

# Suggestions de procédure d'utilisation du logiciel SATO<sup>1</sup>

Document d'information rédigé au bénéfice des étudiants et des chercheurs qui souhaitent s'initier à l'analyse informatisée de données textuelles.

par

Claire Gélinas-Chebat,  
professeure  
Département de linguistique  
Université du Québec à Montréal

Mai 2010.

---

<sup>1</sup> Nous remercions vivement François Daoust pour sa disponibilité et ses commentaires.

## Table des matières

Introduction.....	3
1er partie . Préparation des données textuelles.....	3
Étape 1 : Préparation de votre corpus.....	3
Un corpus simple.....	3
Un corpus plus complexe.....	3
Étape 2 : Saisie de votre corpus.....	4
2.1. Accéder au site web du logiciel :.....	4
2.2. Ouvrir un compte sur SATO.....	4
2.3. Accéder au logiciel.....	4
2.4 Envoyer votre fichier .....	4
Étape 3 : Construction de la matrice.....	5
3.1. Soumettre à SATO.....	5
3.2. Remplir le formulaire.....	6
3.5. Quitter SATO.....	7
Étape 4 : Accès et traitements des données.....	8
4.1. Accès aux données.....	8
4.2. Traitements textuels des données avec SATO.....	8
2e partie. Analyses textuelles à l'aide du logiciel SATO.....	11
Étape 5 : Procédures d'analyses d'un premier niveau .....	11
5.1. Quelques opérations sur le lexique.....	11
5.2. Quelques opérations sur le texte.....	12
Étape 6 : Procédures plus poussées d'analyses avec SATO.....	12
6.1. Quelques opérations sur le lexique.....	13
Étape 7 : La saisie des données et le transfert vers d'autres logiciels.....	16
7.1. Le transfert du lexique dans un tableur (Excel ou autre).....	16
7.2. Le passage de SATO vers Alceste, DTM, ou LEXICO3. ....	16
Informations complémentaires.....	17
8.1. Plan lexique/occurrences.....	18
8.2. Catégorisation.....	18
8.3. Analyse.....	19
8.4. Interface.....	19
8.5. Outils.....	19

## Introduction

Il faut distinguer trois étapes dans l'utilisation du logiciel SATO :

1. **La saisie** de votre corpus par le logiciel : il faut envoyer votre fichier sur le serveur.
2. **La construction de la matrice** lexicque-occurrences : il faut soumettre votre fichier pour le mettre en forme pour SATO.
3. **Les analyses** à l'aide du logiciel. Votre fichier n'a pas à être soumis à nouveau, il est sur le serveur (ce qui a été fait à l'étape 2) et SATO peut (ce qui a été fait à l'étape 3) procéder aux différentes analyses que vous allez demander.

### 1<sup>er</sup> partie . Préparation des données textuelles

#### Étape 1 : Préparation de votre corpus

##### Un corpus simple

Un corpus constitué d'un seul texte : c'est par la simple fonction *copier-coller* que vous pouvez constituer votre premier corpus.

##### Un corpus plus complexe

Un corpus constitué de différents textes, différents locuteurs, différentes sources, etc. : il faut marquer ces variables par des indices dans votre texte pour permettre à SATO de les prendre en compte au moment des analyses. Il s'agit de *propriétés* SATO introduites par un caractère particulier. Pour distinguer plusieurs textes, par exemple, on introduira au début du premier texte : *\*page=texte1* puis *\*page=texte2* pour le second et ainsi de suite. Pour étudier la variable *sexe*, on peut introduire *\*sexe=H* pour les hommes *\*sexe=F* pour les femmes. SATO donne une très grande liberté dans le libellé des propriétés et de ses valeurs.

Pour exclure des portion de texte comme le titre, l'auteur, des détails descriptifs du corpus qu'on veut conserver mais qu'on ne veut pas traiter comme éléments d'analyse, on introduit dans le texte d'origine « \*{ » pour déterminer le début du texte à exclure, qu'on encadre par « }».

Par exemple,

*\*{Texte descriptif de la procédure de la préparation d'un corpus}*

serait exclu de l'analyse si on soumet ce texte à SATO dans sa forme actuelle.

Selon les besoins, il est possible de considérer un corpus comme un seul fichier constitué de plusieurs textes. On peut aussi, créer autant de fichiers que de textes à distinguer.

Une fois cette information codifiée, utilisez la fonction *copier-coller* pour chaque fichier à créer.

## Étape 2 : Saisie de votre corpus

### Démarches à suivre lors d'une première visite :

#### 2.1. Accéder au site web du logiciel :

<http://www.ling.uqam.ca/sato/index.html>

#### 2.2. Ouvrir un compte sur SATO

Dans le menu de gauche, choisir *Première inscription* puis, compléter le formulaire.

Utilisez un pseudonyme et un mot de passe facile à mémoriser puisque vous les utiliserez lors de vos prochaines visites sur le site.

**SATO Version 4.3**

- Ouvrir une session
- Première inscription
- Gestion de compte
- Langue: English | Español

**RESSOURCES**

- Forum
- Projet Visibilité
- Corpus express

**A PROPOS DE SATO**

- Fiche descriptive
- Manuel [format HTML]
- Installation et gestion

**ADMINISTRATION**

- Gestion du site
- SATO sur *strophe*

SATO (CHN RE-ATO) / OUVRIER D'UNE SESSION

Votre pseudonyme :  
(lettres et chiffres seulement)

Votre mot de passe :

Nom du compte partagé (optionnel)\* :  
(lettres et chiffres seulement)

Mot de passe du compte partagé\*\* :

Confirmez la langue de l'interface :

Choisissez une application :

Français  English  Español

Démonstrations

SATO-Calibrage

SATO en mode intégral

APPLIQUER RECOMMENCER

**Remarque**  
Les versions de l'interface en anglais et en espagnol ne sont pas encore disponibles.

Une session SATO doit toujours débiter par cette page.

Pour une première inscription, cliquez sur l'onglet « Première inscription »; pour modifier vos données personnelles ou donner accès à votre compte en mode lecture, choisissez l'onglet « Gestion de compte ».

Pour terminer une session SATO, on doit cliquer sur le lien **Quitter**.

#### 2.3. Accéder au logiciel

Bienvenue sur le serveur SATO-C du Centre ATO de l'UQAM

Nombre de sessions en cours : 5

**SATO en mode Intégral**

Cette application vous permet d'accéder à SATO en mode intégral et de gérer vos fichiers sur le serveur.

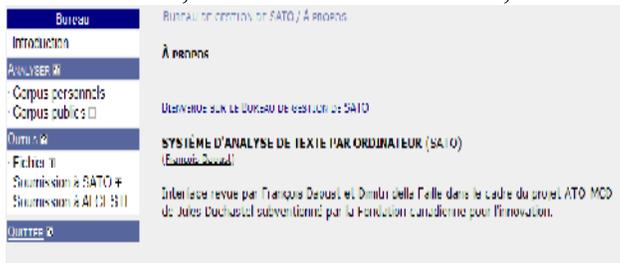
**Entrer**

[Quitter](#)

#### 2.4 Envoyer votre fichier

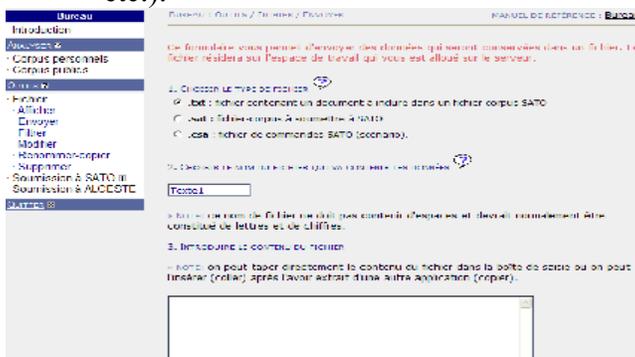
La procédure est la même que vous vouliez constituer un seul fichier pour plusieurs textes, ou un fichier par texte.

Dans le menu, choisir la fonction *Outil*, ouvrir *Fichier* et sélectionner *Envoyer*



Puis procéder aux étapes suivantes :

- sélectionner le type de fichier (maintenir la proposition *.txt*),
- donner un nom à votre fichier (un nom court sans espace). Chaque fichier doit avoir un nom particulier.
- coller le texte à analyser dans l'espace approprié (soit en le tapant directement à l'écran, soit par la fonction *copier-coller* du traitement de texte, du navigateur Web, etc.).



- Procéder à l'envoi de votre fichier, en texte intégral. On peut répéter la procédure pour chacun des fichiers à envoyer.

À cette étape vous avez un ou plusieurs fichiers «texte» (ce sont des documents *.txt*) qui ont été créés dans l'espace SATO.

Dans le menu *Outils*, la rubrique *Fichier*, vous permet de revenir sur votre texte soit pour l'*afficher*, soit pour le *modifier* et le corriger, ou encore pour le *renommer*.

**Les fichiers n'ont pas encore été soumis au logiciel.**

## Étape 3 : Construction de la matrice

### 3.1. Soumettre à SATO

Dans le menu *Outils*,

a. Choisir *Soumission à SATO* puis

b. **Créer et soumettre un corpus**

Même si vos textes sont maintenant dans des fichiers en tant que document *.txt* sur le serveur, il faut maintenant les rassembler sous forme de fichier corpus portant le suffixe *.sat*. Lorsque vous reviendrez ultérieurement, vous pourrez éventuellement choisir *Soumettre un corpus déjà existant*.

### 3.2. Remplir le formulaire

a. **Le nom du fichier.**

Vous devez choisir un nom pour le fichier (lettres minuscules et chiffres seulement). Ce nom, par opposition aux noms donnés aux textes individuels à l'étape précédente, correspond à un fichier vous permettant de traiter plusieurs textes qui auraient été saisis à l'étape précédente. Ces textes *.txt* seront considérés comme des documents (sous-corpus) de votre corpus global.

b. **La langue**

Par défaut, français

c. **Les caractères spéciaux**

Pour le moment utilisez les suggestions de SATO par défaut. Il faut savoir qu'il existe des caractères (des termes) réservés à SATO et que si votre texte en contient, ils seront interprétés par SATO comme une commande. Par exemple, le caractère suivant «\*», est interprété comme l'ouverture d'une commande. Si votre texte en contient, SATO cherchera à interpréter cette commande.

d. **La dimension de la page**

Par défaut : 60 lignes, 68 caractères

e. **Autres déclarations**

Il est possible de définir la prise en compte de certaines variables de votre corpus (âge, sexe, locuteur, etc.) : ce sont des *propriétés* SATO. Le choix du libellé pour la propriété et ses valeurs est laissé à l'initiative du chercheur. Pour les propriétés symboliques (catégorielles), la limite est d'environ 250 symboles, simples ou combinés. Si vous avez introduit de telles *propriétés* SATO dans votre texte initial, par exemple \*sexe=H, \*sexe=F, SATO pourra en tenir compte au moment des analyses.

**Dans la boîte de saisie, il faut inscrire la propriété et toutes les valeurs attribuées en respectant scrupuleusement cette syntaxe :**

*propriété* LIBELLÉ *symbolique pour texte* VALEUR1 VALEUR2

Exemples de syntaxe :

*propriété* sexe *symbolique pour texte* F H

Il est possible d'omettre les valeurs de la propriété dans la déclaration. Dans ce cas, SATO va les ajouter au fur et à mesure qu'il les rencontre dans le texte. Mais, si on a fait une erreur de frappe dans la valeur de la propriété, SATO ne pourra pas s'en apercevoir.

Si vous n'avez pas spécifié de *propriétés* SATO dans le texte original, il n'y a rien à inscrire dans la boîte de saisie.

**f. Choisir un titre pour votre corpus**

Donner un titre à votre corpus vous permettant de mieux vous rappeler en quoi consiste votre corpus. SATO conserve cette information qu'il fera apparaître lors du chargement du corpus en phase d'analyse.

**g. Entrer le contenu de votre corpus**

Sélectionner le ou les documents pertinents qui apparaissent à droite de l'encadré sous la forme :

**\*page=@NOM-DU\_FICHER.txt**

Il aurait aussi été possible d'entrer votre texte par la fonction *copier-coller*.

**Puis envoyer votre corpus** (par la fonction *envoyer* dans le bas de la page écran).

Un fichier est créé sur votre espace de travail. Votre corpus a été **SATOGÉNÉRÉ!**

Vous verrez apparaître une nouvelle fenêtre et c'est ici que vous saurez si votre corpus a été soumis avec succès à SATO.

Si SATO n'a pas décelé d'erreurs, vous pourrez commencer immédiatement vos analyses.

Par contre, si SATO détecte des erreurs, il les listera à cette étape. Dans ce cas, faites vous une copie des commentaires de SATO dans un document Word (avec la fonction *copier-coller*). Cette opération vous permettra de corriger méthodiquement toutes les erreurs listées. SATO détaille les erreurs et il faudra modifier chacune des fichiers *.txt* qui contient ces erreurs. Vous devez *quitter, sans sauvegarde* au bureau SATO (interface en bleu) pour modifier les fichiers *.txt* en erreur : onglet *Outils/Fichier/Modifier*. **Reprenez cette étape jusqu'à ce que SATO génère votre corpus sans indiquer d'erreurs.**

### 3.5. Quitter SATO.

Pour terminer une session SATO, on doit cliquer sur *Quitter*.

**Quitter avec sauvegarde**, si vous voulez sauvegarder les analyses et les annotations que vous auriez faites.

**Quitter sans sauvegarde**, si vous voulez sortir du logiciel. Votre corpus est créé sur l'espace de travail et ne sera pas affecté par cette étape. Il est souvent préférable d'utiliser cette fonction pour ne pas écraser les annotations faites dans des sessions précédentes.

## Étape 4 : Accès et traitements des données

### 4.1. Accès aux données

**Reprendre les étapes suivantes :**

- Accéder au site web du logiciel :  
<http://www.ling.uqam.ca/sato/index.html>
- Ouvrir son compte

SATO (CENTRE ATO) / OUVERTURE D'UNE SESSION

Votre pseudonyme :  
*(lettres et chiffres seulement)*

Votre mot de passe :

Nom du compte partagé (optionnel)\* :  
*(lettres et chiffres seulement)*

Mot de passe du compte partagé\* :

Confirmez la langue de l'interface :

Choisissez une application :

Remarques  
Les versions de l'interface en anglais et en espagnol ne sont pas encore disponibles.

Une session SATO doit toujours débiter par cette page.

Pour une première inscription, cliquez sur l'onglet « Première inscription »; pour modifier vos données personnelles ou donner accès à votre compte en mode lecture, choisissez l'onglet « Gestion de compte ».

Pour terminer une session SATO, on doit cliquer sur le lien **Quitter**.

- Accéder au logiciel

Bienvenue sur le serveur SATO-C du Centre ATO de l'UQAM

Nombre de sessions en cours : 5

**SATO en mode intégral**

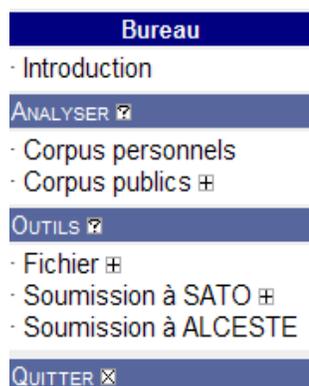
Cette application vous permet d'accéder à SATO en mode intégral et de gérer vos fichiers sur le serveur.

**Entrer**

[Quitter](#)

### 4.2. Traitements textuels des données avec SATO

Choisir l'opération souhaitée à partir du menu



Voici une description succincte des éléments du menu du bureau:

- Sous la rubrique **Introduction** :  
Texte de présentation par François Daoust
- Sous la rubrique **Analyser**  
**Corpus personnels** :  
Dans *Corpus personnels*, vous retrouverez tous les documents que vous avez créés dans votre espace de travail, dont le document que vous avez créé aux étapes précédentes. En sélectionnant *Corpus personnels* vous sélectionner votre corpus et cocher *Analyser le corpus* au bas de l'écran.  
**Corpus publics** :  
Donne accès à des corpus mis en ligne par des chercheurs utilisateurs de SATO.



Chacune des catégories contient un ou plusieurs corpus qui peuvent être traités en SATO intégral.

- Sous la rubrique **Outils**  
*Fichier* : permet différentes opérations sur vos fichiers SATO



### *Soumission à SATO*

Permet d'accéder au logiciel SATO pour *Créer et soumettre un corpus* ou *Soumettre un corpus déjà existant*. Soumettre un corpus déjà existant fera qu'à la fin de la session, le travail fait durant cette session fera disparaître le travail qui aurait été fait dans une (ou des) session(s) antérieures. Cette version écrasera la version précédente. Si, par exemple, vous aviez défini des propriétés dans des sessions antérieures de travail, elles seront perdues et remplacées par les propriétés que vous venez de créer. Le *Journal* permet de retrouver une trace de tout travail qui a été fait depuis le tout début.

La *Foire aux questions* donne accès à de nombreuses réponses à des questions courantes au moment de la soumission d'un corpus à SATO.



Alceste est un logiciel de traitement statistique des données textuelles. L'interface Web permet un accès limité au logiciel qu'il est possible d'acheter auprès du distributeur pour une utilisation autonome.

## 2<sup>e</sup> partie. Analyses textuelles à l'aide du logiciel SATO

Il y a deux entrées possibles :

Soit par *Corpus personnels*

Soit par *Outils, Soumission à SATO*

**Il est préférable de procéder par l'entrée *Corpus personnel* pour conserver les propriétés créés dans des sessions de travail antérieures**

Dans les deux cas on accède au menu suivant, menu qui permettra d'obtenir différents résultats d'analyse à l'aide de SATO :



### Étape 5 : Procédures d'analyses d'un premier niveau

#### 5.1. Quelques opérations sur le lexique

##### 1. PLAN LEXIQUE / OCCURENCES

##### 1.1 Lexique :



##### a. Pour afficher le lexique

Choisir LEXIQUE, AFFICHER

Il est possible de choisir en 1 un filtre de fouille du lexique. Par défaut SATO propose tous les éléments du corpus (c'est la raison du signe \$ ). Par exemple, c\$ vous permet de sortir tous les éléments lexicaux qui commencent par c.

Choisir en 2, sous quelle forme vous voulez faire apparaître le lexique. Par ordre alphabétique, par fréquence (fréqtot), etc.



Projet ATO-MCD : interface de SATO

b. Pour décrire le lexique  
Choisir LEXIQUE, DÉCRIRE  
Informe sur l'alphabet utilisé et le nombre de formes lexicales du corpus.

## 5.2. Quelques opérations sur le texte

a. Pour afficher le texte

Il est possible de choisir un filtre pour sélectionner des parties du texte  
TEXTE AFFICHER

b. Pour décrire le texte

Permet d'avoir l'information sur la constitution du corpus. Un corpus peut être constitué de plusieurs documents et on peut faire apparaître le nombre de documents, le nombre de pages et la répartition du nombre de pages par document.

TEXTE DECRIRE

## Étape 6 : Procédures plus poussées d'analyses avec SATO

## 6.1. Quelques opérations sur le lexique

### a. Pour afficher le lexique

Choisir LEXIQUE, AFFICHER

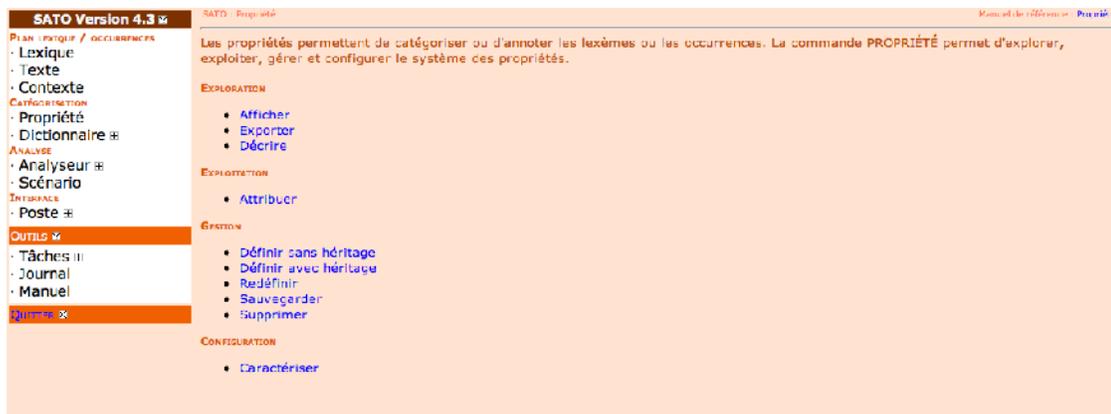
Au point 1, choisir le patron de fouille. Par défaut SATO propose tous les éléments du corpus ( \$ ). On peut appliquer des filtres pour faire ressortir les éléments lexicaux en fonction de différentes caractéristiques. Par exemple c\$ vous permet de sortir tous les éléments lexicaux qui commencent par c. Puis, en 2, choisir les modalités de présentation (alphabétique, fréquentielle, etc.)



Projet ATO-MCD : interface de SATO

### b. Pour ventiler le lexique en *sous-textes*.

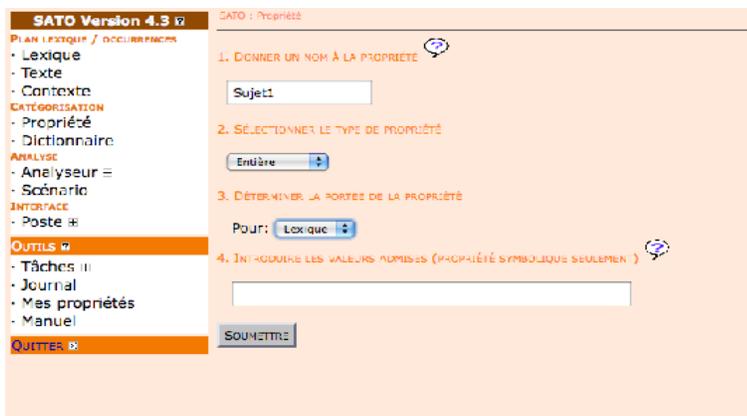
En première étape, il faut ajouter des colonnes (propriétés lexicales) afin de conserver le décompte des formes lexicales dans des parties données du corpus (sous-textes).



Il faut choisir

- PROPRIÉTÉ DÉFINIR sans héritage  
puis, inscrire un nom à la propriété, (par exemple, Sujet1);

ce nom apparaîtra dans le titre de la colonne du tableau du lexique ventilé. Il apparaîtra également comme nom de la propriété.  
 Pour compter des entrées du lexique selon divers critères, on choisira *entière* comme type de la propriété. Puis, cliquez sur *Soumettre*.



On répète cette opération autant de fois qu'on veut ajouter des colonnes au lexique. L'ordre des colonnes ne pourra pas être modifié ultérieurement.

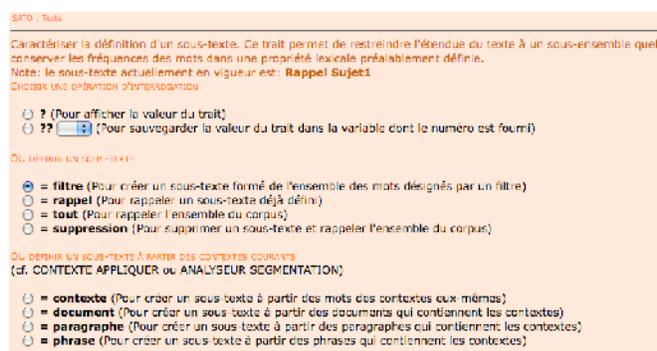
Remarque. Pour annoter le lexique, par exemple par un système de catégories, on doit choisir le type *symbolique* pour un système de catégories ou *libre* pour une annotation en format libre.

Puis, il faut procéder aux étapes suivantes :

### •TEXTE CARACTÉRISER SOUS-TEXTE



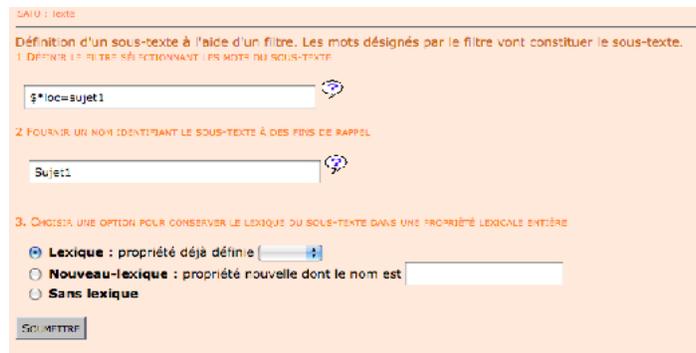
Choisir à la fenêtre suivante = **FILTRE**



Puis, en 1, définir le «filtre» de sélection des éléments lexicaux, nécessairement en lien avec les propriétés préalablement définies.

Par exemple,  $\$*loc=sujet1$  pour sélectionner toutes les interventions du locuteur *sujet1*. Ou,  $\$*sexe=H$  pour sélectionner toutes les parties du texte marquées par cette propriété.

Dans le deuxième champ, donner un nom au sous-texte.



Choisir en 3 le nom de la propriété déjà définie dans le menu déroulant.



Si on a oublié de définir une propriété lexicale (colonne) pour le sous-texte, on peut cocher *Nouveau-lexique* et entrer le nom d'une nouvelle propriété.

On répète cette opération qu'on a créé de colonnes dans le lexique pour calculer les fréquences des mots dans des sous-textes.

Pour modifier la largeur des colonnes au moment de la présentation du lexique, il faut sélectionner

#### •PROPRIETE CARACTERISER



Puis, il faut sélectionner la propriété à caractériser, ici *Sujet1*.

Il sera possible de modifier le trait **LARGEUR** en redéfinissant cette largeur (par défaut 9)

**LARGEUR = 6**

N.B » : Le trait FORMAT permet d'afficher les résultats en pourcentage, choisir dans le menu déroulant  
Définir la valeur du trait : POURCENTAGE.

## Étape 7 : La saisie des données et le transfert vers d'autres logiciels

### 7.1. Le transfert du lexique dans un tableur (Excel ou autre)

Le transfert du **lexique ventilé** dans un tableur peut se faire en procédant aux commandes suivantes successivement :

En premier, sous la rubrique «INTERFACE», choisir POSTE PROTOCOLE puis CARACTÉRISER GÉNÉRAL et changer le protocole à l'étape «Choisir une opération», sélectionner «=>» et dans le menu déroulant choisir TABULAIRE.

Puis, sous la rubrique «PLAN LEXIQUE», choisir LEXIQUE EXPORTER puis définir les lexèmes, (par défaut tous les éléments «\$») et le tri qu'on veut opérer (par ordre alphabétique, fréquentiel, etc.)

Poursuivre en choisissant sous la rubrique «INTERFACE» POSTE EXPORTATION AFFICHER

Il faut noter que le patron par défaut proposé sera un fichier .lis, appuyer sur CONTINUER, choisir le bon fichier puis soumettre en texte brut.

Enfin, par la fonction *COPIER-COLLER* il est possible d'importer d'un coup toutes les valeurs affichées dans une feuille de calcul. Pour un fichier EXCEL, il faut choisir la fonction *COLLAGE SPÉCIAL* dans le menu d'édition, puis *TEXTE UNICODE* pour préserver les colonnes préalablement définies dans SATO.

### 7.2. Le passage de SATO vers Alceste, DTM, ou LEXICO3.

Il est possible de convertir un corpus soumis à SATO dans le format de divers autres logiciels de *textométrie*, en particulier ALCESTE, DTM ou encore LEXICO3. Pour convertir les données d'un format à l'autre, on utilise un langage pivot en XML utilisant des balises du Text Encoding Initiative (*TEI*)

Pour exporter un corpus SATO vers XML-TEI, on ouvre une session d'analyse sur le corpus (interface saumon). On choisit dans le menu *OUTILS* l'option *Exportation*, puis *Exportation en format TEI*. Il suffit alors de choisir le nom du fichier (sans le suffixe *xml* qui sera ajouté automatiquement) et de cliquer sur *Exporter*.

Remarque. Il est à noter qu'on peut aussi se servir de cette section de SATO pour *exporter en format SATO* un corpus annoté suite à des analyses avec le logiciel. Ainsi, par exemple, SATO fournit dans *OUTILS / Locutions et mots complexes* des façons de marquer des expressions. Si on exporte ce corpus enrichi en format SATO pour soumettre cette nouvelle

version à SATO, on aura un lexique qualifié dans lequel on retrouvera nos expressions (termes complexes) plutôt que les mots détachés.

Une fois que l'on a un fichier en format XML-TEI dans son espace de travail, il est possible de le convertir dans le format spécifique d'un certain nombre de logiciels. Pour ce faire, on utilise le menu Outils/fichier/Filtrer du bureau.



Dans le sous-menu de *Filtrer*, on choisit *Conversion de format : de XML-TEI vers ALCESTE, DTM, LEXICO-3, SATO*, Ensuite, on choisit le nom du fichier XML et le logiciel désiré. Le fichier sera créé dans l'espace de travail. On pourra l'afficher et en faire un copier-coller sur son ordinateur pour le soumettre au logiciel choisi.

## Informations complémentaires

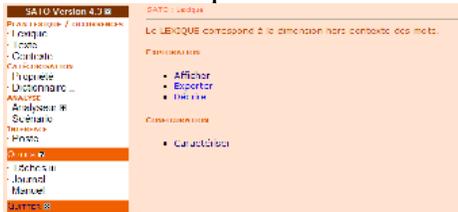
Cette section est consacrée à la description systématique des ressources de SATO telles qu'elles apparaissent :



Ce menu dans SATO donne accès dans la partie «*SATO Version 4.3*», à 4 sections différentes et dans la partie «*Outils*», à 4 sections.

## 8.1. Plan lexique/occurrences

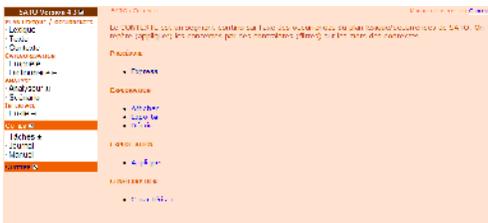
### a. Lexique :



### b. Texte :



### c. Contexte



## 8.2. Catégorisation

### a. Propriété



### b. Dictionnaire



```
* ----- 2010-01-14 15:24:00 -----  
* Bienvenue sur le serveur SATO D du Centre ATO de l'UQAM  
* Dupuis - AN192.168.111.5 (sato-ds@users.sourceforge.net)  
TEXTE DECHIFFRE  
TEXTIQUE DECHIFFRE
```

[Rafraîchir](#)

Rappel d'une commande: entrer la commande. Décocher pour compléter la commande en mode menu.

Soumettre

## c. Mes propriétés

propriété *NoLex* entière pour lexique

propriété *NoOcc* entière pour texte

propriété *Alphabet* symbolique pour lexique fr

propriété *Édition* symbolique pour texte collé (é maj par cap tab 2 3 4 5 6 7 8 no fin)

propriété *Fréqtot* entière pour lexique

propriété *Longueur* entière pour lexique

propriété *Page* *référentielle* pour texte

propriété *Commentaire* libre pour texte

propriété *pub* symbolique pour texte brochure

propriété *Indicateur* symbolique pour texte i01 i02 i03 i04 i05 i06 i07 i08 i09 i10 s01hn s02fn s03ff s04hf s05ff s06hf s07fn s08hn s09hn s10ff s11hf s12hf s13ff s14hn s15fn s16fn s17fn s18ff s19hn s20fn s21hf s22hn s23hn s24hn s25hf s26hn s27fn s28hf s29hf s30ff s31hn s32fn s33hn s34hf s35ff s36hn s37fn s38hn s39hn s40ff s41hf s42hf s43ff s44ff s45xxx s46xxx s47xxx s48xxx

propriété *Thème* symbolique pour lexique apparence arrêt négation concret danger dépendance acc-je maladie mort plaisir publicité tabac nicotine drogue interdiction fumeur soc-ami soc-famille soc-gens liberté envie conscience volonté soc-jeune coûts début santé éducation prévention

propriété *sexe* symbolique pour texte femme homme

propriété *fumeur* symbolique pour texte oui non

## d. Manuel

SATO 4.3. Manuel de référence (mars 2007; dernière modification en décembre 2009)

Système d'analyse de texte par ordinateur

### Table des matières

0. [Modifications récentes au logiciel](#)
1. [Introduction](#)
2. [Présentation du logiciel SATO](#)
3. [Bureau](#)
4. [Préparation d'un corpus pour soumission à SATO](#)
5. [Commandes \(présentation\)](#)
  - PLAN LEXIQUE / OCCURRENCES :
    - o [LEXIQUE](#)
    - o [TEXTE](#)
    - o [CONTEXTE](#)
  - CATEGORISATION :
    - o [PROPRIÉTÉ](#)
    - o [DICTIONNAIRE INDEXÉ](#) | [SÉQUENTIEL](#) | [TABULAIRE](#)
  - ANALYSE :
    - o [ANALYSEUR COMPARAISON](#)
    - o [ANALYSEUR COMPTAGE](#)
    - o [ANALYSEUR DISTANCE](#)
    - o [ANALYSEUR LISIBILITÉ](#)
    - o [ANALYSEUR PARTICIPATION](#)
    - o [ANALYSEUR SEGMENTATION](#)
    - o [ANALYSEUR TAMISAGE](#)
    - o [SCÉNARIO](#)
  - INTERFACE :
    - o [POSTE ÉCRAN](#)
    - o [POSTE EXPORTATION](#)
    - o [POSTE EXTRACTION](#)
    - o [POSTE JOURNAL](#)
    - o [POSTE PROTOCOLE](#)
    - o [POSTE TOUCHE](#)
    - o [POSTE VARIABLE](#)
6. [Outils](#)
7. [Définitions](#)
8. [Liste des erreurs](#)
9. Tutoriels :
  - o [Guide de préparation de corpus pour soumission à SATO \(pdf\)](#)
  - o [Catégorisation lexicale \(pas-à-pas\)](#)
  - o [Catégorisation en contexte \(pas-à-pas\)](#)
  - o [Touches de catégorisation \(pas-à-pas\)](#)
  - o [Sous-texte avec un filtre \(pas-à-pas\)](#)
  - o [Sous-texte à partir de contextes \(pas-à-pas\)](#)
  - o [Scénario \(pas-à-pas\)](#)
  - o [Analyse de la distance \(pas-à-pas\)](#)
10. [Soumission d'un corpus](#) [Foire aux questions](#)
11. [Guide de programmation des interfaces HTML](#)

Suit une liste des modifications récentes au logiciel.