



Comment utiliser le programme d'analyse sémantique Sémato

Marissa Trudel^a

^aÉcole de Psychologie, Université d'Ottawa

Abstract ■ Le logiciel Sémato développé par Pierre Plante à l'Université du Québec à Montréal permet de compléter des analyses sémantiques avec de très larges données textuelles. Cet article a comme but d'être un guide pour faciliter l'utilisation du logiciel par de futurs chercheurs. Des instructions étape par étape sont fournies pour (1) la préparation du texte avant l'analyse; (2) l'ouverture d'un nouveau projet; (3) l'identification, la modification et l'élimination des thèmes; et (4) certaines analyses. Ces explications sont supportées par des figures d'une analyse précédente qui illustrent quelles étapes suivre sur le logiciel.

Keywords ■ Semantic analyses; book content; statistics books. **Tools** ■ Sémato.

marissa.t0799@gmail.com

[10.20982/tqmp.19.1.p014](https://doi.org/10.20982/tqmp.19.1.p014)

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Reviewers
■ One anonymous reviewer.

Introduction

Le logiciel Sémato a été développé par Pierre Plante à l'Université du Québec à Montréal (PLANTE et collègues, 2013). Sémato a été créé pour permettre l'analyse syntaxique et sémantique de textes en français et en anglais. Dans cet article, nous fournissons des instructions sur la façon d'utiliser le logiciel pour effectuer différentes analyses, mais d'abord, certaines définitions pertinentes sont nécessaires pour bien comprendre le fonctionnement du logiciel.

Définitions

Lemme : L'unité de base d'un dictionnaire. Les mots sont parfois au pluriel, les verbes conjugués. Le lemme est le mot sans aucune conjugaison. Par exemple, "chat" est un lemme, et les mots "chat", "Chat" (majuscule) et "chats" (pluriels) sont des variantes du lemme "chat". Idem pour les verbes, "manger" est un lemme, et "manger", "mange", "mangent", "mangèrent", etc. sont tous des variantes du même lemme. L'étape qui consiste à passer au travers de tous les mots d'un texte pour trouver les lemmes de chaque mot s'appelle la "lemmatisation" et se fait de façon automatique.

Thème : Le thème est plus ou moins synonyme de "concept" ou de sens du mot. Certains mots sont ambigus, car ils peuvent avoir plus d'un sens. Par exemple, les

lemmes "manger", "ingurgiter" et "avalier" sont tous les trois liés au sens de "MANGER". Ces trois lemmes sont donc des "ingrédients" du concept "MANGER".

Ingrédients : Lemmes qui sont liés à un concept.

Thème non étanche : Thème dont les ingrédients ne lui sont pas exclusifs. Par exemple, le concept québécois "écœurant" est un thème qui veut dire "dégoutant", mais aussi "épatant!". Le thème écœurant n'est pas étanche, car ces ingrédients se retrouveront aussi dans au moins un autre thème.

Thème étanche : thème dont les ingrédients ne se retrouvent pas dans aucun autre thème. Je ne crois pas que cela existe dans les langues parlées.

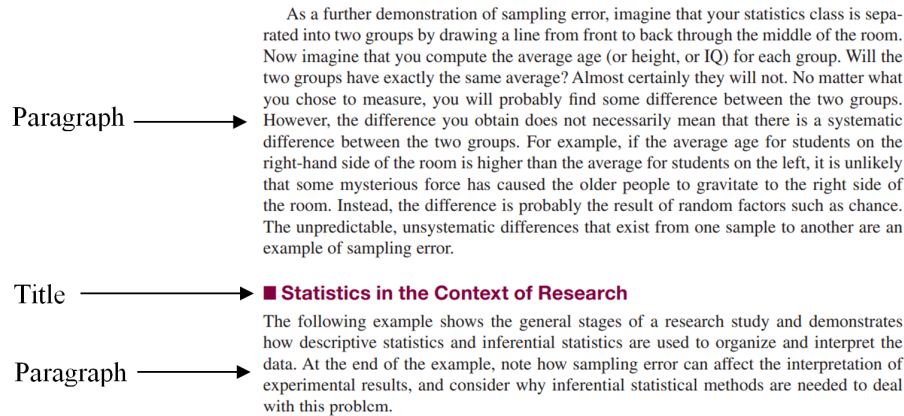
Préparer les textes avant l'analyse

Avant l'analyse, Sémato exige que chaque manuel, disponible sous forme de document PDF soit copié et codé dans un document en texte brut (par exemple, en utilisant Notepad++, Ho, 2023, pour pouvoir identifier les titres et les paragraphes qui sont analysés comme des unités individuelles par le logiciel. On peut faire la distinction entre un titre et un paragraphe, car un titre est souvent en caractère gras ou d'une différente couleur et les paragraphes sont marqués par un alinéa (Figure 1).

Lors du codage des textes (ici, un manuel de statistiques), chaque manuel débute avec les mêmes quatre lignes de code (Figure 2). La première ligne de code est *ntxt=x, qui indique une nouvelle unité d'analyse,



Figure 1 ■ Identification de titre et paragraphe (extrait du manuel *Statistics for the Behavioral Sciences*, 10e édition (2019), par Gravetter & Wallnau, p. 7).



soit un titre ou un paragraphe. La deuxième ligne de code `&type=titre`, précise que l'unité d'analyse est le titre des unités suivantes. Ce code est utilisé pour indiquer tout le titre à travers les manuels. Si l'unité d'analyse est un paragraphe au lieu d'un titre, la ligne de code utiliser pour indiquer tous les paragraphes dans les manuels est `&type=corps`. La troisième ligne de code `&livre=livre1`, indiquent que toutes les unités suivantes sont des données copiées du premier manuel. Lors du codage du deuxième manuel, le code sera `&livre.=livre2`, etc. En plus de pouvoir réaliser des analyses sémantiques par manuel, Sémato est également capable de compléter des analyses par chapitre, indiquer par la quatrième ligne de code, `&chapitre=1`. La cinquième ligne dans la figure est le titre du chapitre tel qu'il est écrit dans le manuel. La Figure 3 montre un extrait d'un manuel et sa version codée dans Notepad++. Une fois qu'une unité est complètement codée, une ligne vide est laissée entre la dernière ligne du paragraphe et la ligne de code suivante.

Ouvrir un Projet

Pour ouvrir un projet, il faut aller sur le site <https://se-mato.uqam.ca/> puis choisir "Sémato texte". Dans la barre de menu en haut, choisir "Projet" (Figure 4). Ensuite il faut choisir "Activation de projet" (Figure 5). Ne cliquez pas sur "Ouverture de projet" qui crée un projet et efface le projet existant. Après ceci, le logiciel demande pour le nom et mot de passe du projet qui est fourni par le gestionnaire du site web (Pierre Plante; Figure 6). Une fois que le nom et mot de passe sont soumis, le message "votre projet est activé" apparaît et il faut cliquer sur le menu "Projet" pour accéder aux données. L'entrée de données a été effectuée par le gestionnaire du site web.

Identification des Thèmes

Lors de l'analyse initiale du corps de texte, Sémato propose de façon automatique un certain nombre de thèmes. Pour voir les thèmes, il faut utiliser le menu "Thème" en haut de page, ce qui vous amène à un tableau contenant tous les thèmes extraits par Sémato (Figure 7).

La colonne à gauche du tableau dans la figure ci-dessus contient les thèmes identifiés, qui débutent par l'indicateur "th p-". En réalité, l'indicateur n'est pas actuellement nécessaire, mais aide à tenir compte des modifications à venir. La colonne à droite identifie les ingrédients de ces thèmes. La colonne du milieu contient la fréquence de ce lemme dans le corps du texte analysé.

Modification et Élimination de Thèmes

Pour débiter la modification ou l'élimination d'un/des thème(s), il faut cliquer sur "Repartir" dans les options sous la barre de menu (Figure 8). Ceci mène à une nouvelle page où il est possible de modifier le nom des thèmes en changeant les composants dans les cases blanches du tableau (Figure 9). Si on efface le tout (nom et ingrédients), ce thème disparaîtra du projet.

Pour ajouter un nouveau thème, il faut aller en bas de page pour retrouver le tableau intitulé "Nouveau Thème" (Figure 10). Il est seulement possible d'ajouter trois nouveaux thèmes à la fois avant d'être obligé de "Soumettre" (Figure 11) les changements et rouvrir le sous-menu Répartir pour rajouter à nouveau des thèmes. Attention de ne pas cocher la case pour fusionner les thèmes.

Lors de n'importe quelle modification, addition ou éliminations, toute la base de données n'est plus à jour et il faut refaire une indexation.



Figure 2 ■ Extrait de Notepad++ des codes requis pour le codage d’un premier manuel.

```
*ntxt=x
&type=titre
&livre=livre1
&chapitre=1
Chapter 1. Introduction to statistics
```

Figure 3 ■ Extrait d’un manuel (gauche) et sa version codée (droite) (l’extrait est du manuel intitulé Statistics for the Behavioral Sciences, 10e édition (2019), par Gravetter & Wallnau, p.10)

1.2 | Data Structures, Research Methods, and Statistics

LEARNING OBJECTIVES 4. Differentiate correlational, experimental, and nonexperimental research and describe the data structures associated with each.

5. Define independent, dependent, and quasi-independent variables and recognize examples of each.

■ Individual Variables: Descriptive Research

Some research studies are conducted simply to describe individual variables as they exist naturally. For example, a college official may conduct a survey to describe the eating, sleeping, and study habits of a group of college students. When the results consist of numerical scores, such as the number of hours spent studying each day, they are typically described by the statistical techniques that are presented in Chapters 3 and 4. Non-numerical scores are typically described by computing the proportion or percentage in each category. For example, a recent newspaper article reported that 34.9% of Americans are obese, which is roughly 35 pounds over a healthy weight.

■ Relationships Between Variables

Most research, however, is intended to examine relationships between two or more variables. For example, is there a relationship between the amount of violence in the video games played by children and the amount of aggressive behavior they display? Is there a relationship between the quality of breakfast and academic performance for elementary school children? Is there a relationship between the number of hours of sleep and grade point average for college students? To establish the existence of a relationship, researchers must make observations—that is, measurements of the two variables. The resulting measurements can be classified into two distinct data structures that also help to classify different research methods and different statistical techniques. In the following section we identify and discuss these two data structures.

```
*ntxt=x
&type=titre
1.2 Data Structures, Research Methods, and Statistics.
```

```
*ntxt=x
&type=titre
Individual Variables: Descriptive Research.
```

```
*ntxt=x
&type=corps
Some research studies are conducted simply to describe individual variables as they exist naturally. For example, a college official may conduct a survey to describe the eating, sleeping, and study habits of a group of college students. When the results consist of numerical scores, such as the number of hours spent studying each day, they are typically described by the statistical techniques that are presented in Chapters 3 and 4. Non-numerical scores are typically described by computing the proportion or percentage in each category. For example, a recent newspaper article reported that 34.9% of Americans are obese, which is roughly 35 pounds over a healthy weight.
```

```
*ntxt=x
&type=titre
Relationships Between Variables.
```

```
*ntxt=x|
&type=corps
Most research, however, is intended to examine relationships between two or more variables. For example, is there a relationship between the amount of violence in the video games played by children and the amount of aggressive behavior they display? Is there a relationship between the quality of breakfast and academic performance for elementary school children? Is there a relationship between the number of hours of sleep and grade point average for college students? To establish the existence of a relationship, researchers must make observations—that is, measurements of the two variables. The resulting measurements can be classified into two distinct data structures that also help to classify different research methods and different statistical techniques. In the following section we identify and discuss these two data structures.
```

Analyse

Cooccurrence des Thèmes par Phrase

Une analyse possible du logiciel est d’obtenir les cooccurrences de lemmes par phrase. Pour compléter cette analyse, il faut premièrement cliquer sur “Requête” dans la barre de menu en haut de page (Figure 4). Ceci mène à une nouvelle page qui demande quelle sorte de requête effectuer, il faut sélectionner “Analyse” (Figure 12). Ceci mène à une autre page où le logiciel demande de préciser quelle sorte d’analyse effectuer (Figure 13). Parmi les op-

tions se retrouve des tableaux croisés qui sont un tableau ou les catégories (par exemple, différent livre ou chapitre) crée dans le corps du texte sont en colonne et les thèmes trouvés par Sémato sont en rangée. Les cooccurrences sont la fréquence d’apparition de tous les duos d’une liste d’éléments à l’intérieur d’une unité de texte (par exemple, a l’intérieure d’une phrase).

Pour procéder à l’analyse, il faut sélectionner “Cooccurrence avancée” qui mène à une autre page où il faut cliquer sur “Tout le corpus” et ensuite “Cooccurrence de thèmes dans les phrases”. Après ceci, il faut sélectionner “Tous les

Figure 4 ■ Barre de menu en haut de page. Notez le disque dur à l’extrême droite qui sera utilisé à la fin de ce texte.





Figure 5 ■ Ouverture d'un projet sur Sémato.



Figure 6 ■ Nom et mot de passe du projet requis pour ouvrir un projet.



thèmes à la fois” sinon, le logiciel demande de choisir manuellement les thèmes inclus dans l’analyse.

Finalement le logiciel demande de sélectionner la catégorie pour laquelle l’écart détaillé sera calculé. Dans notre cas, nous avons sélectionné la catégorie “Livre”. Une fois la sélection des paramètres de l’analyse est complète, il faut être patient avec le logiciel, car cela peut prendre quelques minutes avant que les résultats soient affichés.

Le premier bloc obtenu dans la page de résultats représente les étapes suivies pour obtenir ce résultat (Figure 14). Le deuxième bloc obtenu démontre les lemmes trouvés automatiquement par le logiciel ainsi que leur fréquence (Figure 15). Si on clique sur le titre d’un lemme, cela nous apporte à une nouvelle page qui constitue un tableau ou on peut voir chaque occurrence du lemme dans le texte ainsi que les autres thèmes automatiquement trouvés dans le contenu textuel d’une occurrence (Figure 16).

Le troisième bloc montre les cooccurrences de thèmes (Figure 17). Par exemple, les thèmes “th-p-hypothesis” et “th-p-null” sont la paire de thèmes la plus fréquente (979 phrases contiennent ces deux thèmes; le premier se retrouve 1745 fois dans les textes, et le second 986 fois). La colonne “Cooccurrences attendues” indique combien de fois cette paire devrait être vue si les thèmes étaient appariés par hasard (la formule est le nombre d’occurrences du premier \times le nombre d’occurrences du second divisé par le nombre de phrases examinées, c.-à-d., 22345 phrases). L’écart entre la cooccurrence observée et attendue est 902. Ce chiffre “902” est d’une certaine façon le “score” de cette paire. La colonne probabilité est le résultat d’une formule qui indique avec quelle probabilité on pourrait s’attendre à un écart de 902 si seul le hasard avait créé ces paires (ici, la probabilité est infiniment proche de zéro).

Le bloc qui suit est le réseau de similitude (Figure 18). Cette section essaye de voir s’il n’y a pas des “grumeaux” de concepts. Le logiciel commence toujours avec la paire

la plus “lourde”, soit “th-p-hypothesis” et “th-p-null” avec un “poids” de 902. Le score ici est divisé par le nombre de thèmes dans cette paire (2), soit 451. Le second grumeau se trouve en ajoutant au premier le couple (“th-p-sample” et “th-p-population”). Chacun a un score de 902 et 732 (1634) qu’on divise par le nombre de thèmes (4). Ensuite, le troisième grumeau le plus “lourd” s’obtient en amalgamant le précédent et “meaning”, etc. La règle pour augmenter les grumeaux n’est pas entièrement claire, mais c’est de trouver des poids de plus en plus grands en amalgamant des concepts ou des réseaux précédemment identifiés.

La configuration au rang 36 a un poids de 18988.943 et est considérée comme un optimum (Figure 19). Passé ce point, tout ce qu’on rajoute au réseau diminue son poids. On dira que la configuration 36 est celle qui maximise les cooccurrences tout en minimisant le nombre de thèmes impliqués. Sémato appelle ce réseau la “configuration focus”.

Suivant le quatrième bloc se situent les liens pour les graphiques (Figure 20). Le lien qui termine avec “... courbes.txt” contient les données du quatrième bloc. Les liens qui terminent avec “... graphml” contiennent du code pour faire des graphiques. Il faut installer le logiciel gratuit yED (troisième lien) pour visualiser les réseaux de connexions. Une fois le logiciel installé, le lien “... graphml-2.graphml” présente tous les thèmes regroupés au centre de l’écran (Figure 21). Les rectangles sont des thèmes, et les liens, la cooccurrence entre ces thèmes (de gros chiffres indiquant un lien très fort). Il est possible de déplacer les thèmes manuellement avec la souris, mais il y a plusieurs centaines de thèmes et cela prendrait beaucoup de temps.

Il est possible de demander au logiciel yED de déplacer les thèmes en changeant la configuration du graphique. Dans le menu en haut à gauche il faut choisir l’onglet “Layout” et ensuite on peut choisir la configuration désirée. Par exemple, si on choisit la configuration “Organic” ceci



Figure 7 ■ Tableau de thème généré par Sémato.

Sémato Texte

Projet Thèmes Requête Documentation *Projet statis*

Page des thèmes - Aide Archives AST GTH TEC Ingrédients Arrimages Graphiques
 Actualiser les fréquences Copier Renommer Effacer Composer Fusionner Répartir Filtrer

- Nombre de thèmes : 418, aucun filtre n'est actif sur les noms des thèmes.
- Pour appliquer l'AST sur un thème, cliquez sur le nom du thème. L'AST permet de déployer sémantiquement un thème.
- Pour exécuter un thème en requête et actualiser ainsi sa fréquence et ses arrimages, cliquez sur le bouton cible associé.
- Le nombre associé à chaque thème indique sa fréquence phrasique au corpus (option sous Projet pour fréquence textuelle).
- Une autre option (sous Projet) permet de changer l'ordre de présentation de la liste des thèmes (alphabétique fréquentiel).

th-p-sample	2867	sample
th-p-variance	2814	discrepancy, variability, variable, variance, variant, variation, vary, varying
th-p-meaning	2738	mean, meaning, meaningful, meaningfully, meaningless, sense, significance
th-p-value	2533	value
th-p-outcome	2451	effect, outcome, result
th-p-data	2166	data, database, datum, inform, information
th-n-test	2137	test, testing

Figure 8 ■ Sous-menu de la page “Thèmes”.

Sémato Texte

Projet **Thèmes** Requête Documentation *Projet statis*

Page des thèmes - Aide Archives AST GTH TEC Ingrédients Arrimages Graphiques
 Actualiser les fréquences Copier Renommer Effacer Composer Fusionner **Répartir** Filtrer

ouvre une autre petite page où on peut changer certains aspects du graphique tel que le “Compactness” (si désiré) et ensuite cliquer sur “Ok” (Figure 22). Après ceci, on obtient un graphique avec des thèmes plus dispersés (Figure 23). Le logiciel essaie une disposition visuelle pour que ce soit “organique” avec les nœuds visibles autant que possible. Ces résultats sont approximatifs (une heuristique) et si on était pour le refaire, le résultat peut être différent.

Il est aussi possible d’obtenir un graphique avec une organisation hiérarchique des thèmes qui permet de voir les thèmes qui sont le plus éloignés du nœud central. Ces thèmes ont souvent moins de relations avec les autres et représentent une pauvre relation au domaine d’étude.

Sauvegarder une analyse

Il est possible de sauvegarder les résultats d’une analyse sur Sémato pour ne pas être obligé de refaire l’analyse à chaque fois que l’on consulte la base de données. Une fois une analyse complète et les résultats obtenus, il faut cliquer sur l’image du disque dur à la droite de la barre de menu (vu à la Figure 4). Ceci amène à une nouvelle page où il est possible de retrouver les analyses sauvegardées précédemment et sauvegarder une nouvelle analyse (Figure 24). La deuxième flèche rouge permet de réviser les derniers résultats obtenus par le logiciel et si satisfaits, la troisième flèche rouge indique de nommer l’analyse dans

Figure 9 ■ Tableau du sous-menu “Répartir”.

Nom actuel	Nouveau nom	Ingrédients	Am	Fr
th-m-chose	<input type="text" value="th-m-chose"/>	choice, choose, chose, multiple-choice, option	-	311
th-m-description	<input type="text" value="th-m-description"/>	describe, description, descriptive	-	329
th-m-geography	<input type="text" value="th-m-geography"/>	geographic, geographical, geography	-	4
th-m-hoc	<input type="text" value="th-m-hoc"/>	hoc, post-hoc	-	66
th-m-mean	<input type="text" value="th-m-mean"/>	average, mean	-	2657

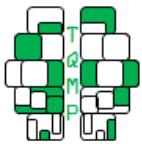


Figure 10 ■ Tableau permettant l'ajout de trois nouveaux thèmes.

Figure 11 ■ Soumettre des modifications aux thèmes.

la case blanche et ensuite soumettre. Cette nouvelle analyse sauvegardée se retrouvera sur la page du disque dur.

Conclusion

Un aspect important du codage du textes lors de l'utilisation de Sémato est la cohérence. Tout aspects textuels doit être coder de la même manière pour obtenir des résultats précis. Pour cette raison il est suggéré de se familiariser avec le contenu sémantique à être analysé pour déterminer des règles d'exclusion et de codage.

La quantité de temps requis pour copier et coder les manuels dans un document en texte brut est une limite du logiciel. Malheureusement, nous n'avons pas pu trouver une façon plus efficace d'effectuer ceci. Il serait idéal d'avoir un

autre logiciel qui pourrait automatiquement transférer les aspects textuels d'un manuel. Après ceci, seules les lignes de codes devront être ajoutées. Ceci éliminerait plusieurs heures de travail de la part du chercheur.

Références

- GRAVETTER, F. J., & WALLNAU, L. B. (2019). *Statistics for the behavioral sciences (10th ed.)* Wadsworth Publishing & Cengage Learning.
- HO, D. (2023, janvier 1). *Notepad++ (Version 8.4.2)*. SourceForge.net. <https://notepad-plus-plus.org/downloads/>
- PLANTE, P., DUMAS, L., & PLANTE, A. (2013). *Sémato : Logiciel d'analyse sémantique des documents textuels*. Université du Québec à Montréal.

Figure 12 ■ Initialiser une nouvelle analyse avec le logiciel Sémato.



Figure 13 ■ Options d'analyse possible par Sémato.

Si vous ne voulez pas sauvegarder vos modifications, quittez cette page sans faire Soumettre.

Cochez ici pour fusionner les thèmes qui ont des ingrédients communs (les thèmes non étanches).
Ne cochez cette case qu'en toute connaissance de cause.

Pour sauvegarder vos modifications, cliquez sur Soumettre.

Figure 14 ■ Résultats, Bloc 1.

Requête	Analyse
Analyse	Cooccurrences avancées
Corpus	Tout le corpus
Cooccurrences	Cooccurrences de thèmes dans les phrases
Catégorie de l'écart détaillé	Livre

Citation

TRUDEL, M. (2023). Comment utiliser le programme d'analyse sémantique Sémato. *The Quantitative Methods for Psychology*, 19(1), 14-24. doi: [10.20982/tqmp.19.1.p014](https://doi.org/10.20982/tqmp.19.1.p014).

Copyright © 2023, TRUDEL. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 09/06/2022 ~ Accepted: 13/11/2022

Figures 16 to 24 follows.

Figure 15 ■ Résultats, Bloc 2 (extrait).

Vocabulaire	Fréquence Totale	Fréquence Texte
ability	65	49
able	111	102
accuracy	58	44
accurate	104	88
accurately	25	25
ace	48	8
adaptation	1	1
add	349	246
addition	30	21
additional	59	51



Figure 16 ■ Fréquence détailler d'un lemme (extrait du lemme "ability").

ability
Fréquence : 65 au total, 49 textes.

- Pour sauvegarder une modification apportée dans une cellule, cliquez illico dans une cellule voisine.

Texte	Texte-Phrase	Contenu textuel	Thèmes automatiques	Thèmes manuels
<i>Les menus pour les ajouts d'arrimage et les retraitts d'arrimage thématique ne peuvent contenir plus de 200 noms de thème. Utilisez les filtres sur les thèmes afin de réduire le nombre de thèmes actifs pour la catégorisation manuelle.</i>				
71-texte	71-3	However, the researcher has no ability to control which participants go into which group—all the males must be in the boy group and all the females must be in the girl group.	th-m-participant, th-p-ability, th-p-girl, th-p-researcher	+ ▾
387-texte	387-5	Nearly half the participants (44%) had extraordinary pitch-naming ability (absolute pitch), and were able to identify most of the tones correctly.	th-m-participant, th-p-ability	+ ▾
435-texte	435-2	You may be somewhat skeptical, especially when you learn that the participants were regular Cornell undergraduate students who were not selected because they had unique ESP abilities.	th-m-participant, th-m-selection, th-m-student, th-m-students, th-p-ability	+ ▾

Figure 17 ■ Résultats, Bloc 3 (extrait)

Table des cooccurrences phrastiques de thèmes

[Les réseaux de similitude de cette table \(plus bas sur cette page\)](#)

Le nombre de cooccurrences présentées a été réduit à 5000.

Thème 1	Thème 2	Fréquence thème 1	Fréquence thème 2	Cooccurrences réalisées	Cooccurrences attendues (hasard)	Écart entre cooccurrences réalisées et cooccurrences attendues	Probabilité	Écart Livre (Livre1)	Écart Livre (Livre2)	Écart Livre (Livre3)	Écart Livre (Livre4)
th-p-hypothesis	th-p-null	1745	986	<u>979</u>	77	902	$=1-$ LOIBINOMIALE(978; 22490; 172057/50580010; VRAI)	270	108	118	399
th-p-sample	th-p-population	2867	1640	<u>941</u>	209	732	$=1-$ LOIBINOMIALE(940; 22490; 235094/25290005; VRAI)	320	120	138	140
							$=1-$				



Figure 18 ■ Résultats, Bloc 4 (extrait).

Réseaux de similitude
Configuration focus trouvée au poids : 18988.943
100 premières inscriptions affichées.

Numéro	Inscription	Seuil	Configuration des réseaux		Poids de la configuration	Nombre de réseaux	Nombre d'inscrits (sur 417)
			Réseau	Poids du réseau			
1	th-p-hypothesis, th-p-null	902	th-p-hypothesis, th-p-null	451	451	1	2
2	th-p-sample, th-p-population	732	th-p-hypothesis, th-p-null th-p-sample, th-p-population	451 366	1634	2	4
3	th-p-sample, th-p-meaning	434	th-p-sample, th-p-population, th-p-meaning th-p-hypothesis, th-p-null	388.66666 451	1679.3334	2	5
4	th-p-test, th-p-hypothesis	419	th-p-hypothesis, th-p-null, th-p-test th-p-sample, th-p-population, th-p-meaning	440.33334 388.66666	1658	2	6

Figure 19 ■ Résultats, Bloc 4, Configuration Focus.

Numéro	Inscription	Seuil	Réseau		Poids de la configuration	Nombre de réseaux	Nombre d'inscrits (sur 417)
			Réseau	Poids du réseau			
36	th-p-equation, th-p-regression	130	th-p-deviation, th-p-difference, th-p-distribution, th-p-meaning, th-p-normality, th-p-outcome, th-p-population, th-p-sample, th-p-treatment, th-p-significant, th-p-frequency	372.18182	18988.943	11	40
			th-p-hypothesis, th-p-null, th-p-test, th-p-rejection, th-p-region, th-p-statement, th-p-optional	351.2857			
			th-p-variance, th-p-independence, th-p-dependent, th-p-relationship, th-p-measurement	240.8			
			th-p-degree, th-p-freedom	165			
			th-p-regression, th-p-line, th-p-equation	96			
			th-p-value, th-p-table	106.5			
			th-p-student, th-p-researcher	97			
			th-p-summation, th-p-square	92			
			th-p-interval, th-p-confidence	73			
			th-p-column, th-p-row	67.5			
			th-p-level, th-p-alpha	65			

Figure 20 ■ Liens au graphique en bas de page.

Fichier construction de la courbe des poids de configurations : [statis/sorties/G515878964245-Global-courbes.txt](#)
 Fichier tableur matrice des similitudes : [statis/sorties/G515878964245-Global-matrice.txt](#)

Fichiers GraphMI (logiciel yEd pour lire les graphml) :
 Fichier graphml (configuration focus) : [statis/sorties/G515878964245-Global-graphml-1.graphml](#)
 Fichier graphml (réseau total) : [statis/sorties/G515878964245-Global-graphml-2.graphml](#)



Figure 21 ■ Graphique créé avec yED où tous les thèmes sont regroupés.

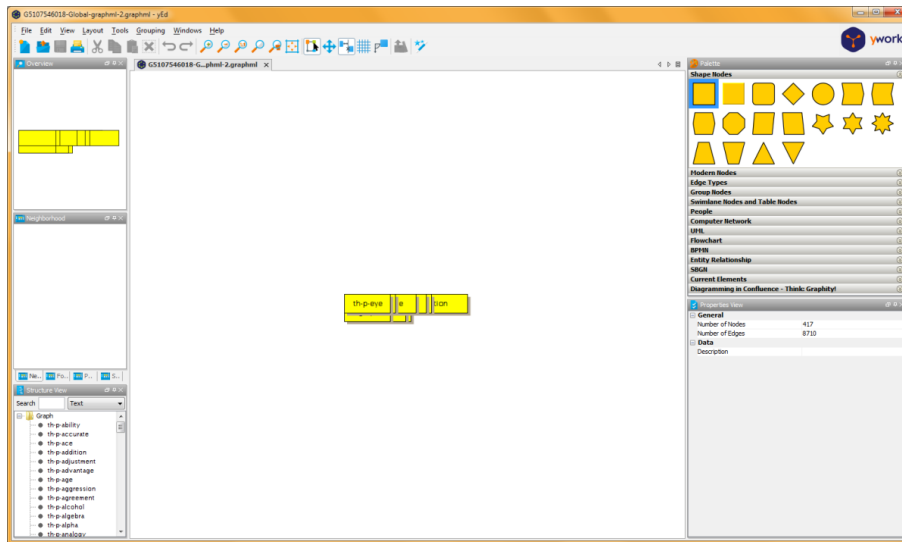
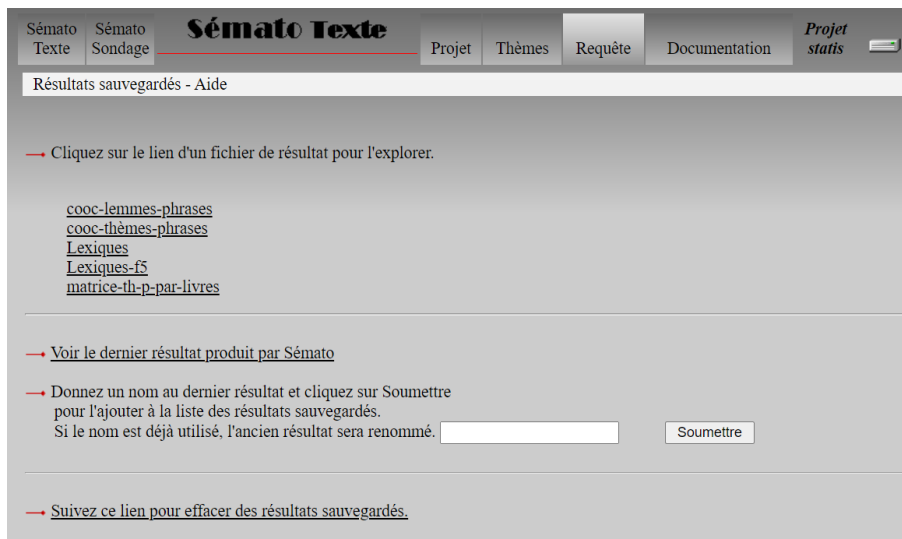


Figure 22 ■ Modifier le graphique avec le logiciel yED.



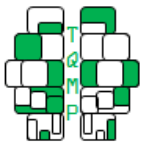


Figure 23 ■ Graphique en organisation organique obtenu avec le logiciel yED.

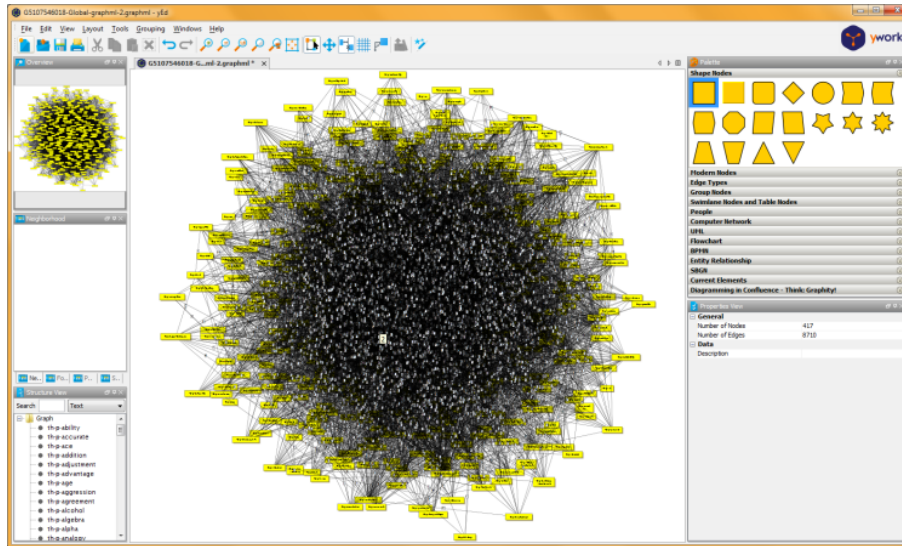


Figure 24 ■ Sauvegarder une analyse sur le disque dur.

Sémato Sémato **Sémato Texte** Projet Thèmes Requête Documentation *Projet stats*

Résultats sauvegardés - Aide

→ Cliquez sur le lien d'un fichier de résultat pour l'explorer.

[cooc-lemmes-phrases](#)
[cooc-thèmes-phrases](#)
[Lexiques](#)
[Lexiques-f5](#)
[matrice-th-p-par-livres](#)

→ Voir le dernier résultat produit par Sémato

→ Donnez un nom au dernier résultat et cliquez sur Soumettre pour l'ajouter à la liste des résultats sauvegardés. Si le nom est déjà utilisé, l'ancien résultat sera renommé. Soumettre

→ Suivez ce lien pour effacer des résultats sauvegardés.