An ontology and a memory island to give access to digital literature works

Yan Rucar, Jean-Gabriel Ganascia

▶ **To cite this version:**

Yan Rucar, Jean-Gabriel Ganascia. An ontology and a memory island to give access to digital literature works. Digital Scholarship in the Humanities, 2019, 34 (Supplement_1), pp.i150-i155. 10.1093/llc/fqy059 . hal-02492770

HAL Id: hal-02492770

https://hal.sorbonne-universite.fr/hal-02492770v1

Submitted on 27 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An ontology and a memory island to give access to digital literature works

Yan Rucar, Jean-Gabriel Ganascia

Abstract: *we have designed an ontology to index a corpus of digital literature works. We have given this ontology the shape of a memory island, a navigable virtual territory where categories are regions and descriptors places, and where archives of these ephemeral works are made accessible.*

A corpus: a technically and conceptually digital literature

To define our concept of digital literature, we will first state what it is not. Not all texts available in an electronic form pertain to the field of digital literature, and distinguishing a native digital literature is no more relevant. As a matter of fact, texts which are written in Word, and published online, can very well be printed without their meaning being altered in the least. Are considered as intrinsically digital the literary forms whose development solely relies on computers. The transfer of these works onto a paper leaf would compromise the integrity of their structures, while weakening their contents. Digital poetry revolves around words moving on the screen, hypertext fiction relies on hyperlinked pages arranged in a tree structure, the generator builds a text through a computer program, network-based fictions are updated in real time by Web contents.

These works are comprised of different technical components and media, which also convey meaning. Some interfaces are conditional paths to the texts, because going through them requires deciphering a complex system, a phenomenon that Espen J. Aarseth has named *ergodicity* [1]. Our task consisted in indexing these works woven with various technical functionings and possessing underlying conceptual depths. We have also made up for the scattering of these works which, despite a few specialized journals

(*BleuOrange*, *The New river*), are mostly displayed on their authors' websites. Indeed, there are but a few digital literature publishing houses: *Eastgate* for example. Many portals are accessible (NT2 laboratory database, Electronic Literature directory), where hundreds or even thousands of links to digital works are stored. However, despite this praiseworthy profusion, we have chosen to investigate only one trend of digital literature aesthetics, so as to display a coherence that cannot be found anywhere else. Also, one of our research axes is a feature common to all digital literary works: ephemerality. Because of the quick obsolescence of software and disintegration of platforms, most of these works have vanished since their inception in 1952, with the generator *Love letters* by Christopher Strachey.

The issue of access guides our research:

1. through a unique portal gathering initially scattered works falling within the same aesthetics
2. through an indexation of the works providing a privileged channel to their meaning
3. through the creation of archives preserving a memory of the works despite their predictable disappearance

Indexation: building an ontology

We have chosen to index digital works that challenge reading, as elusive texts are common in electronic literature, a strategy that Philippe Bootz has called the "aesthetics of frustration" [2]. Through fast-paced displays, blurred mobile signifiers, the saturation of words on the screen surface, misleading interfaces or complex paths to the text, the reading of these works includes constant disruptions, sudden suspensions of meaning,

often nonproductive questionings, yet eventually leading to a semantic area at some point. The ontology has been built using Protégé [3] and is in the OWL format. The following screenshots provide a thorough illustration of this ontology.

Images: screenshots Ontology 1 to 4

This ontology includes descriptors explaining the type of archive created, the concepts of the indexed works, the functioning of the computer programs, the categories of unreadability or illegibility (the way the text is elusive), literary genres (specifically digital), the mode of interactivity, the existence of other versions in other formats or dissemination methods (the text may be available, albeit formally divergent, in a printed book; the work may have been exhibited as an installation in a gallery), and the media used (textual, sound, visual properties). Descriptors give an account of both the aesthetics and materiality of the works. Therefore the indexation is both a critical analysis and the technical description of the corpus, this latter aspect being fundamental with regards to the archiving process. The specification of the technical facets of the works informs their preservation.

The aesthetical aspect is represented by the categories of *concepts* (screenshot 1), *forms of unreadability* (screenshot 2), and *literary genres* (screenshot 2), which clarify the intentions behind the works. The *concept* is the core meaning of the work. Although there are numerous descriptors, three conceptual families stand out, from the utter surrender to the whirlpools visible on the screen to the quest for control over the text. The contemplation of the computer's performance, the technological fascination, is the first conceptual feature, through which the spectator watches the generator cut up the text. The descriptor "immersion in language" defines works that surround the

reader/spectator/inter-actor with barely visible words that are likely to collide, which may deprive the users from their bird's-eye view on language, so as to turn this source of meaning into a disorientating agent. The opposite trend, the search for control, is well summed up by the concept of ergodic literature coined by Espen J. Aarseth. Understanding complex systems is a prerequisite to access the text. This reading mode is the literary equivalent of video game. Literary genres may be classified into four groups: visual poetry, animation, text generator and net-art. Visual poetry and animation plunge the reader into language, while interactivity emphasizes this dimension. The text generator uses the machine as a creative writing device, and net-art weaves together web contents taken from different sources. These literary genres exhibit categories of unreadability pertaining to their own textual constructions. Our notion of unreadability is based on the following definition by Bertrand Gervais: "Unreadability is a situation of deficiency: there is a lack which must be overcome" [4]. Our corpus presents a superficial unreadability of the text which must be addressed through an investigation led by the reader/spectator/inter-actor.

The descriptors of the types of archives (screenshot 1), workings (screenshot 4), modes of interactivity (screenshot 2), media specificities (screenshots 3 and 4), leave the conceptual field to focus on the materiality of screens. Workings may be generative (the computer program builds the text), interactive, or consist of a sequence unfolding (video film, mapping software), or the computer opening to networks (links to online newspapers). The rules of interactivity may be obvious, even stated, or completely hidden, in which case the work transforms before the helpless reader.

Five types of documents are preserved: texts, images, sounds, processes recorded on video, computer programs (source-codes). The textual dimension of the works (screenshot 3) centres around the processes of appropriation and transformation, fragmentation, the pairing of the signified and the graphic signifier. The sound dimension (screenshot 3) may be central or secondary, the descriptors "sound related to the text" and "speech synthesis" apply to digital works regulated by or regulating sound files. The visual dimension (screenshot 4) may be conveyed by the graphic signifier of the text, or be autonomous yet interconnected because images go through the same motions as the words. Iconicity may also be completely independent from language.

Visualization of categories and descriptors: generation of a memory island

The LIP6 laboratory has developed a system generating memory islands from ontologies. Categories become territories and descriptors places, according to a concepts visualization process reminiscent of the ancient arts of memory. The rhetorician would memorize the stages of his discourse through *loci*, images and places located in various parts of a palace. The non-linear navigation on the memory island is similar to this journey through *loci*, as described by Frances A. Yates: "It is essential that the places should form a series […] so that we can start from any locus in the series and move either backwards or forwards from it" [5]. This memorization aid also proves useful when it comes to conceptualization, because it makes the abstract tangible, as well as what lies in its vicinity. Developed by Bin Yang and Chakib Chekrouni, our memory island shows territories that are areas of meaning transversal to the indexed works, and descriptors (features) belonging to these categories. When clicked on, these descriptors open onto a list of works pertaining to this specific property, with direct links to them. One first sees

the whole island map, i.e. a body of literary works with its distinctive features, before going into the details of the corpus. Accessing the meaning of enigmatic works goes together with their exhibition; the cartography is at once an enlightening tool and a portal. This memory island can be accessed on the Internet[1]. Below are close-ups of the island describing the aspects of the works.

Screenshot Memory Island 1: Concepts

Screenshot Memory Island 2: Form of unreadability or illegibility

Screenshot Memory Island 3: Interactivity

Screenshot Memory Island 4: Literary genre

Screenshot Memory Island 5: Sound dimension

Screenshot Memory Island 6: Textual dimension

Screenshot Memory Island 7: Visual dimension

Screenshot Memory Island 8: Working

It is possible to export these views and store them on a computer. There are six zoom levels, as well as a tree structure on the left of the screen, which provides another view of the categories and descriptors of the ontology. When the information icon at the bottom of the screen is activated, a text pops up. This text presents the digital literary aesthetics covered by the ontology and the memory island. Clicking on a descriptor opens a window giving explanations on its function, since part of the critical jargon used in the field of digital literature might be misunderstood by the general public. Contextualization is therefore very strong; not only is the ontology itself very visible, but it is also surrounded by numerous clarifications. These explanatory words will soon be complemented with the

[1] https://ilesdememoire.lip6.fr/beta/Oeuvres/result/resultid.html

authors' answers to questionnaires. One third of the authors have agreed to give their comments, so that the voices of creators counterbalance the bias inherent to every indexation. These questionnaires will soon be accessible through a new region of the island.

Some of the archives are already accessible through the memory island, and more documents will be added. All archives, whether already online or soon to be added, consist of:

- 568 screenshots
- 10 source-codes
- 26 sound recordings
- 32 video recordings
- 9 texts

Beside the fact that they are traces of an ephemeral literature, these documents have an informative aim. Digital works are ever-changing, any return to a previous stage of these reconfigurations is extremely difficult, if not impossible. The archives offer fixed images and sounds, as well as texts which originally escape reading. The teaching of digital literature would greatly benefit from these stable artifacts.

The memory island, a tool to grasp a diversity of meanings

The ontology and its visualization are non-linear practices (because these representations have as many entry points as categories) which aim to give a critical and technical account of a corpus characterized by a multiplicity of perspectives and material aspects. Therefore, the conceptualization of this literature through cartography meets creative forms defying any synthetic approach. The fixation of these works as archives is

obviously their reduction, but it is the only way to preserve a literature written in sand, a memory enabled by the conceptual framework in which these traces are set.

Works cited

[1] Aarseth, Espen J. (1997). *Cybertext: perspectives on ergodic literature*. Baltimore: Johns Hopkins University Press

[2] Bootz, Philippe. (2006). *Vers de nouvelles formes en poésie numérique programmée?*, http://www.rilune.org/images/mono5/4_bootz.pdf

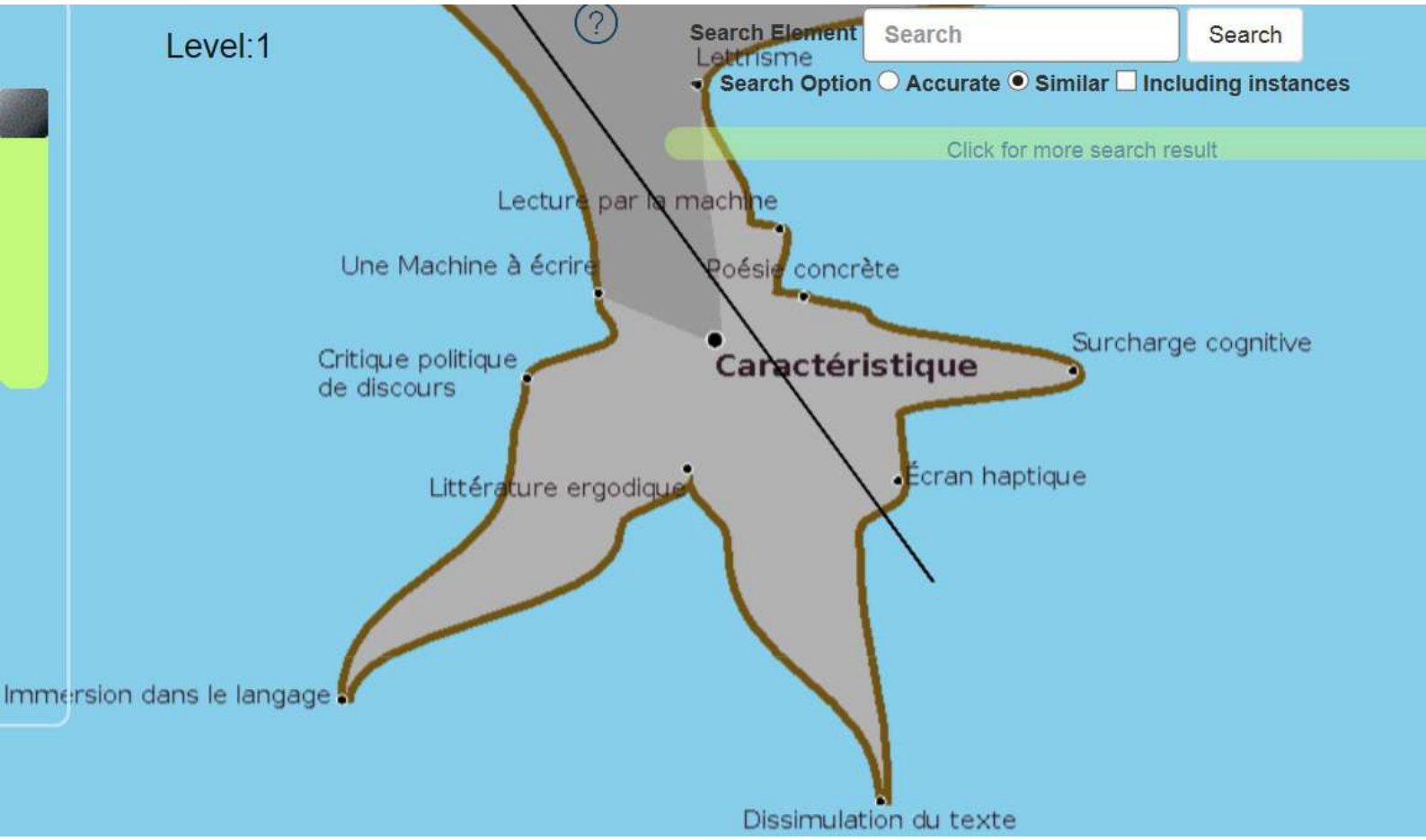
[3] Kapoor, B., Sharma, S. (2010). *A comparative study ontology building tools for semantic web applications*. International journal of Web & Semantic Technology (IJWesT), 1(3)

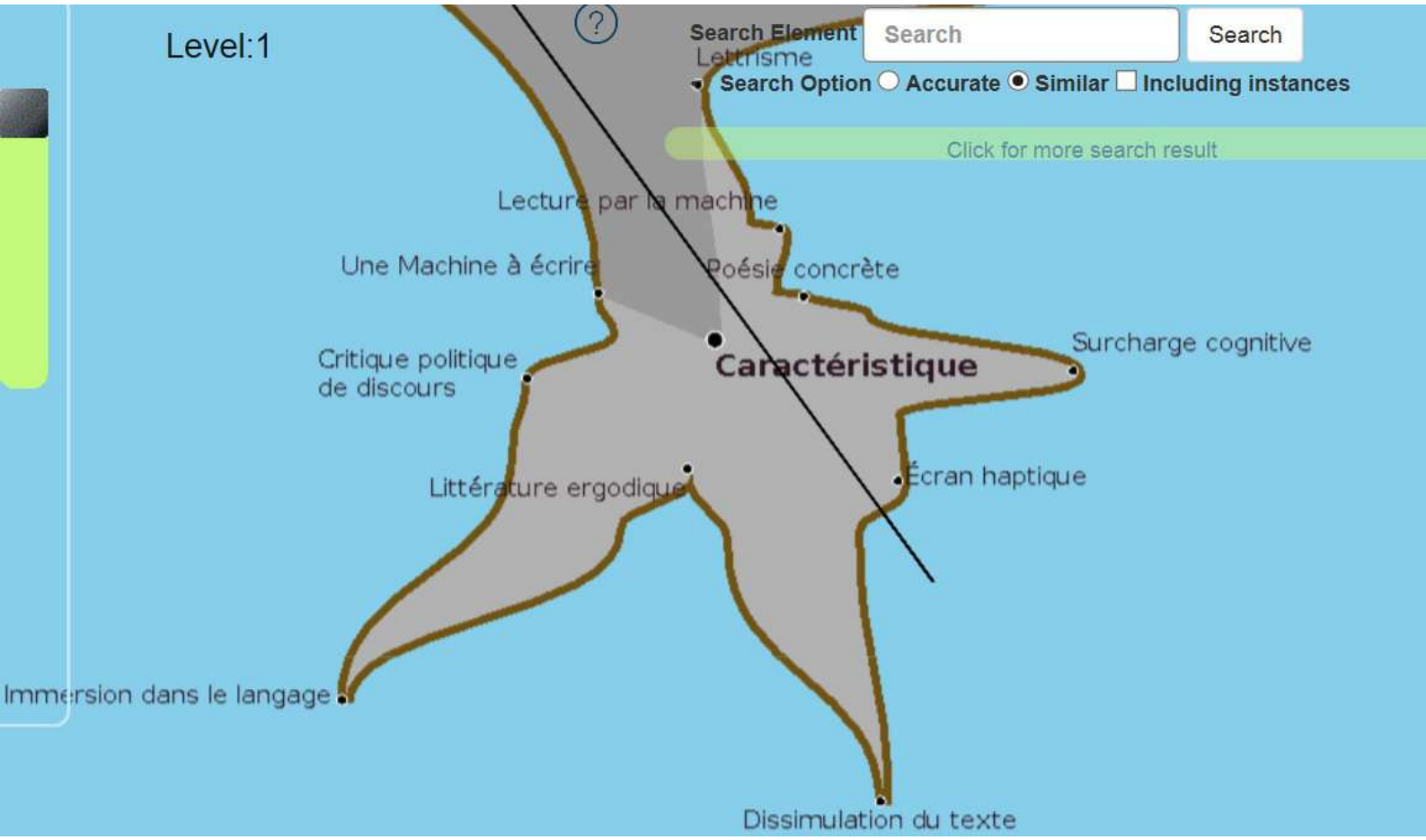
[4] Gervais, Bertrand. (1999). *Presbytère, hiéroglyphes et dernier mot. Pour une définition de l'illisibilité*

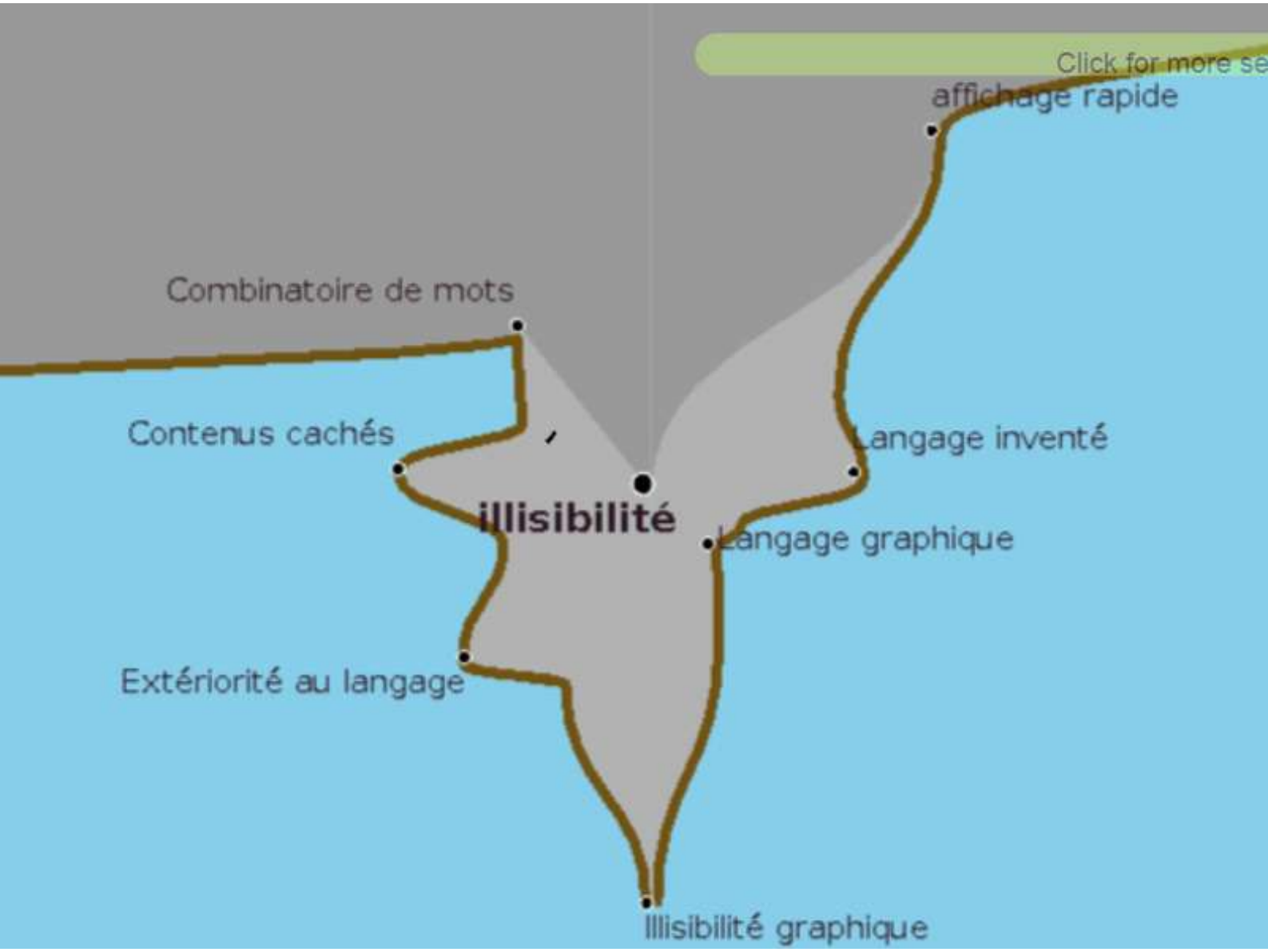
http://blogue.nt2.uqam.ca/cenestecritnullepart/2009/06/11/presbytere-hieroglyphes-et-dernier-mot-pour-une-definition-de-lillisibilite/

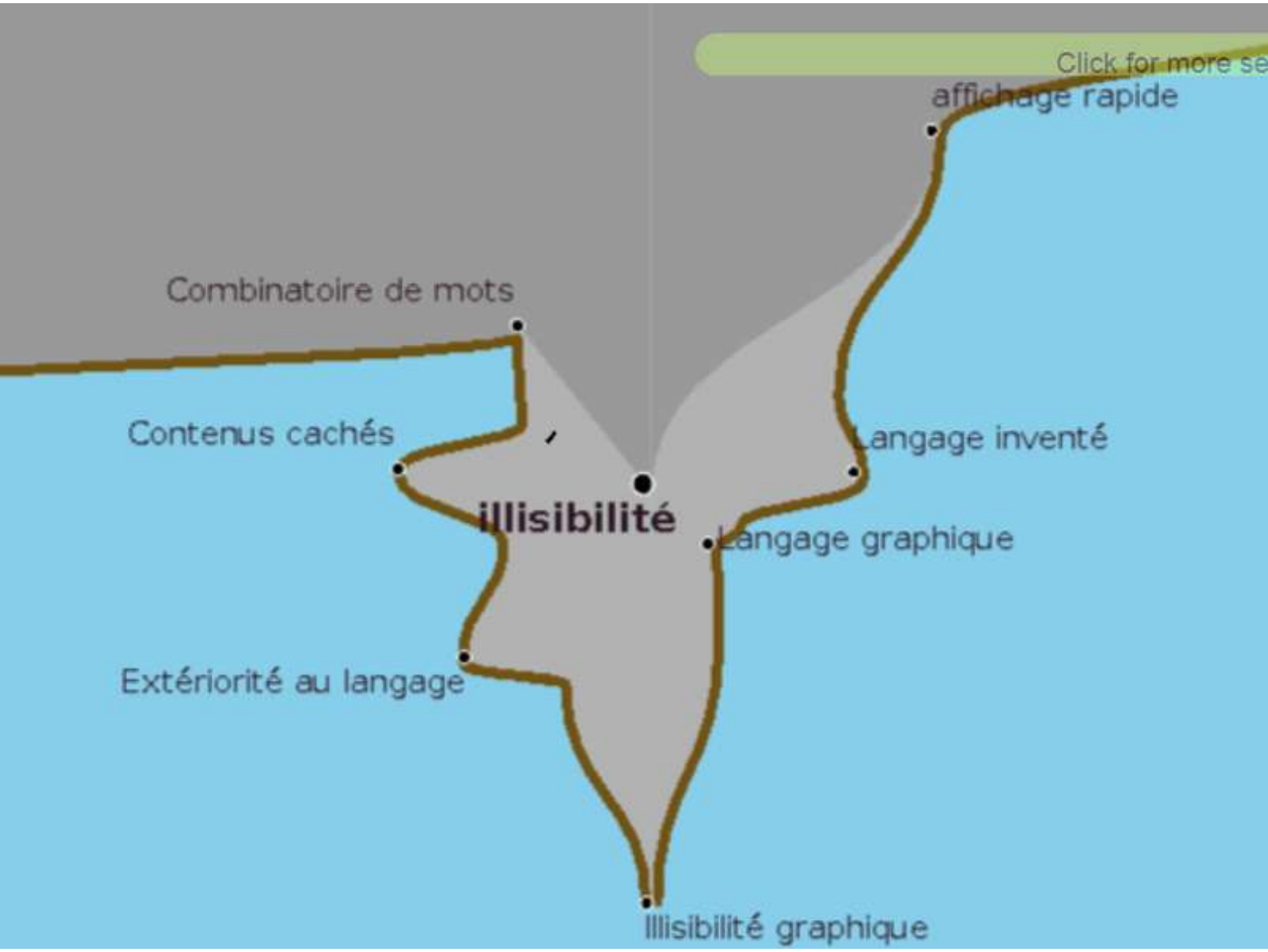
Our translation.

[5] Yates, Frances A. (1966). *The Art of memory*. Londres et New-York: Routledge, p.7


Level:1
Search Element
Search
Search
Lettrisme
Search Option
Accurate
Similar
Including instances
Click for more search result
Lecture par la machine
Une Machine à écrire
Poésie concrète
Surcharge cognitive
Caractéristique
Critique politique
de discours
Littérature ergodique
Écran haptique
Immersion dans le langage
Dissimulation du texte

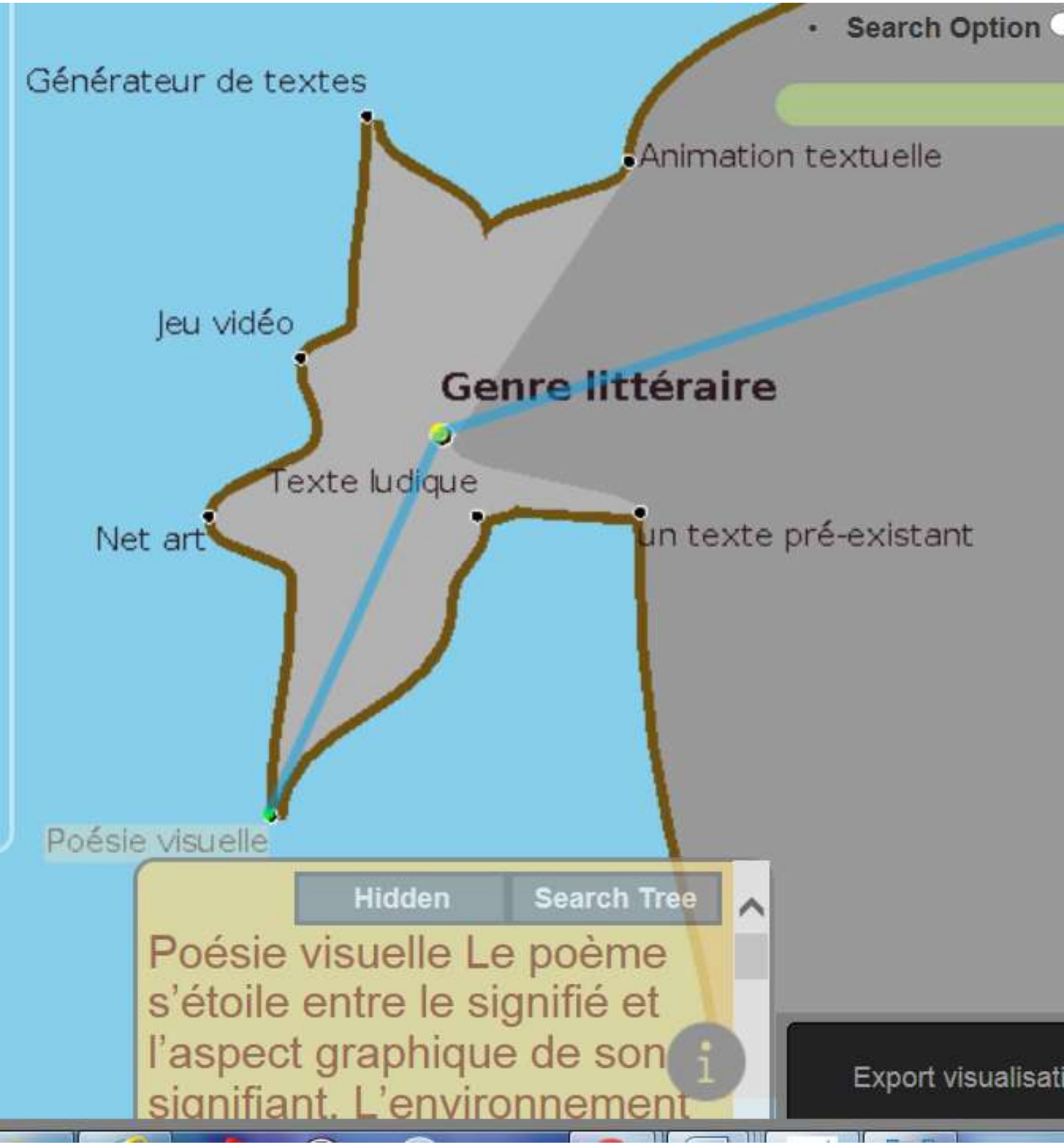

Level:1
Search Element
Search
Search
Lettrisme
Search Option
Accurate
Similar
Including instances
Click for more search result
Lecture par la machine
Une Machine à écrire
Poésie concrète
Surcharge cognitive
Caractéristique
Critique politique
de discours
Littérature ergodique
Écran haptique
Immersion dans le langage
Dissimulation du texte

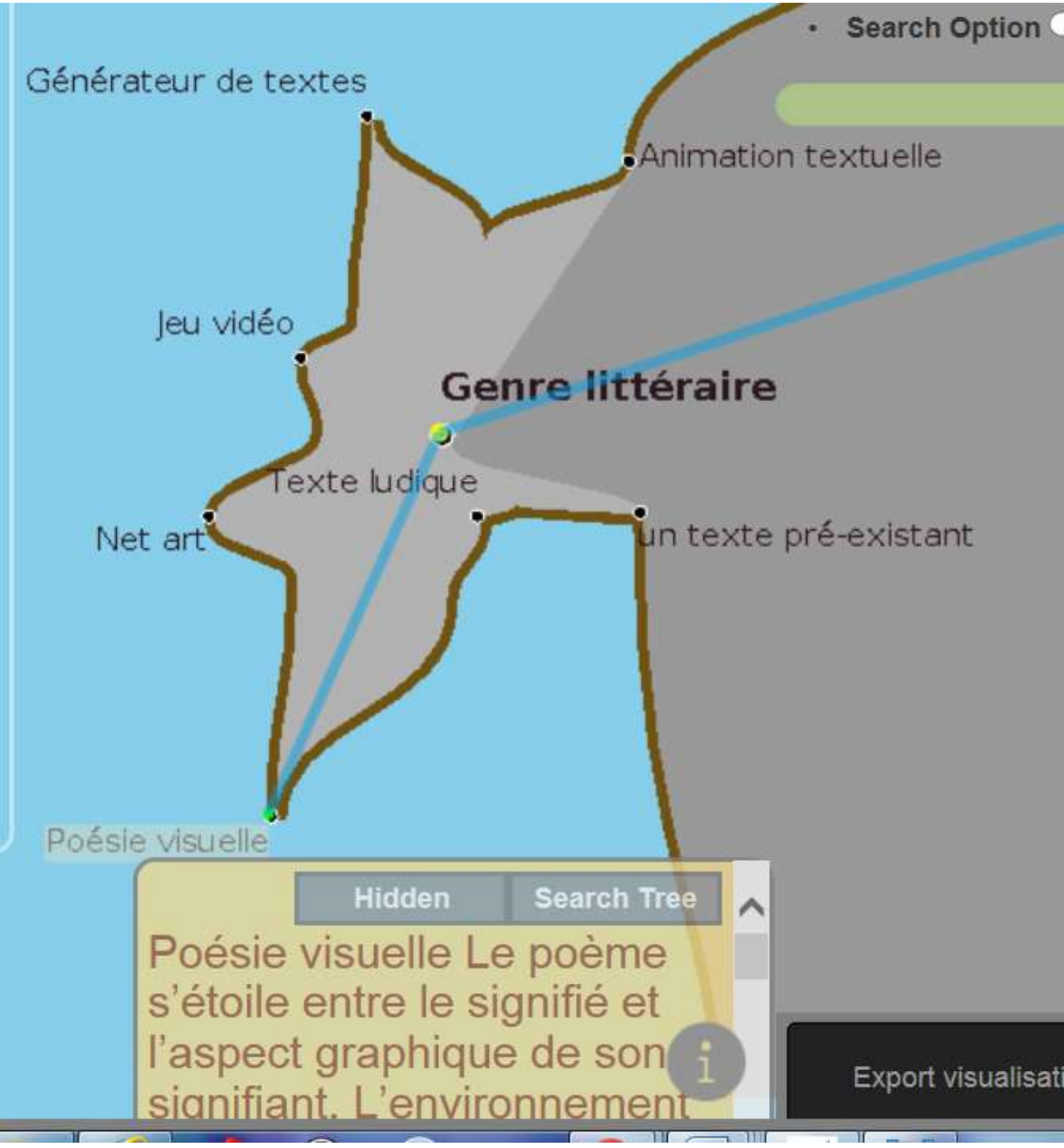

Click for more se
affichage rapide
Combinatoire de mots
Contenus cachés
Langage inventé
illisibilité
Langage graphique
Extériorité au langage
Illisibilité graphique

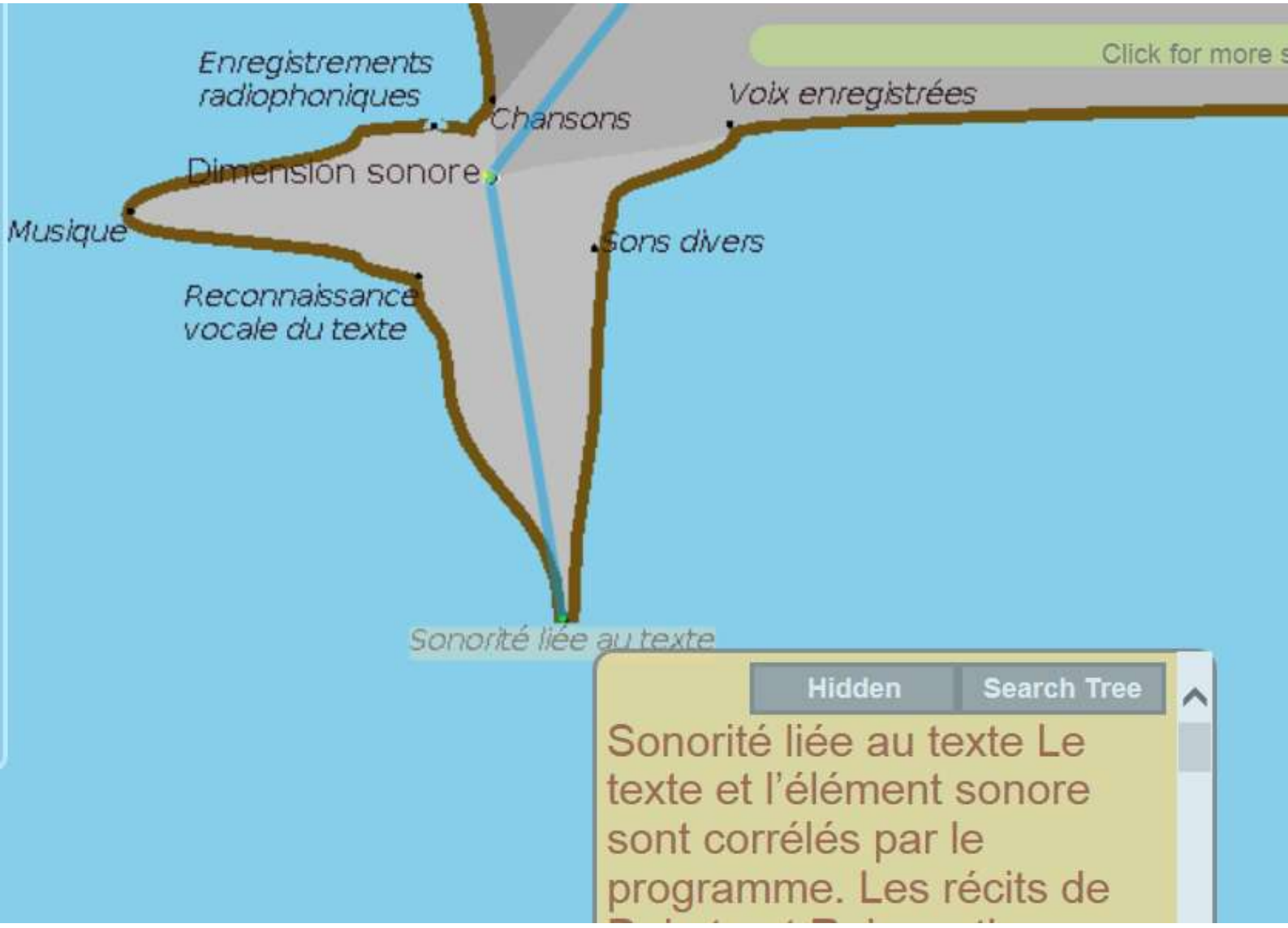

Click for more se
affichage rapide
Combinatoire de mots
Contenus cachés
Langage inventé
illisibilité
Langage graphique
Extériorité au langage
Illisibilité graphique

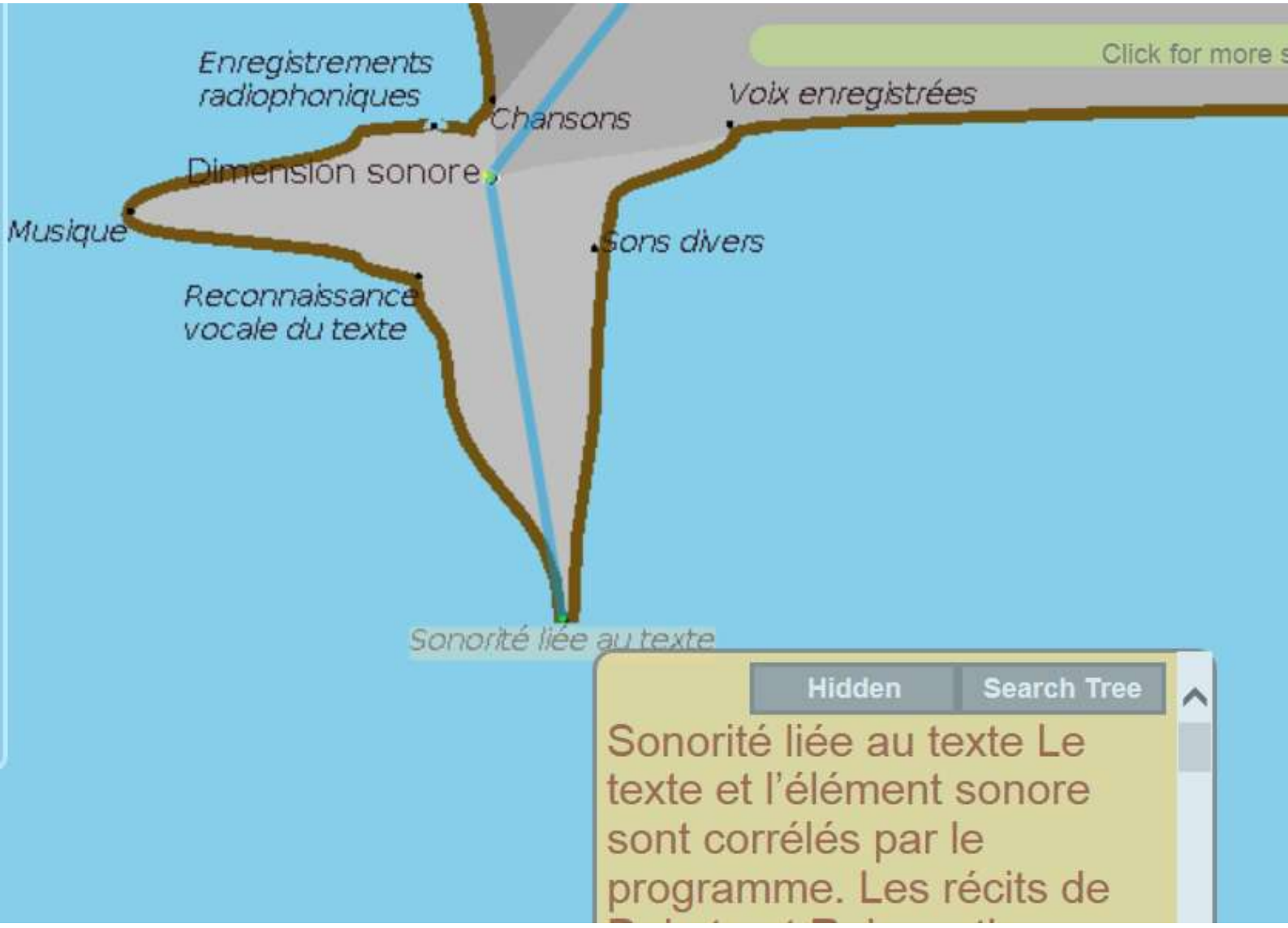
Interactivité consciente
Instances: [
Robolettries
Family tree
Le rabot poète
Soliloquy
Les amis sur le seuil
Dead tower
Interactivité
Interactivité inconsciente

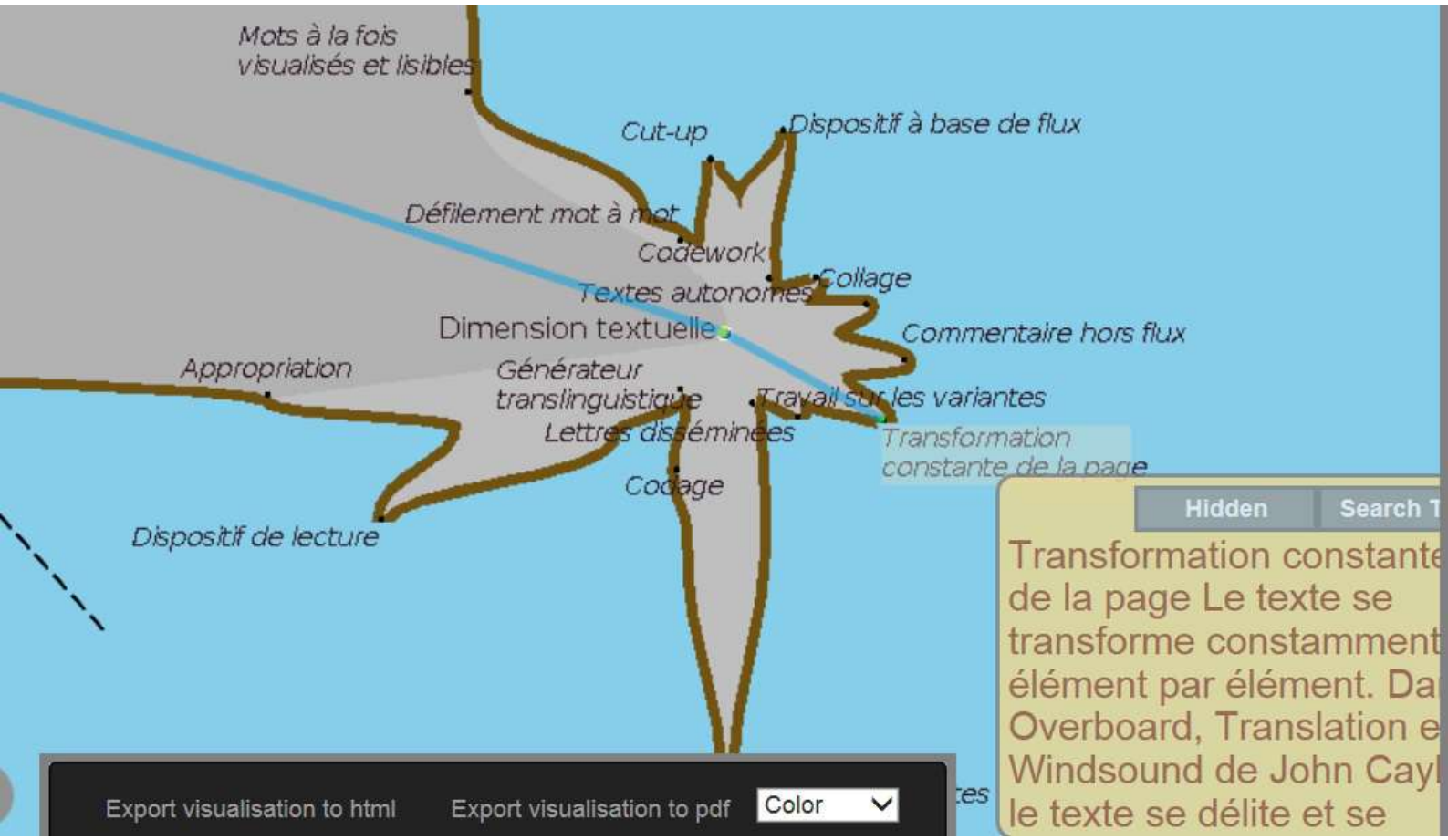
Interactivité consciente
Instances: [
Robolettries
Family tree
Le rabot poète
Soliloquy
Les amis sur le seuil
Dead tower
Interactivité
Interactivité inconsciente

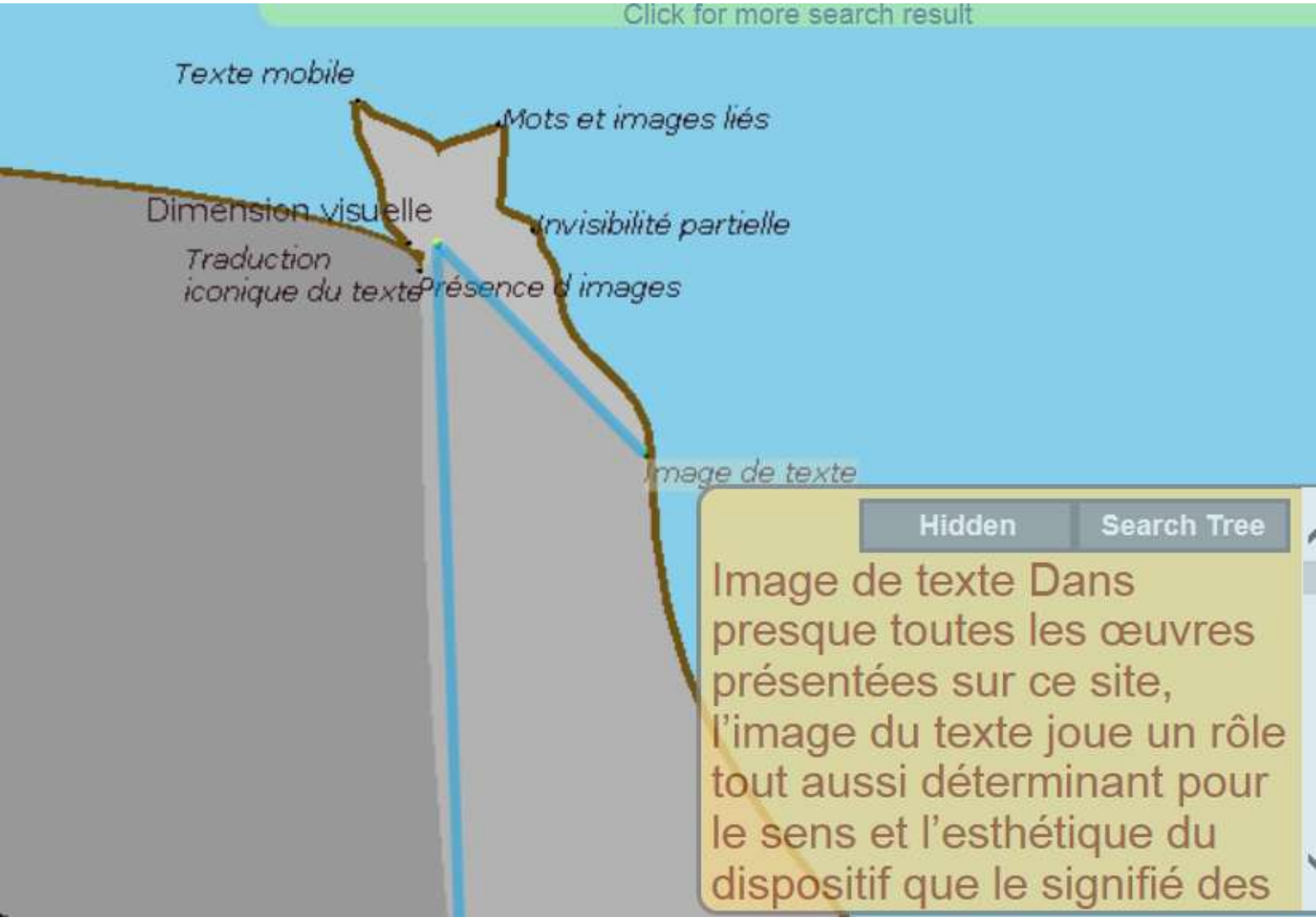

Search Option
Générateur de textes
Animation textuelle
Jeu vidéo
Genre littéraire
Texte ludique
Net art
un texte pré-existant
Poésie visuelle
Hidden
Search Tree
Poésie visuelle Le poème s'étoile entre le signifié et l'aspect graphique de son signifiant. L'environnement
Export visualisati

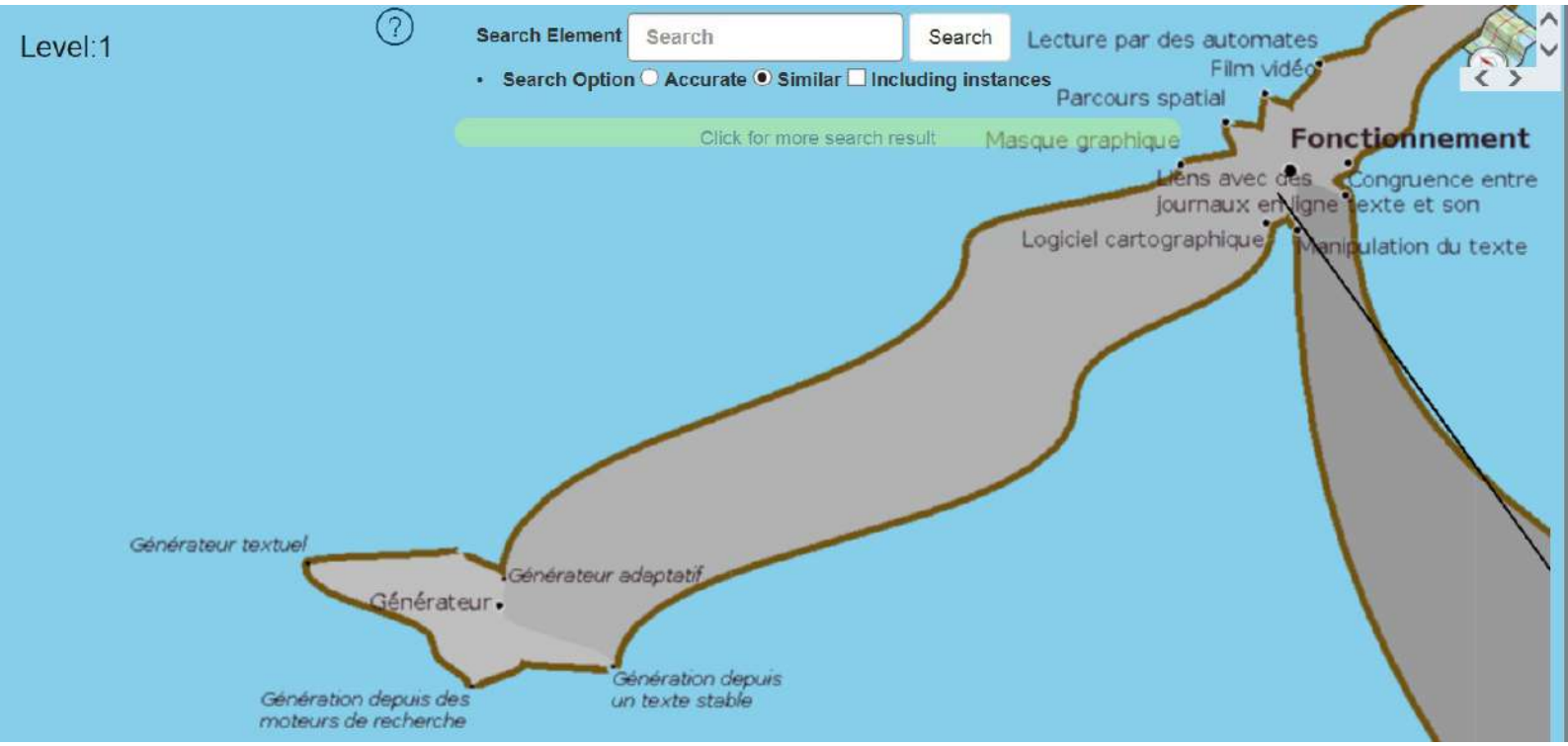

Search Option
Générateur de textes
Animation textuelle
Jeu vidéo
Genre littéraire
Texte ludique
Net art
un texte pré-existant
Poésie visuelle
Hidden
Search Tree
Poésie visuelle Le poème s'étoile entre le signifié et l'aspect graphique de son signifiant. L'environnement
Export visualisati


Click for more s
Enregistrements radiophoniques
Chansons
Voix enregistrées
Dimension sonore
Musique
Sons divers
Reconnaissance vocale du texte
Sonorité liée au texte
Hidden
Search Tree
Sonorité liée au texte Le texte et l'élément sonore sont corrélés par le programme. Les récits de


Click for more s
Enregistrements radiophoniques
Chansons
Voix enregistrées
Dimension sonore
Musique
Sons divers
Reconnaissance vocale du texte
Sonorité liée au texte
Hidden
Search Tree
Sonorité liée au texte Le texte et l'élément sonore sont corrélés par le programme. Les récits de


Mots à la fois visualisés et lisibles
Cut-up
Dispositif à base de flux
Défilement mot à mot
Codework
Collage
Textes autonomes
Dimension textuelle
Commentaire hors flux
Appropriation
Générateur translinguistique
Travail sur les variantes
Lettres disséminées
Transformation constante de la page
Collage
Dispositif de lecture
Hidden
Search T
Transformation constante de la page Le texte se transforme constamment élément par élément. Da
Overboard, Translation e
Windsound de John Cay
le texte se délite et se
Export visualisation to html
Export visualisation to pdf
Color


Click for more search result
Texte mobile
Mots et images liés
Dimension visuelle
Invisibilité partielle
Traduction iconique du texte
Présence d images
Image de texte
Hidden
Search Tree
Image de texte Dans presque toutes les œuvres présentées sur ce site, l'image du texte joue un rôle tout aussi déterminant pour le sens et l'esthétique du dispositif que le signifié des


Level:1
Search Element
Search
Search
Search Option
Accurate
Similar
Including instances
Click for more search result
Lecture par des automates
Film vidéo
Parcours spatial
Masque graphique
Fonctionnement
Liens avec des journaux en ligne
Congruence entre texte et son
Logiciel cartographique
Manipulation du texte
Générateur textuel
Générateur adeptatif
Générateur
Génération depuis des moteurs de recherche
Génération depuis un texte stable

- Archives
  - Screenshots
  - Sound_recordings
  - Source_codes
  - Texts
  - Video_recordings
- Concepts
  - A_writing_machine
  - Cognitive_overload
  - Concrete_Poetry
  - Ergodic_literature
  - Fragmentation_of_language
  - Gap_between_text_and_image
  - Haptic_screen
  - Hidden_text
  - Immersion_in_language
  - Lettrism
  - Machine-performed_reading
  - Machine-simulated_human_language
  - Play_on_the_audible_word_and_the_visible_word
  - Political_criticism_of_discourses
  - Staging_of_the_computer_code
- Dissemination_formats_beside_Internet_and_CD-ROM
  - Installation_in_a_gallery
  - Musical_work
  - Performance
  - Printed_text

- Form_of_unreadability_or_illegibility
  - Combinatory_of_words
  - Cryptic_language
  - Deconstruction_of_language
  - Exteriority_to_language
  - Fast_display
  - Gap_betweeen_audible_words_and_visible_words
  - Hidden_contents
  - Illegibilty
  - Invented_language
  - Lack_of_semantic_relationship_between_words
  - Visualized_Language
- Interactivity
  - Context_aware_interactivity
    - Manipulation_of_the_text
    - Traversal_of_a_space
  - Context_unaware_interactivity
- Literary_genre
  - Codework
  - Digital_reading_of_a_pre-existing_text
  - Net_art
  - Playful_text
  - Sound_Poetry
  - Text_generator
  - Textual_animation
  - Video_game
  - Visual_Poetry