $(w_1, w_2, b) \in rl$ iff $(w_2, w_1, b) \in rl$.

Note that this model integrates associations and oppositions in one single set. They can be retrieved as $\{(w_1, w_2) : (w_1, w_2, b)\}$ with $b = 0, 1$ respectively.

3.3 Operations

With this notation, drawing an analysis consists in beginning with $(t, \emptyset, \emptyset)$ and integrating information step by step by either integrating a new word w or a new relation $r = (w_1, w_2, b)$ in the analysis (t, wl, rl). The addition of a word is formalised by the following operation.

$$addWord : ((t, wl, rl), w) \to (t, wl \cup \{w\}, rl)$$

The addition of a relation is a partial function only defined provided none of $(w_1, w_2, 0)$ or $(w_1, w_2, 1)$ is in rl:

$$addRel : ((t, wl, rl), (w_1, w_2, b)) \rightharpoonup (t, wl \cup \{w_1, w_2\}, rl \cup \{(w_1, w_2, b), (w_2, w_1, b)\}$$

[1] but not mandatory in t

The operations set involves also possibility for correction in an analysis, namely, removing a term or a relation.

$$removeWord : ((t, wl, rl), w) \to (t, wl \setminus \{w\}, rl \setminus \{(w_1, w_2, b) : w_1 = w \vee w_2 = w\}$$

$$removeRel : ((t, wl, rl), (w_1, w_2, b)) \to (t, wl, rl \setminus \{(w_1, w_2, b), (w_2, w_1, b)\})$$

Note that to preserve the constraints on the inclusion of words appearing in the relations in the set of words, removing a term implies the withdrawal of all the relations in which it is involved.

4 INTERMEDIAL CONCEPTUAL MODEL

In an unpublished study of the first paragraph of a description of the Transition (Transition Network, 2021), the Words dictionary made of 21 words (Head, Heart, Hands, find, best information, evidence available, collective intelligence, compassion, value, paying attention, emotional, psychological, relational, social aspect, tangible reality, practical projects, build, new healthy economy, good intentions, charity, old economy) and 26 symetrical relations between them.

Intermediate transposition aims to present the same information in different media in a way that facilitates its appropriation by the analyst while stimulating reflection. In section 5, we explore several visual transpositions. Preliminary, this section describes the enrichment of the conceptual model to take into account visual aspects of the transposition regardless the specific visual format.

After the introduction of the *Analysis Project* model in subsection 4.1, subsection 4.2 enhances the enriched views on the project which can be derived from the model. Then, section 4.3 formalises the model transformations and their use in the process of drawing an analysis.

4.1 The Knowledge Model

Integrating the elements useful for conducting a structural analysis as described in the section 2, the text-enriched model, which we call *Analysis Project* (or *AP*) consists of a *Text*, *FieldWord*s, *FieldRelation*s, and mappings to a representation domains defined by functions of *DomMap*. Let's first introduce these sets.

Notation 1. *Let*

- *FieldWord* $=$ *Word* $\times$ *Colour denote the set of pairs of words associated with a colour,*

- $FieldRelation = Word \times Word \times Boolean$ *denote the set of pairs of words associated with a boolean denoting if the relation is a association or an opposition*
- $DomMap = FieldWord \rightharpoonup Coordinate$ *denote the set of the partial functions defined on (a subset of) Fieldword into a visualisation domain Coordinate.*

Definition 1. *Formally, the minimal definition of a analysis project can be given by a tuple* (t, wd, rd, ml) *in* $AnalysisProject = Text \times P(FieldWord) \times P(FieldRelation) \times P(DomMap)$ *where*

- $t \in Text$*, is the text object of the analysis,*
- $wd \in P(FieldWord)$ *is the set of purposeful words*[2] *associated with colours. wd is also called the* Words dictionary*,*
- $rd \in P(FieldRelation)$ *is a list of pairs of Words appearing in wd with a boolean denoting if the relation reflects an evocation or an opposition (also called association and dissociation). rd is also called the* Relations dictionary
- $ml \in P(DomMap)$ *is a list of partial function defined on (a subset of) the words in wl into (possibly different) visualisation domains.*

A few notations are associated with this model.

Notation 2. *Given a words dictionary wd, and a specific FieldWord w, we denote by*

- wd_i*, the* i^{th} *element of the list,*
- *w.word, the word item of the FieldWord pair.*
- *w.colour, the colour item of the FieldWord pair.*

Similarly, given a relations dictionary rd, and a specific FieldRelation r we denote by

- rd_i*, the* i^{th} *element of the list,*
- $r.w1$*, the first word item of r,*
- $r.w2$*, the second word item of r,*
- *r.rel, the boolean denoting the type (association or opposition) of r.*

Beside the project analysis itself, the analysis drawing process, as well as its presentation integrates knowledge ($k \in Knowledge$) which basically consists in structured knowledge as dictionaries, lemmatisers, synonyms dictionaries, antonyms dictionary and formalised domain knowledge as well as explicit and implicit expert knowledge.

[2]For sake of simplicity, we call here "word" the the basic unit used in the analysis. In a linguistic perspective, it could be more properly called "lexical item". Indeed, according to the analyst's object of interest it can be either a specific word, a lemma or a fixed set of words

4.2 Enriched Views of an Analysis Project

Based on the minimal representation introduced in the previous section, enriched concepts are built that integrate various element of the Analysis Project Field tuple. A few notations are introduce to support their formal definition.

Definition 2. *Given a word w, exploiting the linguistic knowledge formalised by a dictionary in k, we call* $deriv(w)$ *the set including all its lexical derivation if w is a lemma, the singleton involving only w if w is a derived word or a compound expression.*

In the following, we extend the seminal notation $\sum_{i=1}^{n}$ to the $Append$ function on list

Notation 3. *We denote by*

$$Append_{i=1}^{n} description(...i...)$$

the list composed of the n elements obtained by instantiating i from 1 to n in the description.

Definition 3. *For a given* AnalysisProject $p = (t, wd, rd, gl)$*, one defines the following objects.*

- Enriched Text*: Crossing the text t and wd, the enriched text proposed a version of the text in which the words that appear in wd (or whose lemma appears) are tagged and associated with the same colour as in wd. Formally,*

$$\begin{aligned} &Enriched\_Text: \\ &\qquad Text \times P(FieldWord) \rightarrow TaggedText \\ &Enriched\_Text(t, wd) \\ &\quad = \ Append_{i=1}^{lengt(t)} coloured(t_i) \end{aligned}$$

where

$$\begin{aligned} &coloured(t_i) \\ &\quad = \ t_i < wd_j.colour > \\ &\qquad \text{if } t_i \in deriv(wd_j.word) \text{ for any } j, \\ &\qquad t_i \text{ otherwise} \end{aligned}$$

- Field Dictionary*: enriching the world dictionary wd with knowledge extracted from the text t, the field term dictionary is the list of purposeful words associated with their colour and a natural number denoting the number of occurrences of the word in the text (or the number of its declination if the word is a lemma). This is formalised as follows.*

$$\begin{aligned} &Field\_dictionary: \\ &\qquad Text \times P(FieldWord) \\ &\qquad \rightarrow P(FieldWord \times Colours \times \mathbb{R}) \end{aligned}$$

$$
\begin{aligned}
&Field\_dictionary(t, wd)\\
&\quad = Append_{i=1}^{length(wd)}(w_i.word, w_i.colour,\\
&\qquad \#\{w \in t : w \in deriv(wd_i.word)\})
\end{aligned}
$$

- Field Relation Dictionary*: the relation dictionary, rd, where each word is augmented with the colour of the word in wd and a Boolean denoting the presence/absence of the word (or one of its derivative) in the text. Formally*

$$
\begin{aligned}
&Field\_Relation\_Dictionary :\\
&Text \times P(FieldRelation)\\
&\quad \rightarrow P(FieldWord \times Boolean \times FieldWord\\
&\qquad \times Boolean)
\end{aligned}
$$

$$
\begin{aligned}
&Field\_dictionary(t, rd)\\
&\quad = Append_{i=1}^{length(rd)}\\
&\qquad (r_i.w1, (\exists w \in t : w \in deriv(r_i.w1)),\\
&\qquad r_i.w2, (\exists w \in t : w \in deriv(r_i.w2)))
\end{aligned}
$$

- Field Relation Matrix*: the symmetric matrix M of dimensions $|wd| \times |wd|$ where each element $m_{i,j}$ is a colour that indicates the existence and the type a relation $(wd_i, wd_j) \in rd$. The possibles values are red for an opposition, green for an association and grey if no relation. Formally, denoting Col the set $\{Red, Green, Grey\}$,*

$$
\begin{aligned}
&Field\_Relation\_Matrix :\\
&\quad P(FieldWord) \times P(FieldRelation) \rightarrow \mathcal{M}(Col)
\end{aligned}
$$

$$
\begin{aligned}
&Field\_dictionary(wd, rd)\\
&\quad = \{m_{i,j} = \ Red\ iff\ (wd_i, wd_j, 0) \in rd,\\
&\qquad\qquad Green\ iff\ (wd_i, wd_j, 1) \in rd,\\
&\qquad\qquad Grey\ otherwise\}
\end{aligned}
$$

- Visualisation *which is a graphical representation of a DomMap $\in gl$ enriched with information involved in p as the colours associated with words, their occurrence in the text and the relations between the words.* [3]

Direct access to these enriched objects provides to the human scientist an enriched perception of the text, and guides him into the interpretation process.

[3]At this stage, for the sake of genericity, the model allows any kind of domain for visualisation. In the following we describe more precisely some of the mapping implemented in Evoq.

4.3 Project Transformations and Analysis Project Construction

The model presented above formalises a post-structuralist enhanced model of a text, as the result of an analysis, and the enriched views on the information. We describe in section 5 how it efficiently serves the intermedial transposition. At the formalisation level, a question remains to address: how to build such a model from a fresh text? Expressed in our formalism: how to transform an original $(t, \emptyset, \emptyset, \emptyset, k)$ into an analysis (t, wd, rd, ml, k) and its associated enriched objects?

Let us remind that we do not aim to offer a tool that will fully automate the analysis but a tool that will, on the one hand, stimulate the analyst and, on the other hand, facilitate some tasks of encoding or research. That is to say that only a part of the knowledge k requested to lead the analysis can be fully formalised and implemented, an important part of it remains in the analyst's brain.

With this in mind, the analysis process can be formalised as a sequence of transformation that will, step-by-step, transform a analysis project field (t, wd, rd, ml, k), and

- *extract words* from the text and/or the knowledge using the text and/or the (integrated or human) knowledge, and
- *add these words*, with their associated colour, to wd,
- *identify* relations among the words, from the text and/or the knowledge, and
- *add these relations*, with their associate boolean, to rd.
- *create or adapt* spatial manifestations of the words and the relations to express semantic fields and their tension.

Whether the operations are performed by a human analyst or proposed by an AI agent, our aim here is to show that they can be supported by the proposed formal model.

The current version of the Evoq platform offers a basic function allowing the extraction of the most frequent words from the text, with or without the exclusion of stopwords. The version including all words can be modelled by

$$FrequentWords(t, wd, rd, ml, k) : AP \rightarrow P(FieldWord)$$

$$t \rightarrow \{(w, black) : \#\{i : t_i \in deriv(t_i)\} \geq m\}$$

where t_i denotes the ith word of the text t, m denotes the arbitrary minimal number of occurences k involves the capability to retrieve derivatives of a word.

And the one removing stopwords by

$$FrequentKeyWords(t,k) : Text \to P(FieldWord)$$

$$t \to \{(w, black) : \#\{i : t_i \in deriv(t_i) \setminus Stopwords\} \geq m\}$$

with the same conventions, and *Stopwords* being a list of stop words involved in k.

The addition of a (colored) word (w, col) is defined only if w is not already a member of wd

$$addWord : ((t, wd, rd, ml, k), (w, col)) \rightharpoonup (t, wd \cup \{(w, col)\}, rd, ml, k)$$

The addition of relation is the direct adaptaion of the one on the basic model and similarly defined only if one of $(w_1, w_2, 0)$ or $(w_1, w_2, 1)$ is in rl:

$$addRel : ((t, wd, rd, ml, k), (w_1, w_2, b)) \to (t, wd \cup \{(w_1, w_2\}, rd\{(w_1, w_2, b), (w_2, w_1, b), ml, k)\}$$

The operations for correction in an analysis, namely, remove a term or a relation follow also directly from these on the basic model.

$$\begin{aligned} removeWord : ((t, wd, rd, ml, k), w) \\ \to \quad & (t, wd \setminus \{(w, col) : col \in Colours\}, \\ & rl \setminus \{(w_1, w_2, b) : w_1 = w \vee w_2 = w\}, ml, k) \end{aligned}$$

$$\begin{aligned} removeRel : ((t, wd, rd, ml, k), (w_1, w_2, b)) \\ \to \quad & (t, wd, rd \setminus \{(w_1, w_2, b), (w_2, w_1, b)\}, ml, k) \end{aligned}$$

A last operation consists in modifying the colour of a word already present in the word dictionnary, this is formalised by the partial function

$$\begin{aligned} changeColour : ((t, wd, rd, ml, k), (w, col)) \\ \rightharpoonup \quad & (t, (wd \setminus \{(w, c) : c \in Colours\}) \cup \{(w, col)\}, \\ & rd, ml, k) \end{aligned}$$

5 SEMANTIC LANDSCAPE: INTERMEDIAL TRANSPOSITION

In this section we highlight how our formalisation offers a unifying model and supports the intermediary transposition between different visualisations associated with an Analysis Project which constitute its semantic landscape.

This section is illustrated with the analysis of the text available on (Transition Network, 2021) that presents the principles and value guiding the movement called *Transition Network*. As mentioned above the basic element of this analysis are the text t, the set of 21 words wl, and a list of 52 relations rl.

In the intermedial model, words are associated with colours in the word dictionary wd, the relation dictionary $rd = rl$. The list of mapping ml and elements of the knowledge k that complete the Analysis Project $AP = (t, wd, rd, ml, k)$ are described below.

After a literature review of text-related visualisation, (Clarinval et al., 2018) proposed a selection of visuals and studied their respective suitability to support the structural analysis.

This section present in turn, enriched text, words and relation dictionaries, chord diagram, adjacency matrix and node-link diagram.

5.1 Enriched Text

The text presentation built into Evoq provides a direct transposition of the Enriched text defined by the *Enriched_Text* function in the subsection 4.2. To provide this visualisation, the domain mapping constructs the enriched text and then calls a function that transforms each tagged word into the html expression that applies the required formatting.

Figure 1 illustrates the visualisation of the beginning of the text of our running example[4].

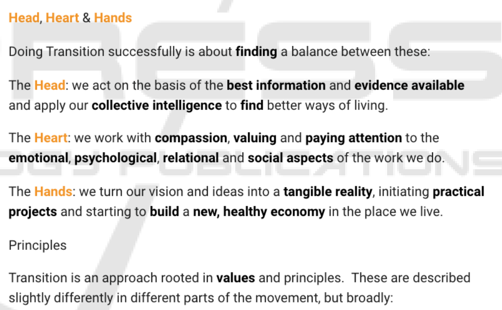

Figure 1: Enriched text visualisation.

5.2 Dictionaries

Dictionaries can be seen as simple lists and presented as such. However, as they are not isolated objects but elements of a complete analysis project AP, the Evoq software propose enriched view of the dictionary. Namely, the word dictionary wd is completed with the number of occurrences of each of its words in the text t using the *Field_dictionary* function and the relation dictionary rd is completed with information related to the words using the *Fied_relation_dictionary* function.

Here again, the mappings consist in computing the enriched elements and then turning the results into

[4]The text is an extract from (Transition Network, 2021) used to build our running example

Couleur	Lemme	Occurrences
	head	3
	heart	3
	hands	3
	(old economy)	0
	(charity)	0
	(good intentions)	0
	best information	1
	evidence available	1
	collective intelligence	1
	find	4
	compassion	1

Figure 2: First items of Word Dictionary Visualisation.

Terme 1 ↑ 1	Force	Terme 2
hands	=	tangible reality
hands	=	practical projects
hands	=	build
hands	=	new, healthy economy
hands	=	(charity)
hands	=	(old economy)
head	<>	heart
head	<>	hands
head	=	best information

Figure 3: First Items of Relation Dictionary Visualisation.

HTML code that construct the tables. Figures 2 and 3 respectively present the first items of the word dictionary and relation dictionary of our running example. Note that, in figure 2 the absence of a word in the text is revealed both by the null number of occurrences and the parenthesis surrounding the word. Similarly, in figure 3, the words are presented with the colour defined in *wd* and are also surrounded by parenthesis if they are absent from the text t.

5.3 Adjacency Matrix

Relations being at the heart of the structural analysis, several visuals are proposed from the relation dictionary. The adjacency matrix proposes an enriched vision from the *Field_Relation_Matrix* function defined in subsection 4.2. After computing the Field Dictionnary and Field Relation Matrix, the mapping create an html tabular having the coloured words as columns and lines headers and each cells receiving the coloured (red/green/grey) defined by the Field Relation Matrix corresponding cell, but for the diagonal elements which are dark grey coloured.

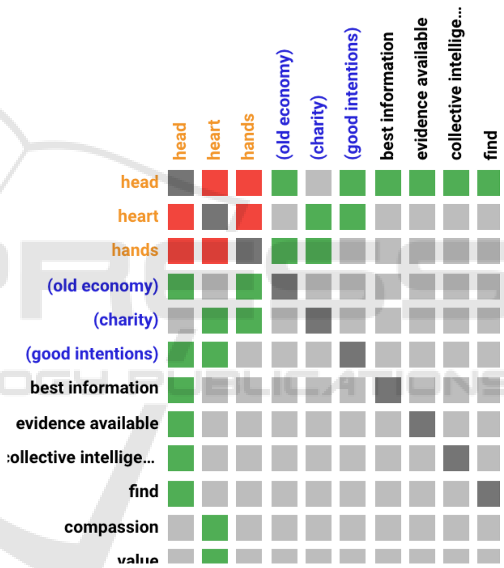

Figure 4: Top-left corner of the adjacency matrix.

5.4 Node Link Diagram

A last visual associated with the set of relationships is to present the relation dictionary as a node-link diagram. For this visual the mapping consists of

- projecting each word in the dictionary into a two-dimensional space
- writing the words in the corresponding colour and surrounding them with brackets if necessary, according to the convention explained above
- draw the arcs corresponding to the relations, in red for oppositions, in green for associations.

After the computation of the Field Relation Dictionary, the last two aspects do not pose any critical problems. The question of the placement of points in space is much more delicate. A large literature deals with the question of the placement of the points of a graph in particular with the objective of limiting the crossings of the arcs. The dictionaries of relations integrating two types of relations, these solutions proved to be unsatisfactory in terms of interpretability of the produced diagram. The Shock Wave algorithm (Cauz et al., 2021) has been specifically developed to visualise relation dictionaries. It places the nodes to highlight the structural axes represented by the opposition relations.

Figure 5 presents the node-link diagram created for the Transition example.

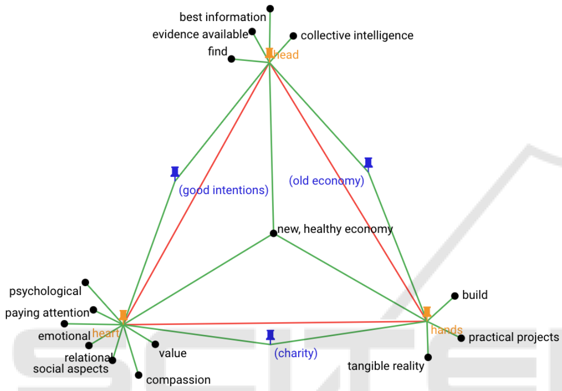

Figure 5: Node-Link diagram.

Each of the visualisations integrates its own interaction mechanisms which give access to the functions described in the section 4.3 : add, delete or modify a word, add or delete a relation (for more details see (Clarinval et al., 2018)). The node-link diagram also offers the specific possibility of modifying the mapping function by changing the position of a node with a simple drag and drop.

6 CONCLUSION

The qualitative analysis of unstructured texts is an essential task for human science researchers. It remains largely unexplored by the IT world, which offers these researchers essentially support tools for the manipulation of texts and elements of analysis that do little to enrich the analyst's reflection or focus on quantitative approaches (Lejeune, 2010).

Through the contribution presented in this work, we develop an innovative perspective which aims to propose a formal modelling of the analyst's approach in order to offer him a variety of visual supports that multiply the views on the analysis in progress while preserving their coherence.

To achieve this ambition, the formal modelling of the analysis integrates in its structure the theory of language mobilised for the analysis: structural analysis. This paper proposes a formal model of knowledge that incorporates this theory. It also illustrates how this model allows for intermediary transposition and ensures the overall consistency of the semantic landscape. The transpositions that make up the semantic landscape provide rich tools both for communicating analyses and for stimulating the creativity of the researcher.

This work served as the basis for the development of the Evoq tool (Clarinval et al., 2018; Linden et al., 2020) available on https://evoq.info.unamur.be/login.

REFERENCES

Cauz, M., Albert, J., Wallemacq, A., Linden, I., and Dumas, B. (2021). Shock wave: a graph layout algorithm for text analyzing. In Patrick Healy, M. B. and Bonnic, A., editors, *Proceedings of the 21st ACM Symposium on Document Engineering*. ACM Press.

Clarinval, A., Linden, I., Wallemacq, A., and Dumas, B. (2018). Evoq: a visualization tool to support structural analysis of text documents. In *Proceedings of the 2018 ACM Symposium on Document Engineering*, United States. ACM Press.

Derrida, J. (1967). *L'écriture et la différence*. Editions du Seuil, Paris.

Elleström, L., editor (2010). *Media Borders, Multimodality and Intermediality*. Palgrave Macmillan, London.

Elleström, L. (2014). *Media transformation. The Transfer of Media Characteristics among Media*. Palgrave Macmillan, London.

Gibson, H., Faith, J., and Vickers, P. (2013). A survey of two-dimensional graph layout techniques for information visualisation. *Information Visualization*, 12(3-4):324–357.

Kucher, K. and Kerren, A. (2015). Text visualization techniques: Taxonomy, visual survey, and community insights. In *Proceedings of IEEE Pacific Visualization Symposium (PacificVis)*, pages 117–121. IEEE.

Lejeune, C. (2010). Montrer, calculer, explorer, analyser. ce que l'informatique fait (faire) à l'analyse qualitative. *Recherches Qualitatives*, 9:15–32.

Lejeune, C. (2021). http://www.squash.ulg.ac.be/lejeune/. web page dedicated to Qualitative Data Analysis Softwares, accessed:2022-03-29.

Linden, I., Wallemacq, A., Dumas, B., Deville, G., Clarinval, A., and Cauz, M. (2020). Text as semantic fields: Integration of an enriched language conception in the text analysis tool evoq®. In Fabiano Dalpiaz, J. Z. and Loucopoulos, P., editors, *Research Challenges in Information Science - 14th International Conference, RCIS 2020, Proceedings*, volume 385 of *Lec-*

ture Notes in Business Information Processing, pages 543–548.

Rajewsky, I. (2002). *Intermedialität*. A.Franke UTB, Stuttgart.

Rajewsky, I. (2005). Intermediality, intertextuality, and remediation: A literary perspective on intermediality. *Intermédialités / Intermediality*, (6):43–64.

Transition Network (2021). https://transitionnetwork.org/about-the-movement/what-is-transition/principles-2/. accessed:2022-03-30.

Wolf, W. (1999). *The Musicalization of Fiction. A Study in the Theory and History of Inter mediality*. Amsterdam ; Atlanta, GA : Rodopi.