



Rapport de recherche

2004

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Note de synthèse concernant EVOQ: un logiciel d'analyse et de
visualisation de données textuelles. Rapport final d'activité à destination de
la Fondation Francqui

De Saint-Georges, Ingrid

How to cite

DE SAINT-GEORGES, Ingrid. Note de synthèse concernant EVOQ: un logiciel d'analyse et de
visualisation de données textuelles. Rapport final d'activité à destination de la Fondation Francqui. 2004

This publication URL: <https://archive-ouverte.unige.ch/unige:30775>

Rapport d'activité bourse Francqui

Note de synthèse concernant EVOQ : un logiciel d'analyse et de visualisation de données textuelles

Ingrid de Saint-Georges

ReCCCoM

Facultés Universitaires Notre-Dame de la Paix

8, rempart de la Vierge, 5000 Namur

Contact info : (ingrid.desaint-georges@fundp.ac.be)

ABSTRACT. *The ReCCCoM has developed expertise in two types of domains. On the one hand, it has carried out research in the domain of crisis management and risk perception; on the other hand, in the domain of language and perception in organizations. This research has led the ReCCCoM to develop its own tools for data collection and analysis. EVOQ, the software developed in this context, is a tool to construct cognitive maps of conceptual domains, bringing qualitative, structural and phenomenological approaches to bear on the analysis of textual data. At the heart of the program, the main algorithm allows the user to produce graphical visualisation of the results obtained. The 2-D and 3-D visualization modules (in 'semantic landscapes') free the user from linear representations privileged by other systems for cognitive cartography (such as for example, 'decision-explorer'). It allows exploring the terms of reference constitutive of the speakers' cognitive world in a non-linear fashion. This note, based on the research carried out over the last 8 years by Jean-Marie Jacques, Anne Wallemacq and their previous and current collaborators, provides a brief synthesis of where the program stands with the latest development now operational. It also seeks to position the software with regards to research done in the field of qualitative data analysis. It finally seeks to provide a basis for a future user's manual.*

KEYWORDS : *discourse analysis, qualitative approach, cartography, cognition*

RESUME. *Les recherches du ReCCCoM dans le domaine de l'étude des crises et de la perception des risques, d'une part, et sur le langage et la perception dans les organisations, d'autre part, ont amené*

ses initiateurs, Jean-Marie Jacques et Anne Wallemacq à développer leurs propres outils de collecte et de traitement des données. EVOQ, le logiciel conçu dans ce cadre est un logiciel de cartographie cognitive, combinant approches qualitatives, structurales et phénoménologiques, dans l'analyse de données textuelles. Au centre de ce programme, l'algorithme principal permet une saisie graphique des résultats. La visualisation (en 2-D ou en 3-D sous forme de paysages sémantiques) permet de sortir des formes linéaires classiques privilégié par d'autres systèmes de cartographie cognitive (type 'decision-explorer'). Elle permet de mettre en évidence de façon non-linéaire les termes de référence structurant l'univers cognitif du locuteur. La présente note constitue une brève synthèse concernant les derniers développements du programme actuellement opérationnels. Il vise à positionner le logiciel par rapport à d'autres logiciels existant dans le champ de l'analyse des données qualitatives. Sur base du travail de test du logiciel effectué, il cherche aussi à poser les bases de la description du programme en vue de la rédaction d'un futur « manuel de l'utilisateur ».

MOTS-CLES : analyse du discours, méthode qualitative, cartographie, approche cognitive

1. Introduction

Les recherches conduites au sein du ReCCCoM dans le domaine de l'étude des crises et de la perception des risques, d'une part, et sur le langage et la perception dans les organisations, d'autre part, ont amené les initiateurs du ReCCCoM à élaborer leurs propres outils de collecte et de traitement des données (Wallemacq *et al.* 2004 : 4). Dans cette note, nous souhaiterions résumer les fonctionnalités les plus récentes du logiciel EVOQ développé dans ce cadre et synthétiser ainsi le travail post-doctoral effectué en rapport avec le logiciel dans le cadre de notre mandat Francqui de collaborateur scientifique postdoctoral étranger. Cette note vise donc à reprendre les grandes lignes du logiciel telles qu'elles ont été déjà décrites par ailleurs, en apportant un nouvel éclairage issu plus particulièrement du domaine de l'analyse de données textuelles assistée par ordinateur. Elle cherchera aussi à poser les bases pour un futur manuel de l'utilisateur. Nous contribuerons ainsi à documenter un nouveau tournant et une nouvelle étape dans le développement du logiciel.

Plus précisément, trois dimensions serviront de fil conducteur à notre exposé du fonctionnement du logiciel. Nous commencerons par situer le logiciel en rapport avec d'autres approches logicielles existantes. Plutôt que de faire un bilan exhaustif des logiciels d'aide à l'analyse de données qualitatives, nous chercherons à montrer où s'inscrit EVOQ dans le paysage des logiciels disponibles et ce qui constitue sa spécificité d'un point de vue épistémologique (section 2). Dans un second temps, nous exposerons les fondements théoriques et opératoires du programme et décrirons la fonction de visualisation, au cœur de logiciel (section 3). Nous renverrons enfin aux principaux résultats d'analyses obtenus dans le domaine de la gestion des crises et de l'étude des perceptions dans l'organisation en guise d'illustration. L'objectif visé dans cette dernière partie ne sera pas tant les résultats pour eux-mêmes mais l'éclairage qu'ils permettent de donner aux concepts qui auront été présentés au préalable (section 4).

2. L'analyse des données qualitatives

Depuis un peu plus d'une vingtaine d'années, les sciences sociales et de gestion ont vu proliférer les outils d'aide à l'analyse issus de la micro-informatique. La démocratisation de moyens de calculs particulièrement adaptés à l'analyse de quantités importantes d'information a eu un impact non négligeable non seulement pour l'analyse de données quantitative mais aussi dans le traitement de données qualitatives, et notamment de l'information non structurée¹. D'autres ont déjà mis en évidence les grandes tendances dans ces recherches. Sans prétendre à l'exhaustivité, on peut distinguer ainsi plusieurs approches méthodologiques et logicielles dont les lignes de faite peuvent être brièvement résumées (dans la réalité les limites entre les catégories sont évidemment plus floues et poreuses que la catégorisation retenue peut le donner à penser).

Dans le champ anglo-saxon, les approches informatisées de données qualitatives sont souvent rassemblées sous l'acronyme CAQDAS (*Computer-Assisted Qualitative Data Analysis*). Introduite par Fielding et Lee (1995), cette étiquette désigne dans un premier temps les outils d'aide au codage. Gibbs, Friese et Mangabeira (2002) notent que plusieurs générations de logiciels se sont succédés. Les logiciels de première génération sont ainsi conçus principalement pour la simple gestion des données (ils permettent la réalisation de recherches, de tris et de comptage simples et peu élaborés). Ils permettent de naviguer dans les données et d'en brosser à grands traits une première « image » assez limitée. Les logiciels CAQDAS de seconde génération ajoutent à ces opérations de base des fonctionnalités pour coder les données, manipuler le texte, effectuer des recherches et générer des résultats sous formes de rapports. Certains programmes, dits de troisième génération, vont plus loin que ces manipulations qui restent d'ordre *mécanique* plutôt que *conceptuel* (Thomson 2002). Ces programmes, appelés parfois « theory builders » or « model builders » visent à être un tremplin pour générer de nouvelles idées théoriques, tester des hypothèses et/ou permettre la collaboration avec d'autres chercheurs (*data sharing*) (Gibbs, Friese et Mangabeira 2002). De manière générale, les CAQDAS anglo-saxons peuvent donc se concevoir principalement comme des outils d'aide au dépouillement des données, visant à alléger des opérations souvent coûteuses en terme de temps pour le chercheur confronté à une masse importante de données ainsi qu'à stimuler son travail de réflexion au cours de ce processus (surtout dans le cas des CAQDAS de troisième génération).

Dans le domaine francophone, le champ de recherche en rapport avec le développement logiciel de traitement de données textuelles se présente sous une forme plus éclatée. Jenny (1997 ; cf. également Lejeune 2004) distingue par exemple sept approches attestées dans l'analyse des données textuelles : les approches 1) de type *lexicométrique* (comparant des profils lexicaux et s'appuyant sur la statistique d'analyse factorielle et la classification automatique proposée par Benzécri), 2) de type *socio-sémantique* (codant des unités de sens dans le corpus, les catégorisant sur plusieurs dimensions et les soumettant ensuite à un

¹ Le terme « information non structurée » est un terme qui apparaît depuis quelques années pour désigner des données qui ne sont pas classées ou identifiées par des balises ou un format particulier et qui représentent dès lors des difficultés particulières dans le domaine de la fouille textuelle et de l'analyse assistée par l'informatique.

traitement statistique type « arbre de décision » ou « data mining »), 3) de type *réseaux associés* (identifiant les acteurs/actants au moyen de mots et réseaux de mots qui leurs sont associés, de leur 'profil d'association'), 4) de type *analyse propositionnelle et prédicative* (s'intéressant aux logiques de construction de la cohérence dans le texte, aux opérateurs argumentatif, aux liens de causalités), 5) de type *ingénierie textuelle* (facilitant l'indexation et recherche par mot clé de l'ensemble des emplois d'un mot ou d'un terme, dégagement de thèmes à visée de documentation). On recense aussi 6) des logiciels développés (comme les CAQDAS) plus spécifiquement pour le *traitement d'enquêtes* et les observations de type sociologique et 7) des *systèmes-experts*, orientée vers l'étude de problématiques spécifiques et prédéfinies par le chercheur suivant des modèles théoriques qui leurs sont propres.

D'un point de vue historique, comme le notent Duchastel *et al.* (1989), ce sont les approches statistico-morphologiques qui ont dominé le champ des recherches jusque vers le milieu des années '50. Des années 50 aux années 70, les recherches se sont ensuite déplacées vers le développement de modules syntaxiques dans le traitement des données en langues naturelles, concurrencées dès les années 60 par le développement de modules logico- sémantiques. Depuis les années 70, une quatrième tendance s'est confirmée, concernant la recherche sur « la représentation et l'organisation de la connaissance en faisant appel à des modèles cognitifs » (Duchastel *et al.* 1989). C'est dans la mouvance des développements sémantiques et de gestion des connaissances qu'il convient sans doute de situer le logiciel EVOQ.

Cet ensemble de recherches, par la diversité des méthodes utilisées et des objectifs poursuivis, révèle bien la complexité des problèmes à résoudre ainsi que la multiplicité des points de vue possible dès qu'on s'intéresse au traitement du discours en langue naturelle et à l'analyse des représentations textuellement encodées.

Un dilemme commun, semble-t-il, pour toutes ces approches est de concevoir des modes de traitements de grandes quantités de texte (traitement en général d'ordre quantitatif) tout en restant dans un système d'interprétation relativement « ouvert », préservant la dimension ambiguë, polysémique et la sensibilité aux variations contextuelles qui caractérisent les données textuelles (traitement qualitatif) et qui est l'objet de l'analyse du discours. D'où la nécessité, comme le montrent bien Teil et Latour (1995) de construire des systèmes qui visent le résultat suivant :

« ... to give qualitative workers a Computer Aided Sociology CAS tool that has the same degree of finesse as traditional *qualitative* studies but also has the same mobility, the same capacities of aggregation and synthesis as the *quantitative* methods employed by other social sciences. »

Le passage au filtre des procédures de standardisation nécessaire pour l'application de méthodes quantitatives ne permet souvent pas ce traitement plurivoque, en « finesse », et proche de l'étude de cas, d'où l'intérêt de construire des systèmes qui se veulent avant tout des outils **d'aide à l'analyse**, apportant une intelligibilité nouvelle dans l'analyse des données tout en laissant à l'analyste l'autonomie et la responsabilité du travail d'interprétation.

3. EVOQ : Objet et fondements théoriques

3.1 Le domaine d'application de départ

Le projet EVOQ trouve son origine dans le champ de l'analyse des organisations et plus particulièrement dans le domaine de l'étude des crises et de la perception des risques au sein de l'organisation (Wallemacq *et al.* 2004 : 13). Face à la crise, les décideurs sont souvent démunis et il n'est pas rare qu'ils contribuent à amplifier la crise par sa mauvaise gestion et des dysfonctionnements humains (liés, par exemple au stress, à des conflits intrapsychiques ou interpersonnels, des problèmes de coordination, etc.) (Jacques *et al.* 1994, Jacques et Gatot 1997).

« L'événement [déclencheur de la crise], le processus qui y mène et les conséquences de la crise peuvent être perçus et interprétés différemment suivant les intervenants. Différents intervenants peuvent qualifier une même crise de normale, de critique, de majeure, d'incident ou encore de catastrophe et ceci, en fonction de leur expérience, de leurs connaissances, de leur vécu et de leur environnement. » (Gatot 2000 : 5).

Pour la gestion préventive des crises, il est dès lors essentiel de porter une « attention particulière à l'étude des structures et des processus cognitifs individuels qui sont à l'origine de la représentation mentale, du traitement de l'information et de la prise de décision » (Gatot 2000 : 5).

A de nombreux égards, le discours libre des décideurs et autres acteurs de la gestion de crise est sans doute la manifestation la plus tangible des représentations mentales et des perceptions individuelles qu'ils mobilisent dans leur gestion. Le ReCCCoM s'est donc intéressé à recueillir des matériaux textuels permettant d'étudier les valeurs, préconceptions, cadres de référence dans lesquels s'ancraient le raisonnement de ces gestionnaires de crise et permettant d'accéder aux cadres conceptuels des locuteurs, à la fois « dans leur singularité individuelle et dans leur appartenance culturelle et sociale » (Moscarola 2001). Cette approche cognitive et de perception adoptée, le ReCCCoM s'est trouvé confronté dès le départ à la nécessité d'un outil d'aide à l'analyse et à la lecture des données textuelles. Face aux données qualitatives collectées au cours d'entrevues et d'observation participantes, et plus généralement face à de grands corpus de données non-structurées, « se pose en effet la question du sens, de la signification des propos, et de l'interprétation des écrits...comment s'y prendre pour en tirer une connaissance, comment garantir à celle-ci une légitimité scientifique ? » (Moscarola 2001). Face à ce problème méthodologique et épistémologique, le ReCCCoM a privilégié la construction d'un outil d'aide à l'analyse, « favorisant la participation active de l'analyste du texte », privilégiant « l'analyse du contenu par rapport à la connaissance purement formelle de la langue » (Duchastel *et al.* 2001 : 8), adoptant une perspective ouverte sur le texte et favorisant le travail herméneutique du chercheur en gardant une forme « naturelle » d'intelligibilité. Le produit de ces réflexions est le logiciel EVOQ, en cours de développement depuis 1997 sous l'impulsion d'Anne Wallemacq et de Jean-Marie Jacques. Divers articles écrits par ces deux auteurs présentant le logiciel sous ses différentes dimensions, la discussion des aspects épistémologiques du programme dans la section 3.2 ci-dessous s'inspire largement de ces descriptions (cfr. notamment Jacques *et al.* 2004, Wallemacq et Jacques 2004, Latiers *et al.*

à paraître, Wallemacq *et al.* 2004). La présentation des fondements opérationnels d'EVOQ dans la section 3.3. constitue néanmoins une synthèse concernant les derniers développements logiciels du programme qui n'ont pas encore fait l'objet d'une description dans un article.

3.2 Fondements épistémologiques

Le traitement des questionnaires, des enquêtes, des entrevues semi-dirigées et des observations sociologiques et plus généralement de grandes quantités d'informations non structurées pose des difficultés particulières. Différentes méthodes d'analyse mettant l'accent sur différentes règles et outils conceptuels existent pour en faciliter le dépouillement. Moscarola (2001) suggère ainsi trois angles d'approche du texte, inspirés de la sociologie et de l'analyse assistée par l'informatique. Les méthodes d'analyse de contenu par exemple, traitent l'information sur base d'une grille d'analyse précise qui sert à quadriller le texte pour en repérer et catégoriser les éléments constitutifs. Une analyse statistique peut être couplée à cette approche du texte pour en faire ressortir les structures statistiquement significatives. S'intéressant aux structures linguistiques statistiquement les plus fréquentes, ces méthodes lexicométriques permettent de déterminer quels sont les usages de termes récurrents chez un auteur et leur fréquence, quels termes sont sur-représentés ou sous-représentés chez un locuteur donné et quels sont les tendances d'un locuteur à associer certains termes ensemble. Comme le note Timimi (1998), cependant, le seul critère de la fréquence des termes ou de segments s'accommode mal notamment d'une réalité linguistique de base : le fait qu'il existe des traits de synonymie entre ces mots. Par exemple, les liens sémantiques entre des termes distincts comme « formation, enseignement, éducation » échappent forcément aux approches de type statistique même si par ailleurs, ces approches peuvent donner des résultats très probants concernant les profils lexicaux ou les réseaux sémantiques caractéristiques d'un texte².

EVOQ privilégie une approche complémentaire à ces outils logiciels de tendance positiviste. Pour réfléchir à la question du sens des propos des locuteurs, le logiciel s'inspire à la fois de principes issus de la phénoménologie et du structuralisme, dans une perspective constructiviste. Dans ce cas, on part non plus du principe qu'un texte exprime pleinement son sens de manière explicite, mais on considère plutôt que les univers de perception du locuteur sont toujours des systèmes de relations. De la même manière que l'image d'un puzzle n'est reconstitué que lorsque toutes les pièces sont organisées en une image signifiante (Piret, Nizet et Bourgeois 1996), l'univers de perception d'un locuteur ne peut être pleinement saisi que lorsqu'on analyse non seulement les sens exprimés explicitement dans un texte mais aussi les sens implicitement portés par la langue (les présuppositions, les évocations, les sens plus ou moins cachés mais qui déterminent l'interprétation de ce qui est dit). Comme le décrit Moscarola (2001), dans une telle approche :

« Il s'agit non plus d'aborder le texte en extension avec l'ambition d'en capter l'essence mais de le saisir en intention et de reconstruire les mondes possibles du

² Voir à ce propos les résultats obtenus par un logiciel comme Alceste, mis au point par Max Reinert (1997)

locuteur. Explorer la pluralité des significations inscrites dans chaque fragment du texte, le faire résonner dans toutes les significations qu'il implique plutôt que raisonner sur le sens qu'il convient de lui attribuer »

Plus spécifiquement, dans cette perspective (post) structuraliste, les significations d'un terme sont considérées non pas comme inhérentes à ce terme mais toujours « relationnelles », « colorées » à la fois par le réseaux des évocations qui peuvent lui être associés ainsi que par les « systèmes de différence » dans lesquels il s'inscrit (voir Wallemacq *et al.* 2004 : 18-25). Comme le notent Wallemacq *et al.* (2004 : 25), le sens se construit dans le champ de force constitué par la définition réciproque des termes. Ils citent Derrida :

« Le concept signifié n'est jamais présent par lui-même et de lui-même... tous les concepts s'inscrivent dans une chaîne ou dans un système à l'intérieur duquel ils font référence à l'autre aux concepts, grâce au jeu systématique des différences » (Derrida 1982, cité dans Hassard 1993 : 14)

Reconstituer l'univers de perception d'un locuteur dans ce type d'approche **structuraliste** revient donc à identifier les principes d'identité et de différence au cœur des représentations qui sont assemblées par lui et de s'intéresser au jeu infini des équivalences « horizontales » au travers desquelles le sens est constitué (Wallemacq *et al.* 2004 : 23). En bref, EVOQ considère que les représentations des locuteurs sont toujours un système organisé d'« oppositions » et de d'« associations » et s'intéresse à la manière dont ces systèmes de relation sont articulés du point de vue du locuteur dans le texte examiné. Le chercheur est donc responsable de la lecture informée du texte et du codage des relations de « conjonctions » (lorsque le locuteur associe certains éléments du texte avec d'autres) et de « disjonctions » (lorsque certains éléments du texte sont opposés à d'autres éléments du texte par le locuteur). Dans le processus de codage, le chercheur est non seulement invité à coder les sens exprimés explicitement (les « sens de surface ») mais aussi à reconstituer les termes opposés probables, qui peuvent être absents ou différés dans le texte (les « sens au rebond ») (Wallemacq *et al.* 2004 : 27). Ces systèmes de différences sont souvent socialement partagés : en examinant les expressions automatiques d'idées par les locuteurs individuels, on accède aussi aux présupposés et termes de référence communs aux groupes sociaux dans lesquels les locuteurs se meuvent (ce que les ethnométhodologues appellent le « code »). Le discours de l'individu porte ainsi la trace des divisions et des relations prévalentes dans les groupes dont il fait partie et dans lesquels il a été socialisé (Wallemacq *et al.* 2004 : 18).

Dans ce contexte, le logiciel est conçu comme une aide pour accéder et mettre en évidence les cadres conceptuels et les termes de références individuels ou collectifs « dans lesquels s'ancrent le raisonnement des membres d'une organisation ou d'un groupe social » (Wallemacq *et al.* 2004 : 13). Il vise à mettre en évidence la manière dont le monde est mis en forme pour un groupe ou un individu donné (Latiers *et al.* 2004), dans quels termes il est constitué (y compris des termes contradictoires (Jacques *et al.* 2004) et invite donc à réfléchir sur la façon dont des termes sont « naturalisés » dans un discours particulier.

3.3.1. Le module « texte »

[illegible]

Figure 1 : Le module « texte »

⁴ Réalisées par Patrick Alexandre et Marc Eloy.

Le module texte supporte toute longueur de texte et est donc apte à recevoir de grands corpus (à titre d'exemple, le chargement d'un texte moyen de 1.300.000 caractères ne prend que quelques secondes sur un processeur performant). On peut bien entendu insérer une collection de textes pour en comparer les contours sémantiques. Sous l'onglet « text », on trouve de plus un accès à un outil de recherche simple (fonction « find »), permettant de connaître le nombre d'occurrences d'un terme et de retrouver celles-ci dans le document en cours d'analyse.

3.3.2. Le module « dictionnaire »

La saisie de termes dans le module « dictionnaire » constitue la première étape du traitement des données avec EVOQ. Comme nous l'avons déjà dit, dans la perspective structuraliste qui sous-tend philosophiquement le logiciel, le travail analytique consiste à mettre en lumière les réseaux d'interrelations tissés par le locuteur entre différents termes dans son discours, en repérant particulièrement les relations de type :

- 1) complémentaires, dans le cas des relations d'*associations* ou de *conjonctions* (qu'est-ce que le locuteur range avec quoi ?) ;
- 2) antagonistes, dans le cas des relations d'*opposition* ou de *disjonctions* (qu'est-ce que le locuteur oppose à quoi ?) ;

Ce travail herméneutique n'est pas automatisé. Seul l'analyste en effet, est capable de reconstruire l'épistémologie sous-jacente à un discours donné. Néanmoins, différents outils d'aide à la structuration des relations existent dans la dernière version du programme (nous y reviendrons plus bas).

Dans le cadre d'EVOQ, la saisie des relations de conjonctions et de disjonction se fait sur base des principes de la méthode d'analyse structurale telle qu'elle est décrite dans Piret, Nizet et Bourgeois (1996). Tout utilisateur d'EVOQ s'imprègne de facto de ces principes par la seule utilisation du logiciel qui induit automatiquement une analyse de type structurale.

Concrètement, la partie dictionnaire comporte trois fenêtres, situées dans la partie inférieure de l'interface: une fenêtre à gauche (la liste alphabétique), une fenêtre au centre (le dictionnaire proprement dit) et un menu pour agir sur le dictionnaire (colonne de droite).

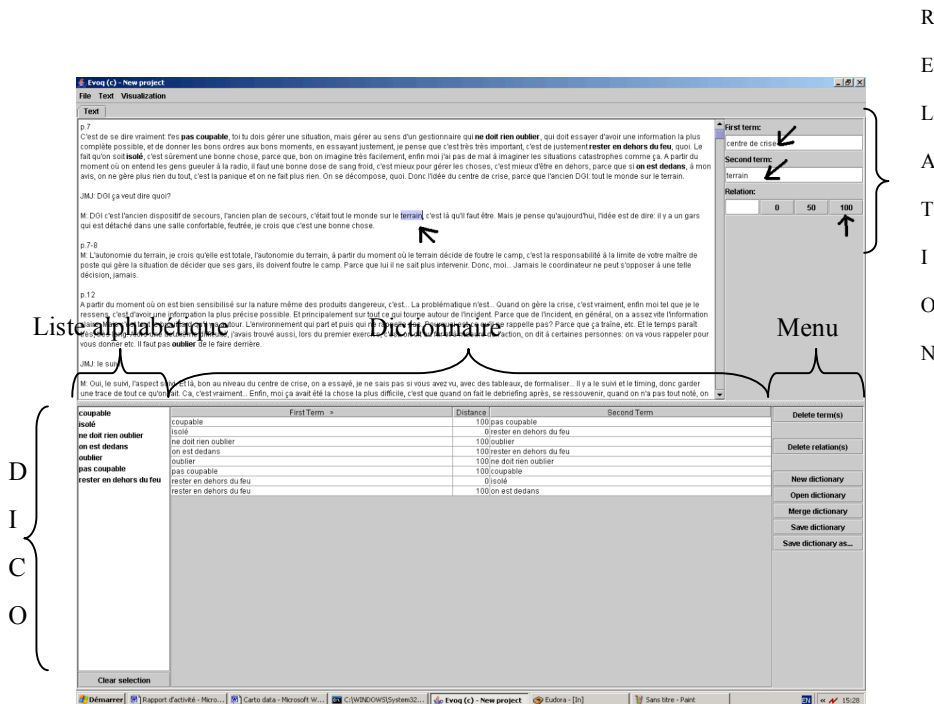


Figure 2 : Le module « dictionnaire »

Il est possible d'introduire un terme dans le dictionnaire de trois manières différentes ce qui rend le système souple et ergonomique. Une première possibilité consiste à double-cliquer sur un mot de la partie « texte ». On peut aussi entrer les termes via le clavier, ou encore sélectionner un terme ou une série de termes en les sélectionnant avec la souris. Dans ces trois cas de figure, les termes s'affichent alors en haut à droite de l'interface, dans l'une des deux boîtes de dialogue (« first term » ou « second term ») de la partie « relations ».

Lorsque deux termes ont été sélectionnés parce qu'ils sont mis en relation par le locuteur, l'analyste va déterminer si cette relation est une relation de type antagoniste ou de type complémentaire. Il appuiera sur le bouton « 100 » pour indiquer une relation d'opposition ou de disjonction entre deux termes et sur le bouton « 0 » pour souligner au contraire une relation de conjonction ou d'association entre les termes (il est à noter que ces nombres font référence à des distances purement arbitraires).

Une fois les relations encodées, elles apparaissent dans la fenêtre du milieu de la partie « dictionnaire » de la visualisation. Celle-ci constitue le « dictionnaire proprement dit », c'est-à-dire la liste des termes et leurs associés ou opposés. La liste des termes déjà présents dans le dictionnaire s'affiche également sous la forme d'une liste alphabétique dans la fenêtre de gauche de la partie « dictionnaire ». Cette liste alphabétique permet de retrouver facilement quels mots ont déjà été encodés. Il est également possible d'introduire un terme dans le dictionnaire en le sélectionnant dans la liste alphabétique puis en le collant dans la partie « relations ». Supprimer un mot de la liste alphabétique se fait en sélectionnant le terme et en l'effaçant via le bouton « clear selection ».

Une dimension importante du dictionnaire à noter est la suivante :

« Toute entrée ou modification du texte suscite automatiquement une reconnaissance des termes présents dans le dictionnaire ; les termes sont

alors affichés en gras. L'utilisateur peut librement sélectionner les unités sémantiques du texte et les ajouter au dictionnaire, ainsi que les relations les concernant. » (Wallemacq *et al.* 2004 : 33).

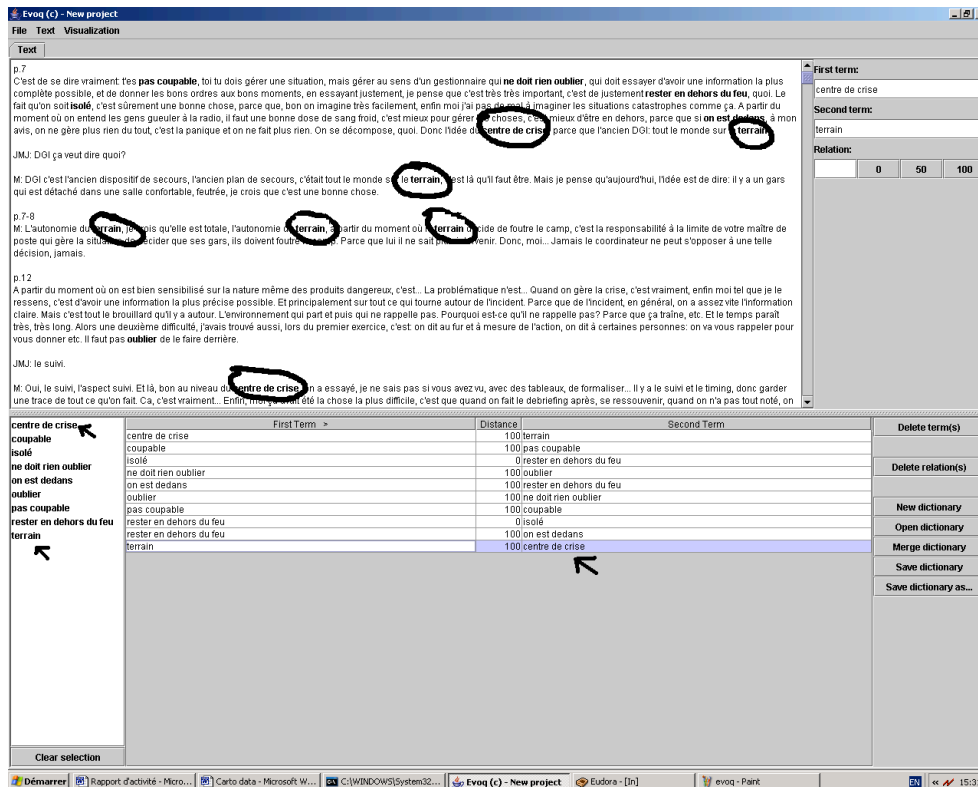


Figure 3 : La procédure de reconnaissance automatique des termes du dictionnaire

Cette procédure de reconnaissance automatisée des termes déjà présents dans le dictionnaire devient un avantage lorsqu'on commence à avoir plusieurs dictionnaires concernant un domaine donné.

« L'interface est ici conçue pour constituer les dictionnaires sémantiques par accumulation : à chaque nouveau texte introduit dans un contexte donné, les termes déjà présents dans le dictionnaire sont mis en évidence, et l'utilisateur peut progressivement enrichir l'ensemble des relations sémantiques. Ainsi, l'identification du réseau sémantique d'un texte introduit est à chaque utilisation plus automatique et complète ». (Wallemacq *et al.* 2004 : 33).

En ce qui concerne la partie « dictionnaire proprement dit » plusieurs éléments sont à noter. Tout d'abord, les termes peuvent être classés de trois façons différentes : par ordre alphabétique croissant, par ordre alphabétique décroissant ou par ordre d'entrée des termes dans le dictionnaire. La navigation entre les différentes formes de classement se fait en appuyant sur les signes « > » et « < » situés en haut des colonnes constituant le dictionnaire. On peut choisir aussi de grouper tous les termes associés, ou tous les termes opposés. Cela

se fait en cliquant sur la tête de colonne « distance ». Les relations entre les termes étant réciproques, elles apparaissent toujours dans les deux directions dans le dictionnaire :

Noir	100	Blanc
Blanc	100	Noir

Figure 4 : Réciprocité des relations

On peut enfin, via le menu interactif situé dans la colonne de droite de l'écran « dictionnaire » effectuer les manipulations suivantes : supprimer un ou plusieurs termes (via le bouton « delete term(s) ») ; supprimer une ou plusieurs relations (via le bouton « delete relation(s) »), enregistrer le dictionnaire (via le bouton « save dictionary (as) »). On peut également ouvrir un dictionnaire existant (via le bouton « open dictionary ») ou créer un nouveau dictionnaire (via le bouton « new dictionary »).

On le voit, la procédure de constitution du dictionnaire vise donc la participation active de l'analyste mais au fur et à mesure que les dictionnaires sont constitués, les procédures de reconnaissances automatique vont croissant.

Constituer le dictionnaire est déjà une forme d'analyse mais celle-ci va trouver son véritable sens dans le module de visualisation. Décrivons-le à présent.

3.3.3. Le module de « visualisation »

Le cœur du programme réside dans le module de « visualisation » des dictionnaires sémantiques. La visualisation vise à apporter une intelligibilité nouvelle dans la lecture du texte. Elle vise à être un outil pour sortir des schémas linéaires classiques de lecture et s'intéresse au « réseaux de sens rhétoriquement et socialement construits » dans le texte (Blommaert et Verschueren 1998). Elle va permettre d'identifier comment s'organisent les réseaux de sens pour le locuteur. Les cartes de relations produites par le module de visualisation d'EVOQ sont de plusieurs types : 1) représentations du réseau sémantique sous formes de cartes dynamiques en deux dimensions qui s'auto-organisent sur base du dictionnaire sémantique, 2) cartes bidimensionnelles où les lignes d'oppositions entre termes et concepts sont marquées par des traits délimitant les territoires sémantiques à l'intérieur de la carte, 3) paysages sémantiques en trois dimensions, 4) constellations dynamiques en 3 dimensions des relations entre les termes—voir ci-dessous (cfr. Wallemacq et Jacques 2000 pour une discussion des implications liées à différentes formes de visualisation).

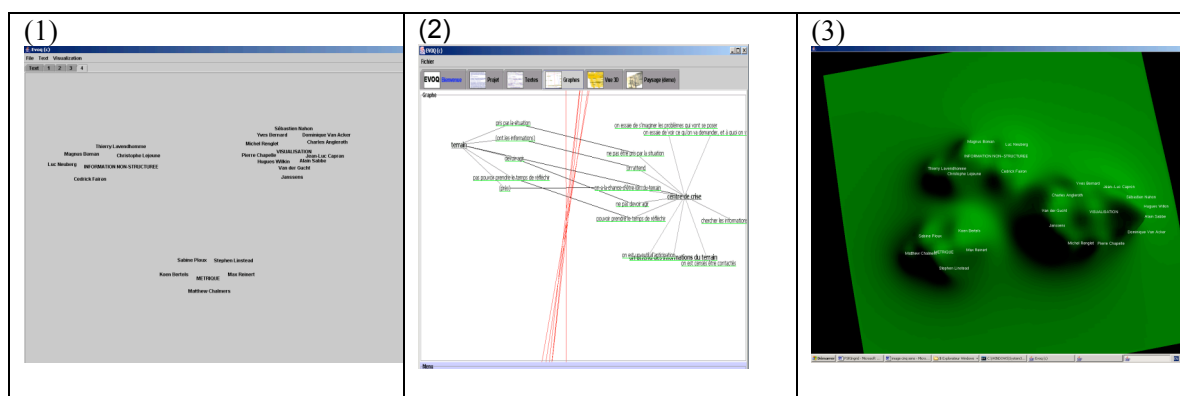


Figure 5: Le module de “visualization”

Visualisation 2-D

Pour créer une visualisation, il suffit d'ouvrir le menu déroulant situé sous l'onglet « visualization ». Cette opération donne le choix entre deux possibilités : créer une visualisation en 2-D (« new 2D ») ou une visualisation 3-D (« new 3D »). Choisir dans le menu « new 2D » fait apparaître une boîte de dialogue où l'utilisateur est invité à donner un nom à la visualisation qu'il crée. Le système permet d'attribuer soit un nom soit un chiffre ou un nombre pour faire référence à la visualisation. Lorsque de nombreuses visualisations sont créées, ce processus d'indexation permet de retrouver rapidement les visualisations déjà enregistrées. Une fois qu'un nom est attribué et que l'on a pressé sur « OK », le programme crée automatiquement la visualisation.



Figure 6 : Les visualisations comme graphes dynamiques

Une visualisation en 2-D « contient initialement l'ensemble des termes reconnus par le dictionnaire dans le texte. Il s'agit donc de la région du réseau sémantique « occupée » directement par le texte » (Wallemacq *et al.* 2004 : 35).

Les unités de base du graphe sont :

- 1) des mots ou des expressions
- 2) des relations
- 3) une dynamique d'organisation

Lorsque l'application « visualization » est lancée, les mots encodés dans le dictionnaire forment d'abord une masse indifférenciée de termes en mouvements (une représentation globalisante et non-structurée de l'univers de référence contenu dans le texte, comme par exemple celle reproduite sur le schéma ci-dessus en T1). Rapidement, ces termes se structurent autour des relations de conjonctions/disjonctions qui ont été encodées dans le dictionnaire, si bien qu'après quelques secondes on peut discerner les pôles d'attractions entre les termes (comme en T4 par exemple).

Les relations de conjonctions sont marquées visuellement du fait que des mots qui sont associés dans le dictionnaire s'attirent l'un l'autre sur le graphe. Si on sélectionne un terme et qu'on le déplace dans le graphe, tous les mots qui y sont associés se déplacent simultanément dans son sillage. Les mots opposés par contre se repoussent de telle manière que sur le graphe on ne pourra jamais se faire rejoindre deux termes antagonistes. Les lignes d'attraction et de répulsion entre les termes structurent donc dynamiquement les graphes.

Ces représentations visuelles permettent aussi de saisir intuitivement et directement ce qui constitue le *texte*, son *contre-texte* (tous les mots opposés à un premier ensemble de termes) et le *sous-texte* (implicitement codé). Il est possible en effet d'introduire des termes qui ne sont pas directement présents dans le texte mais dont on peut inférer qu'ils soient opposés à des termes donnés. Ces termes implicites apparaissent alors en italique dans le graphe.

Enfin, une série de manipulations sont possibles sur le graphe à partir de la fenêtre « visualization ».

En mode « *position* », un simple click du bouton gauche avec la souris sur un terme sélectionné permet en faisant glisser le terme de le stabiliser dans le graphe dynamique. Le terme cesse alors de se déplacer dans le graphe et un astérisque apparaît qui signale que le terme a été fixé. Un simple clic sur le terme le libère à nouveau. Cette fonction permet de jouer avec les termes pour organiser les graphes spatialement.

En mode « *highlight* », l'utilisateur peut mettre en évidence des termes clés pour son analyse. Il peut organiser librement l'information pour en faire ressortir les structures les plus significatives en fonction de son objectif d'analyse. Cette fonction facilite aussi la présentation des résultats (sous forme d'article ou de présentation power point).

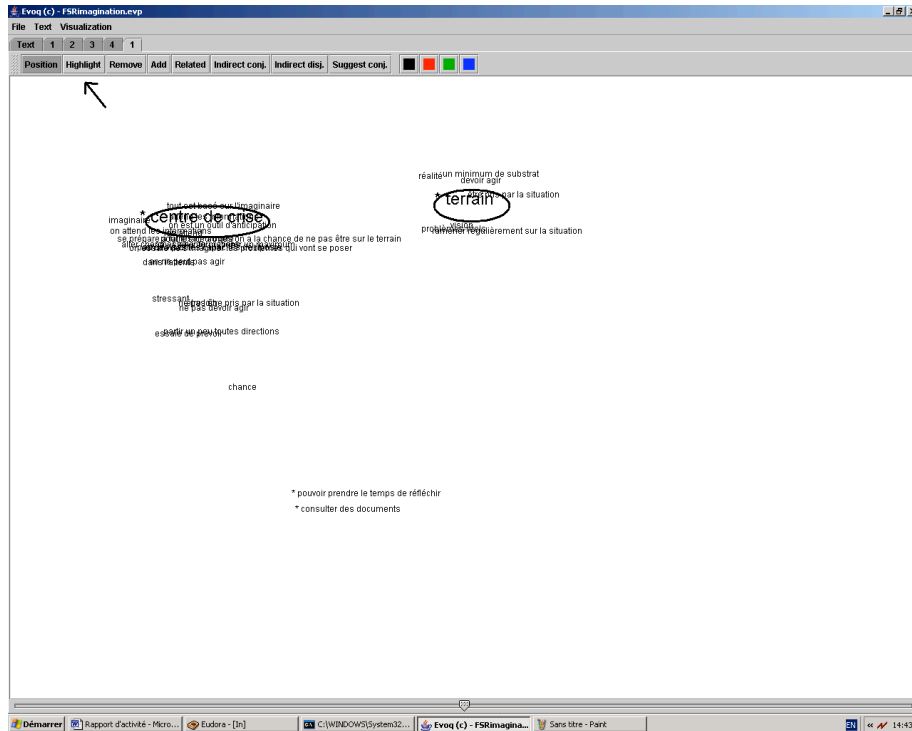


Figure 7 : Le mode « highlight »

En mode « *remove* », l'utilisateur peut supprimer librement tout terme dans le graphe en cliquant dessus avec le bouton gauche de la souris.

Le bouton « *add* » permet d'intervenir dans le graphe pour y ajouter un terme. L'utilisateur sélectionne d'abord un mot dans le graphe (par exemple « centre de crise ») puis introduit un nouveau terme via la boîte de dialogue qui s'ouvre lorsqu'il est sous l'onglet « *add* » (par exemple, « gestion »). Une seconde boîte de dialogue l'invite à préciser si le terme ajouté est à considérer comme un opposé ou un associé du terme initial (par exemple, en lui imputant la distance « 0 » ou « 100 »). Le nouveau terme s'affiche alors dans le graphe et se positionne en fonction de la distance choisie.

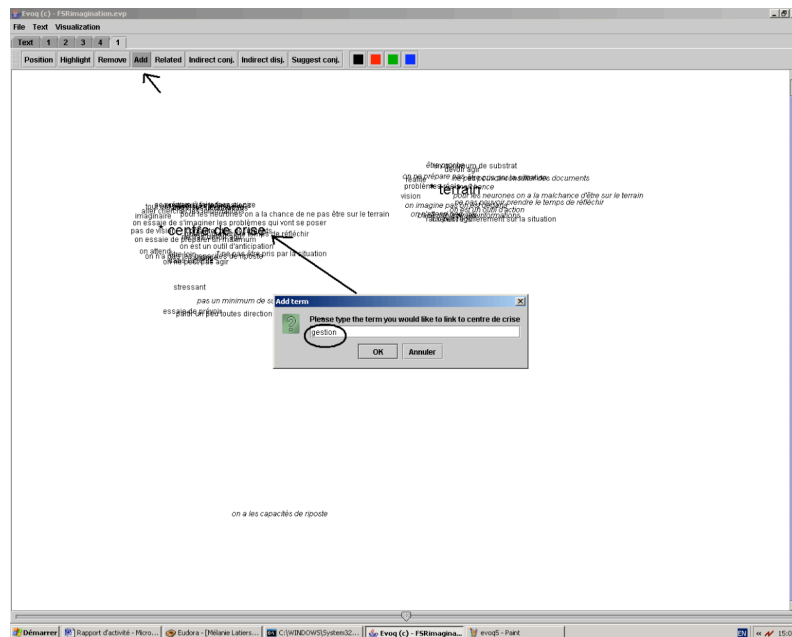


Figure 8 : Le mode « add »

En mode « *related* », l'utilisateur peut choisir un terme dans le graphe et faire apparaître tous les termes qui lui sont implicitement associés. Il « peut choisir de « découvrir » les connotations d'un terme donné ou de tous les termes du graphe simultanément, et répéter l'opération autant de fois que désiré » (Wallemacq et al. 2004 : 35). Les termes s'affichent alors en italique.

Les boutons « indirect disjonction » (*indirect disj.*) et « indirect conjonction » (*Indirect disj.*) « permettent l'inférence de disjonctions ou de conjonctions par transitivité. En cliquant sur un terme dans ce mode, une boîte de dialogue s'ouvre suggérant des relations possibles avec d'autres termes par rapport à un terme donné. L'utilisateur peut sélectionner dans la liste des propositions quelles relations il souhaite ajouter au dictionnaire, et les entrer dans celui-ci via le bouton « add to dictionary ». Ici le logiciel calcule automatiquement en fonction des liens entre des termes introduits dans le dictionnaire si d'autres termes ne pourraient pas y être associés ou opposés.

Enfin, en mode « couleurs » dans la barre d'outil, il est possible d'attribuer des couleurs aux mots d'un dictionnaire. Cette fonction permet de mettre en évidence certains termes au cours de l'analyse. Elle permet aussi, par exemple, de souligner quels termes proviennent du discours de tel ou tel locuteur dans le cas d'une fusion de dictionnaires. Dans ce contexte, il apparaît en effet essentiel d'être capable de déterminer quels termes doivent être attribués à quels locuteurs. Lorsque cette fonction sera automatisée (ce qui est en cours de préparation), l'utilisateur pourra ainsi non seulement mettre à plat les termes de référence dans lesquels raisonnent un locuteur et les comparer avec les termes de référence d'un autre locuteur pour un même domaine. Ou pourra également observer comment un

code peut en envahir un autre, quels sont les effets d'intertextualité ou de colonisation d'un discours par un autre.

On peut également accéder à chacune des fonctionnalités décrites ci-dessus (*highlight, remove from graph, related terms, add term to graph, free all, all related terms*) via le menu contextuel accessible en cliquant sur le bouton droit de la souris lorsqu'on sélectionne un terme dans le graphe.

Visualisation en paysages et visualisations 3-D

Pour créer une visualisation en « paysage sémantique », il suffit de sélectionner la fonction « landscape » dans le menu contextuel accessible à partir de la visualisation en 2 dimensions que nous venons de décrire. La sélection de la fonction « landscape » fait apparaître un paysage en trois dimensions où les disjonctions entre les termes sont représentées sous la forme de vallées, tandis que l'agrégat des termes associés donne lieu à la création de « sommets ».

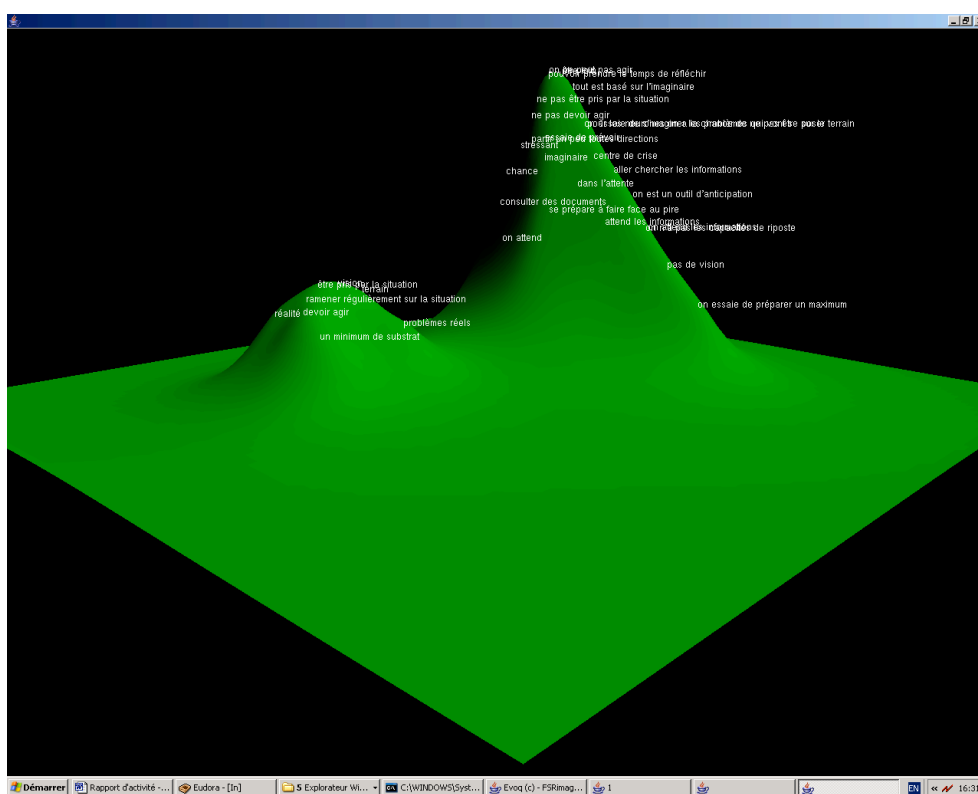


Figure 11 : La visualisation en paysages sémantiques

Enfin, l'onglet « visualisation » permet de créer aussi un autre type de représentations : la représentation en 3D. Il s'agit ici d'une représentation présentant les termes de référence dans lesquels locuteur raisonne sous forme de constellation mouvante :

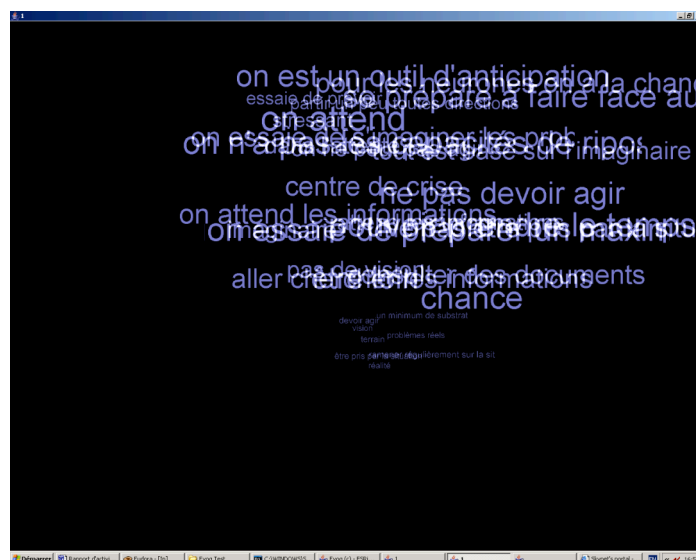


Figure 12 : La visualisation en constellation

L'utilisateur peut se déplacer à l'intérieur de ce graphe, le faire tourner sur son axe.

Chacune de ces visualisations a pour fonction de stimuler la créativité du chercheur (voir Latiers *et al.* à paraître) en lui permettant littéralement de « « mettre à plat » les termes de référence dans lesquels le locuteur raisonne » (Wallemacq *et al.* 17). Via la médiation de ces visualisations, l'utilisateur a accès au « code » d'une manière proche de l'expérience vécue tout en l'abordant à un plus grand niveau de différenciation et d'analyse que lors de la lecture initiale du texte. Le code apparaît sous la forme d'un milieu englobant ou accidenté, d'une série de relations ou de reliefs, chacune de ces métaphores induisant un jeu avec le texte initial obligeant le chercheur à interroger son contenu sous de nouveaux auspices. Dans cette prise de distance avec le texte, se situe la possibilité d'ouvrir le champs des interprétations.

4. EVOQ : Quelques résultats

Pour conclure, nous avons choisi de souligner quelques domaines de recherche facilités par l'utilisation du logiciel EVOQ et qui ont déjà été mis en lumière dans plusieurs articles déjà publiés par les membres du ReCCCoM (cfr, par exemple, Wallemacq *et al.* 2004, Latiers *et al.* à paraître, de Saint-Georges *et al.* 2004, et les autres articles de ces mêmes auteurs déjà cités dans cette note).

- Un premier domaine est celui du repérage du *texte*, du *contre-texte* et du *sous-texte* propres à un discours étudié. Les représentations visuelles générées par EVOQ permettent de repérer très clairement la manière dont les idées d'un locuteur sont organisées et quels cadres conceptuels sous-tendent son discours et ses actions. Un groupe de mots qui s'auto-organisent dans une région topographiquement distincte du graphe constituera par exemple le texte, tandis que les termes s'organisant géographiquement à l'opposée du texte informeront l'utilisateur sur le contre-texte. Les mots apparaissant en italique formeront le sous-texte de l'ordre de l'implicite, articulé par le locuteur. A travers cette mise en lumière, on pourra ainsi « chercher à interroger la culture propre d'un groupe ou d'une organisation pour mieux la comprendre et la faire évoluer » (Wallemacq *et al.* 2004).
- Un second domaine de recherche possible concerne le repérage des *contradictions* dans un texte. Grâce à l'utilisation d'EVOQ, il est facile de constater les moments où un locuteur oppose des idées ou des termes qu'auparavant il avait associé ou vice-versa. Dans le domaine des organisations et de la gestion des crises, ce repérage des contradictions par exemple, est fort utile pour mettre le doigt sur des modes de perception du risque et de la crise qui pourraient mener à des dysfonctionnements dans la gestion de celle-ci.
- Au-delà des contradictions, certaines oppositions dans les termes utilisés relèvent parfois véritablement du *paradoxe*. Dans ce cas, l'organisation des termes de référence conduit alors à une impasse conceptuelle (Jacques *et al.* 2004). Le repérage des paradoxes et des solutions créatives mises en place

pour pour les acteurs pour les dépasser peuvent elles aussi être mises en lumière grâce à EVOQ.

- En plus des distinctions que EVOQ permet de lever et de distinguer, le logiciel permet aussi de voir ce qui « va ensemble », c'est-à-dire d'analyser la manière dont des univers de discours s'articulent de manière interne ; en même temps qu'ils se constituent par des choix signifiants parmi un ensemble d'autres sélections qui auraient été signifiantes. Foucault utilise le terme « ordres du discours » pour faire référence à ces épistémologies qui s'articulent ainsi dans le discours. Or, il n'existe pas vraiment de grille d'analyse qui permette de relever ces éléments constitutifs d'un ordre du discours sans réifier les termes qui le compose. EVOQ permet de mettre en lumière ces « nœuds de significations », ces « moments d'articulation historiquement situés » qui se cristallisent dans le discours.
- Grâce à EVOQ, on peut aussi saisir empiriquement les phénomènes d'*intertextualité* : quels sont les éléments du code qui transcendent la structure des discours individuels ? De quelle manière des mondes conceptuels se construisent-ils sur d'autres mondes ou se font-ils l'écho d'autres termes de référence. En observant comment certains termes colonisent un discours, l'envahissent ou au contraire disparaissent d'un discours à l'autre, on peut s'interroger sur les relations que les discours individuels entretiennent avec des « discours sociaux », des « cadres représentationnels », qui les précèdent et les englobent.
- EVOQ permet enfin de saisir finement les relations de *polysémie* par l'observation du sens qu'un mot acquière en relation avec les autres termes apparaissant dans son sillage.

EVOQ promet donc d'être un outil puissant pour mettre en évidence les cadres cognitifs sous-tendant les actions et les prises de décision des acteurs de l'organisation quels que soient le contexte organisationnel étudié.

Références

COOPER, R. 1989. Modernism, post modernism and organizational analysis 3 : The contribution of Jacques Derrida. *Organization Studies* 10/4: 479-502.

DUCHASTEL, Jules, DUPUY, Luc et François DAOUST. 1989. Système d'analyse de contenu assistée par ordinateur (SACAO). In : Actes du colloque 'La description des langues naturelles en vue d'applications linguistiques'. Québec : centre international de recherche sur le bilinguisme, 1989 : 197-210.

GATOT, L. 2000. Crise et cognition. Conception d'une gestion préventive. Thèse de doctorat. Faculté des sciences économiques, sociales et de gestion. Département des Sciences de gestion. Research Center for Crisis and Conflict Management. Presses Universitaires de Namur.

GIBBS, Graham R., FRIESE, Susanne et Mangabeira, Wilma C. 2002. The use of new technology in qualitative research. Introduction to Issue 3(2) of FQS [35 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social research*. [On-line

Journal], 3(2). Available at: <http://www.qualitative-research.net/fqs-texte/2-02/2-02hrs-g-e.htm>

JACQUES, J.M et L. GATOT. 1997. De l'incident à la catastrophe : un modèle organisationnel. *Urgences* 1997 ; XVI : 24-30, Elsevier, Paris

JACQUES, J.M., GATOT, L., DECHARNEUX, T., SEPULCHRE, O., 1994. Les facteurs aggravants et les crises technologiques dans le secteur de la chimie, *Environnement et Société*, 12, pp.30-55, 1994.

JACQUES, J-M, LAURENT, N., WALLEMACQ Q. 2004. Paradoxes, dilemmes et contradictions. Une mise en lumière au moyen du logiciel EVOQ in Purnelle G, Fairon, C, Duster A (éds). *Le poids des mots. Actes des 7eme journée internationales d'analyse statistique des données textuelles (JADT 04)*.

JENNY, Jacques. 1997. Méthodes et pratiques formalisées d'analyse de contenu et de discours dans la recherche sociologique française contemporaine. Etat des lieux et essai de classification. In : le Bulletin de Méthodologie Sociologique (B.M.S., Numéro 54, Mars 1997, p.64-112. LASMAS-IRESCO (CNRS), 59, rue Pouchet, 75017 PARIS.

LATIER, M, BRUYNINCKX V., WALLEMACQ A., LAURENT N., VAN ASSCHE, V. et JACQUES J-M.. A paraître. Software, Springboard, Creativity. *International Studies of Management and Organization*.

LEE Raymond M. et FIELDING Nigel G. 1995. User's experience of qualitative data analysis software. In : Udo Kelle (ed.) *Computer-aided qualitative data analysis : theory, methods and practice* (pp.29-40), London : Sage, 1995

LEJEUNE, Christophe. 2004. Méthodes qualitatives : Panorama. <http://www.smess.egss.ulg.ac.be/lejeune/logiciels/panorama.html>

MOSCAROLA, Jean. 2001. Contributions des méthodes de l'analyse qualitative à la recherche en psychologie interculturelle. Sphinx et MCA. VIIIe Congrès International de l'Aric Genève.

PIRET, Anne, Jean NIZET et Etienne BOURGEOIS. 1996. L'analyse structurale. Une méthode d'analyse de contenu pour les sciences humaines. DeBoeck Université.

REINERT, M. 1997. Les « mondes lexicaux » et leur « logique » à travers l'analyse statistique de divers corpus. *Lexicometrica*, numéro 0, 1997.

TEIL, Geneviève et LATOUR, Bruno. 1995. The Hume-Machine: can association networks do more than formal rules ? In: SEHR, volume 4, issue 2: *Constructions of the mind*. June 1995. Available [http: http://www.stanford.edu/group/SHR/4-2/text/teil-latour.html](http://www.stanford.edu/group/SHR/4-2/text/teil-latour.html)

THOMPSON, Robert. 2002. Reporting the Results of Computer-assisted Analysis of Qualitative Research Data [42 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], 3(2). Available at: <http://www.qualitative-research.net/fqs-texte/2-02/2-02thompson-e.htm>

TIMIMI, Ismaïl. 1998. Gouverner par les sondages et analyser par la paraphrase. Méthode TIAD. Papier présenté aux JADT 1998.

WALLEMACQ A, ET JACQUES J.M., 2004. Language and landscape, towards new architectures of thought. In: Stephen Linstead and Alison Linstead (eds). Thinking organization. Routledge.

WALLEMACQ A. et JACQUES, J.-M. 2000. Landscapes of words, in : Combes Ch., Grant, D., Keenoy T. et Oswick Cl. Organizational Discourse: word-views, work-views and world-views, proceedings of the conference held at The Management Center, King's College, University of London, 2000, July 26-28, London: KMPC, 275-276.

WALLEMACQ, Anne, JACQUES J.-M, BRUYNINCKX, Vincent. 2004. Dans le sillage des mots...EVOQ, Logiciel de cartographie cognitive. Namur : Presses Universitaires de Namur.