Le Trameur

Propositions de description et d'implémentation des objets textométriques

Serge Fleury

Février 2013

L'objectif principal de la textométrie est de compter des éléments (des *contenus* textuels) dans des ensembles (des *contenants* regroupant des unités élémentaires d'un texte ou des zones de texte couvrant un certain nombre ou un certain type d'unités élémentaires). Le processus de comptage nécessite au préalable d'identifier les *contenus* et les *contenants*. Ce préalable consiste à expliciter une *segmentation* du texte conduisant à la mise au jour d'une ***Trame*** (les contenus) et d'un ***Cadre*** (les contenants).

Ce document met au jour une description des objets textométriques et les méthodes mises en œuvre dans le ***Trameur*** pour travailler sur et avec ces objets dans une perspective textométrique.

Mots-clés : trame, cadre, sélection, annotation

Sommaire

1. La textométrie : préambule 3
2. Constituer des ressources textométriques (coder les textes) 5
2.1 Exemple avec Lexico3 5
2.2 Exemple avec Le Trameur 9
2.2.1 Format texte brut 9
2.2.2 Format dit « Lexico » 9
2.2.3 Format XML 11
2.2.4 Matérialisation graphique du cadre 13
3. Les accès aux textes en machine 18
4. Les objets textométriques 20
4.1 Trame et Cadre textométriques 20
4.2 Sélection 21
4.3 Annotations 22
5. Propositions d'implémentation dans le Trameur 24
5.1 Préparer les textes aux explorations textométriques 24
5.2 La Trame et le Cadre dans le Trameur 26
5.3 Les annotations dans le Trameur 27
5.4 Implémentation des propositions : le moteur textométrique 27
5.3 L'objet Sélection dans le Trameur 29
6. Modifier des ressources textométriques (corriger les textes dans la machine) 30
Mise à jour locale de la Trame 30
Mise à jour globale de la Trame 31
Ajouter des annotations sur la Trame 32
Mise à jour d'une sélection de positions sur la Trame 32
7. Echanges de Bases Textométriques 35
8. (Re)Modelage d'une Base Textométrique 37
9. Références bibliographiques 39
10. Annexes 40
10.1 Base lexicométrique exportée via Le Trameur (dormeur – texte brut) 40
10.2 Base lexicométrique exportée via Le Trameur (dormeur – TEI) 45

1. La textométrie : préambule

Source : [Söze-Duval, 2008]

> ***La démarche textométrique repose sur l'hypothèse, vérifiée à partir de très nombreuses expériences, que pour comparer les différentes parties d'un ensemble de textes, que l'on peut considérer comme autant de contenants, il est utile d'observer, au sein de ces textes, les variations de fréquence de systèmes d'unités textuelles : lexèmes, graphèmes, etc., que l'on peut considérer comme des contenus. L'analyse de ces variations constatées au plan statistique fournit, en général, un éclairage précieux sur les ressemblances et les oppositions pouvant exister entre ces textes sur des plans d'analyse qui intéressent plus directement les chercheurs qui entreprennent ces expériences (linguistes, analystes de discours, sociologues, politologues, etc.).***
>
> ***Une description formelle de ces deux systèmes d'unités (contenants et contenus) permet d'obtenir, à l'aide de procédures informatisées, des décomptes qui se présentent sous formes de vastes tableaux statistiques (à une ou à plusieurs dimensions). Dans la pratique, la taille de ces tableaux constitue un obstacle à toute utilisation directe par le chercheur. Une série de méthodes d'analyse statistique adaptées au traitement des données textuelles (analyses textométriques) doivent alors être requises pour aider le chercheur à repérer des points saillants, liés à la répartition statistique des contenus au sein du corpus, et, si possible à établir des synthèses statistiques qui constitueront des points de départ précieux pour la mise en évidence des principales dimensions de variation du corpus analysé.***
>
> ***Les contrastes entre textes (ou parties de textes) mis en évidence au plan textométrique trouvent leur explication dans différents plans d'analyse textuelle (lexicologie, syntaxe, sémantique, stylistique, analyse du discours, etc.). Le retour systématique au texte s'impose alors comme seule possibilité de contrôle et de validation des différences entrevues au plan statistique. Pour cette raison, les logiciels de textométrie articulent en général plusieurs ensembles de méthodes : certaines sont destinées à produire des synthèses statistiques ; les autres sont mobilisables pour obtenir des restitutions du contexte, organisées autour des points saillants du texte mises en évidence par les premières.***

L'objectif principal de la textométrie est de compter des éléments (des *contenus* textuels) dans des ensembles (des *contenants* regroupant des unités élémentaires d'un texte ou des zones de texte couvrant un certain nombre ou un certain type d'unités élémentaires).

Les *contenus* se réalisent sous la forme de ressources textuelles (une séquence de caractères organisée en phrases, en paragraphes etc.).

Les *contenants* existent sous la forme de système de masques ou de calques que l'on peut définir sur les contenus. Il s'agit de systèmes d'annotations que l'on peut définir sur tout ou partie des zones textuelles, ces annotations constituant en retour des accès sur les parties textuelles qu'elles définissent (le marquage des phrases ou des paragraphes étant un exemple d'annotation particulier pour décrire un certain niveau de la structure du texte).

Le processus de comptage nécessite au préalable d'identifier[1] les *contenus* et les *contenants*. Ce préalable consiste à expliciter une *segmentation* du texte conduisant à la mise au jour d'une ***Trame*** sur laquelle des annotations pourront se greffer ultérieurement.

[1] On verra aussi qu'il est possible avec le *Trameur* de revenir sur les paramétrages initiaux permettant par exemple de définir la *Trame* en modifiant ultérieurement et dynamiquement les unités initialement produites.

2. Constituer des ressources textométriques (coder les textes)

L'utilisation d'un outil de textométrie nécessite en amont de coder les textes (un éditeur classique suffit en général) en intégrant les textes à traiter et les différents éléments permettant d'organiser ces textes et pour y mettre au jour les *contenus* et les *contenants*.

2.1 Exemple avec Lexico3

Dans le cas de **Lexico3** par exemple, cette étape est largement décrite dans la documentation du logiciel ; nous en intégrons ci-dessous un long extrait :

Source : Manuel Lexico3 http://www.tal.univ-paris3.fr/lexico/

> *L'analyse lexicométrique étudie la répartition dans les textes d'unités complexes (lemmes, segments répétés, cooccurrences, types généralisés). Cependant, une segmentation en formes graphiques constitue une première étape nécessaire qui permet tout à la fois :*
>
> - *d'obtenir une première estimation des principales caractéristiques lexicométriques du corpus (nombre d'occurrences, de formes, d'hapax, fréquence maximale);*
> - *de réaliser les premières typologies sur les parties du corpus;*
> - *de localiser les erreurs qui subsistent après les premières corrections.*
>
> *Pour réaliser cette segmentation en formes graphiques, il faut définir des normes. Dans le cas de **Lexico3**, ces normes sont particulièrement simples.*
>
> *Le texte doit être enregistré sous la forme d'un fichier **texte seulement** (*.txt)*[2].
>
> ***Délimiteurs / non-délimiteurs***
>
> *Dans le corpus soumis aux traitements lexicométriques, une forme graphique est une suite de caractères non-délimiteurs, encadrée par deux caractères délimiteurs. Cela veut dire que les formes graphiques -dont on va décompter les occurrences- sont entièrement définies par la liste des délimiteurs retenus par l'utilisateur. L'identification se produit lorsque les chaînes comprises entre deux délimiteurs de formes sont identiques. Ainsi, si l'on n'effectue pas de prétraitement approprié, Vache n'est pas identique à vache, et c'est-à-dire est différent de c'est à dire.*
>
> *La partie mécanique de la segmentation automatique est considérablement simplifiée par la réalisation du principe simple :*
>
> *un signe = un statut*
>
> *Chaque signe typographique doit être susceptible de recevoir un statut (délimiteur ou non délimiteur) fixé une fois pour toutes au début de la procédure.*

[2] On écarte les fichiers de type document (*.doc) et autres formats créés par traitement de texte, car ceux-ci intègrent un en-tête renfermant diverses informations, sur la mise en forme notamment.

Ces principes entrent parfois en conflit avec les conventions typographiques usuelles. Le tiret de dit-il n'a pas le même statut grammatical que celui de garde-manger. Il en va de même pour l'apostrophe de aujourd'hui (qui devrait être considéré comme non-délimiteur), dont le statut diffère de celle qu'on trouve dans la séquence j'aime.

Lexico3 *propose une liste de délimiteurs par défaut qui peut être modifiée par l'utilisateur: -—_:;/.,?¿!¡*$"+=(){}. L'espace (blanc) est automatiquement ajouté à cette liste. Une fois la liste des délimiteurs fixée, les autres caractères : a, b, c,... deviennent des caractères non-délimiteurs.*

Toute suite de caractères non-délimiteurs, bornée à ses deux extrémités par des délimiteurs, est considérée comme l'occurrence d'une forme à relever et à classer.

Minuscules, majuscules, apostrophes

*Pour des visées particulières, l'utilisateur peut combiner les conventions de pré-traitement et les options de segmentation pour influer sur le type des formes produites par la segmentation. On peut par exemple remplacer systématiquement, lors d'un traitement préalable, toutes les majuscules du texte par une astérisque suivie de la minuscule correspondante (ex : Moi devient *moi). Une segmentation intégrant le caractère * parmi les délimiteurs confondra les occurrences des séquences Moi et moi; une segmentation pour laquelle l'astérisque n'appartient pas à cette liste produira des décomptes distincts pour les deux séquences.*

Sections du texte

Au-delà des partitions logiques repérables dans le texte, celui-ci contient des marques de rupture qui constituent sa respiration (phrases, paragraphes, etc.). ***Lexico3*** *offre la possibilité de promouvoir un caractère délimiteur (ou plusieurs d'entre eux) au rang de délimiteur de section. Ce précodage permet d'étudier ensuite la répartition des occurrences d'une unité lexicométrique parmi les sections ainsi constituées.*

NB : L'insertion systématique de caractères délimitant des sections peut être réalisée en utilisant la fonction ***Remplacer*** *d'un traitement de texte[3].*

Clés /Balises

Au cours d'une étude lexicométrique, on cherche à comparer les fréquences des formes dans les différentes parties d'un corpus. Pour rendre possible ces comparaisons, le texte doit comporter des balises indiquant les délimitations logiques du corpus.

Les parties définies par l'utilisateur peuvent être chronologiques, comme dans l'exemple du Père Duchesne, (cf. section 1.2, « Prise en main rapide ») mais aussi thématiques.

Coder une clé.

Une clé (ex : <Auteur=Dupond>) se compose de 5 éléments :

---

[3] On changera ici systématiquement les caractères retour-chariot par la séquence retour-chariot+blanc+caractère §.

1	**<**	*un chevron ouvrant*
2	***Auteur***	*le **type** de la clé*
3	**=**	*le signe "égal"*
4	***Dupond***	*le **contenu** de la clé*
5	**>**	*un chevron fermant*

Exemples : <Année=1998>, <Auteur=Jean_de_la_Fontaine>

L'insertion de clés constitue une phase importante dans la préparation du texte. Les clés introduites permettront ensuite à l'utilisateur d'effectuer des comparaisons à partir des parties du corpus qu'elles découpent.

Normes de dépouillement

Pour procéder à des dépouillements statistiques à partir des textes ainsi stockés, il faut définir des normes d'identification des unités textuelles. Comment identifier les occurrences d'un même type au fil du texte ? Plusieurs normes sont envisageables qui s'appuient chacune sur des savoirs, des pratiques, des perspectives différentes.

*Le dépouillement en **formes graphiques** (identification automatique des occurrences d'une même chaînes de caractères) est particulièrement facile à décrire et à mettre en œuvre.*

*Le dépouillement en **lemmes**, s'appuie sur des ressources extérieures (dictionnaires de lemmes, analyseurs syntaxiques).*

*Certains logiciels proposent également le regroupement d'occurrences qui peuvent être rapportées à une même **racine** ou **n-gramme** à l'aide de processus plus au moins automatisés.*

*Au-delà du simple dépouillement en formes graphiques, **Lexico3** permet de recenser différents types d'unités textuelles.*

*Les **segments répétés** : suites de formes graphiques identiques attestées plusieurs fois dans le texte).*

*Les **cooccurrences** : couples de formes présentes dans les mêmes contextes (phrase, sections, etc..*

*Les **types généralisés** ou **Tgen(s)**: unités de dépouillement définies par l'utilisateur à l'aide d'outils lui permettant d'effectuer automatiquement des regroupements d'occurrences du texte (ex : les occurrences des formes qui commencent par la séquence de caractère patr : patrie, patriotes, patriotisme, etc.).*

Exemple : le corpus Duchesne

***Tableau 1 :** Exemple de codage de corpus*

```
<An=1793> <Numero=220> <S03=0> <Epg=1>
```

```
§ la grande colère du *père *duchesne , de voir que les mouchards de *la-*fayette et
tous les fripons soudoyés par la liste civile, veulent rétablir les compagnies de
grenadiers et de chasseurs, pour égorger les *sans-culottes et les chasser des
assemblées de *section .ses bons avis aux *lurons des *faubourgs pour qu' ils arrachent
les moustaches postiches à ces grenadiers de la vierge *marie , qui veulent rétablir la
royauté.
<S03=1>
§ millions de tonnerre, nous ne mettrons donc jamais les fripons à la raison ? ils
<Epg=2>ont laissé tomber leurs masques et nous les voyons à nu. serons nous encore dupes
des fripons? quand je voulais faire la conduite de *grenoble à tous les talons rouges
quand je disais, du soir au matin, que tous les ci-devant ne cesseraient de nous trahir,
n' avais je pas raison, foutre?
§ je me suis toujours plus défié des nobles convertis que des émigrés. c' est pour nous
frapper de plus près que ces gredins sont restés au milieu de nous. ils ont fait les
chiens couchants pour mieux nous tromper. jamais, foutre, ils n' ont cessé de s'
entendre avec les ennemis du dehors. ce sont eux qui nous ont mis à chien et à chat, qui
ont brouillé les cartes dans les trois assemblées nationales, et corrompu les
représentants du peuple. si nous avions eu assez d' estoc pour les envoyer tous à
*coblentz au commencement de la révolution, nous n' aurions pas acheté notre liberté par
des flots de sang; nous aurions depuis longtemps une constitution; la paix et le bonheur
régneraient dans notre république.
```

Dans ces fichiers-textes fournis à titre d'exemple,

la clé Sda permet de coder l'année durant laquelle le texte a été publié.

la clé Numero permet d'introduire un numéro de livraison qui respecte l'édition originale du texte (96 livraisons numérotées de 255 à 351 pour le corpus DUCHn.txt, 6 numéros pour le sous-corpus text1.txt).

la clé Epg permet le passage à une autre page, conformément à la pagination de l'édition originale du corpus.

la clé S03 permet de distinguer les portions de texte qui sont des titres et des chapeaux (S03=0) du texte proprement dit (S03=1).

le caractère paragraphe § marque le début de chacun des paragraphes du texte.

*le caractère * permet d'identifier les majuscules du document original.*

2.2 Exemple avec Le Trameur

Le balisage du texte est une étape cruciale avant l'utilisation des outils textométriques, il nécessite des choix qui conditionnent les traitements ultérieurs.

Les balises insérées au fil du texte induisent des systèmes de parties permettant ensuite d'examiner des contrastes de vocabulaires dans les parties ainsi définies. Nous intruisons ci-dessous les différents systèmes de marquages de partie permis avec Le ***Trameur*** .

Dans les paragraphes qui suivent, le processus balisage du corpus est associé à la notion de ***Cadre*** qui sera explicitée *infra*.

2.2.1 Format texte brut

Exemple n°1 :

```
          Le dormeur du val
C'est un trou de verdure où chante une rivière,
Accrochant follement aux herbes des haillons
D'argent ; où le soleil, de la montagne fière,
Luit : c'est un petit val qui mousse de rayons.

Un soldat jeune, bouche ouverte, tête nue,
Et la nuque baignant dans le frais cresson bleu,
Dort ; il est étendu dans l'herbe, sous la nue,
Pâle dans son lit vert où la lumière pleut.

Les pieds dans les glaïeuls, il dort. Souriant comme
Sourirait un enfant malade, il fait un somme :
Nature, berce-le chaudement : il a froid.

Les parfums ne font pas frissonner sa narine ;
Il dort dans le soleil, la main sur sa poitrine,
Tranquille. Il a deux trous rouges au côté droit.

           Arthur Rimbaud
```

Le texte est une simple suite de caractères organisée en lignes. Le texte contient une ***Trame*** mais pas de ***Cadre*** explicite.

On peut cependant envisager de mettre en avant des sections du texte, le caractère "retour à la ligne" permettant d'isoler facilement des zones particulières du texte (les vers, les non-vers).

2.2.2 Format dit « Lexico »

Exemple n°2 :

```
<STRUCTURE="TITRE">Le dormeur du val
<STRUCTURE="TEXTEPOEME">
<GROUPEVERS="1">
<LIGNE="VERS01">C'est un trou de verdure où chante une rivière,
<LIGNE="VERS02">Accrochant follement aux herbes des haillons
<LIGNE="VERS03">D' argent ; où le soleil ; de la montagne fière,
<LIGNE="VERS04">Luit : c' est un petit val qui mousse de rayons.
<GROUPEVERS="2">
<LIGNE="VERS05">Un soldat jeune, bouche ouverte, tête nue,
<LIGNE="VERS06">Et la nuque baignant dans le frais cresson bleu,
<LIGNE="VERS07">Dort ; il est étendu dans l'herbe, sous la nue,
<LIGNE="VERS08">Pâle dans son lit vert où la lumière pleut.
<GROUPEVERS="3">
<LIGNE="VERS09">Les pieds dans les glaïeuls, il dort. Souriant comme
<LIGNE="VERS10">Sourirait un enfant malade, il fait un somme :
<LIGNE="VERS11">Nature, berce-le chaudement : il a froid.
<GROUPEVERS="4">
<LIGNE="VERS12">Les parfums ne font pas frissonner sa narine ;
```

```
<LIGNE="VERS13">Il dort dans le soleil, la main sur sa poitrine,
<LIGNE="VERS14">Tranquille. Il a deux trous rouges au côté droit.
<STRUCTURE="AUTEUR">Arthur Rimbaud
```

Le texte est accompagné ci-dessus d'un ensemble de jalons, marqueurs… délimitant des zones, des parties. Le texte contient une ***Trame*** et un ***Cadre*** (balise de début explicite pour chaque partie, balise de fin implicite).

Ce format se distingue cependant de celui permis dans le logiciel **Lexico3** sur un point très important. Le système de balise utilisé dans le logiciel **Lexico3** permet le recouvrement de parties. Dans l'exemple qui suit (extrait du corpus Le Père Duchesne, *cf* tableau précédent), la balise *page* (<Epg=1>) "traverse" une partie qui la suit au fil du texte (la partie ouverte par la balise <S03=1>) ; dans l'extrait ci-dessous, le texte souligné est celui de la partie (<Epg=1>), ce contenu textuel recouvre pour une part la partie <S03=0> et pour l'autre la partie <S03=1> :

```
<An=1793> <Numero=220> <S03=0> <Epg=1>
§ la grande colère du *père *duchesne , de voir que les mouchards de *la-*fayette et
tous les fripons soudoyés par la liste civile, veulent rétablir les compagnies de
grenadiers et de chasseurs, pour égorger les *sans-culottes et les chasser des
assemblées de *section .ses bons avis aux *lurons des *faubourgs pour qu' ils arrachent
les moustaches postiches à ces grenadiers de la vierge *marie , qui veulent rétablir la
royauté.
<S03=1>
§ millions de tonnerre, nous ne mettrons donc jamais les fripons à la raison ? ils
<Epg=2>ont laissé tomber leurs masques et nous les voyons à nu. serons nous encore dupes
des fripons? quand je voulais faire la conduite de *grenoble à tous les talons rouges
quand je disais, du soir au matin, que tous les ci-devant ne cesseraient de nous trahir,
n' avais je pas raison, foutre?
```

Ce type de recouvrement n'est pas permis dans le ***Trameur*** car toute partie se clôt systématiquement si la plus petite partie qui la contient se termine. Il reste possible de poursuivre la définition d'une partie ultérieurement en la ré-ouvrant à l'endroit idoine.

Dans l'exemple qui suit :

```
<A=1><B=1>.......................
<A=2><B=1>.......................
```

La partie ouverte par la balise ouvrante <B=1> (contenue dans la partie ouverte par la balise <A=1>) est close implicitement par l'ouverture de la balise <A=2> qui clôt aussi la précédente partie A. Le même exemple précédent aurait pu être défini ainsi :

```
<A=1><B=1>.......................</B></A>
<A=2><B=1>.......................</B></A>
```

Dans ce type de marquage des parties, les balises de fin de partie sont explicitement inscrites au fil du texte.

Le ***Trameur*** permet donc d'omettre l'écriture explicite des marqueurs de fin de partie (même s'il semble plus raisonnable de le faire pour faciliter la préparation et la lecture du texte) et dans ce cas il clôt implicitement toute partie contenue dans une partie qui se termine.

L'exemple précédent peut donc être réécrit sous la forme suivante et pris en charge de manière similaire par le ***Trameur*** :

Exemple n°2 avec "balisage explicite des parties" :

```
<STRUCTURE="TITRE">Le dormeur du val</STRUCTURE>
<STRUCTURE="TEXTEPOEME">
<GROUPEVERS="1">
<LIGNE="VERS01">C'est un trou de verdure où chante une rivière,</LIGNE>
<LIGNE="VERS02">Accrochant follement aux herbes des haillons</LIGNE>
<LIGNE="VERS03">D' argent ; où le soleil ; de la montagne fière,</LIGNE>
<LIGNE="VERS04">Luit : c' est un petit val qui mousse de rayons.</LIGNE>
</GROUPEVERS>
<GROUPEVERS="2">
<LIGNE="VERS05">Un soldat jeune, bouche ouverte, tête nue,</LIGNE>
<LIGNE="VERS06">Et la nuque baignant dans le frais cresson bleu,</LIGNE>
<LIGNE="VERS07">Dort ; il est étendu dans l'herbe, sous la nue,</LIGNE>
<LIGNE="VERS08">Pâle dans son lit vert où la lumière pleut.</LIGNE>
</GROUPEVERS>
<GROUPEVERS="3">
<LIGNE="VERS09">Les pieds dans les glaïeuls, il dort. Souriant comme</LIGNE>
<LIGNE="VERS10">Sourirait un enfant malade, il fait un somme :</LIGNE>
<LIGNE="VERS11">Nature, berce-le chaudement : il a froid.</LIGNE>
</GROUPEVERS>
<GROUPEVERS="4">
<LIGNE="VERS12">Les parfums ne font pas frissonner sa narine ;</LIGNE>
<LIGNE="VERS13">Il dort dans le soleil, la main sur sa poitrine,</LIGNE>
<LIGNE="VERS14">Tranquille. Il a deux trous rouges au côté droit.</LIGNE>
</GROUPEVERS>
</STRUCTURE>
<STRUCTURE="AUTEUR">Arthur Rimbaud</STRUCTURE>
```

2.2.3 Format XML

Les documents au format XML sont pris en charge par le ***Trameur***. Le système de balisage permis par ce format offre par définition un système de marquage privilégié pour représenter l'organisation logique des documents.

Exemple n°3 :

```
<?xml version="1.0" encoding="iso-8859-1"?>
<poeme>
<STRUCTURE niveau="TITRE">Le dormeur du val</STRUCTURE>
<STRUCTURE niveau="TEXTEPOEME">
<GROUPEVERS nb="1">
<LIGNE type="VERS01">C'est un trou de verdure où chante une rivière,</LIGNE>
<LIGNE type="VERS02">Accrochant follement aux herbes des haillons</LIGNE>
<LIGNE type="VERS03">D' argent ; où le soleil ; de la montagne fière,</LIGNE>
<LIGNE type="VERS04">Luit : c' est un petit val qui mousse de rayons.</LIGNE>
</GROUPEVERS>
<GROUPEVERS nb="2">
<LIGNE type="VERS05">Un soldat jeune, bouche ouverte, tête nue,</LIGNE>
<LIGNE type="VERS06">Et la nuque baignant dans le frais cresson bleu,</LIGNE>
<LIGNE type="VERS07">Dort ; il est étendu dans l'herbe, sous la nue,</LIGNE>
<LIGNE type="VERS08">Pâle dans son lit vert où la lumière pleut.</LIGNE>
</GROUPEVERS>
<GROUPEVERS nb="3">
<LIGNE type="VERS09">Les pieds dans les glaïeuls, il dort. Souriant comme</LIGNE>
<LIGNE type="VERS10">Sourirait un enfant malade, il fait un somme :</LIGNE>
<LIGNE type="VERS11">Nature, berce-le chaudement : il a froid.</LIGNE>
</GROUPEVERS>
<GROUPEVERS nb="4">
<LIGNE type="VERS12">Les parfums ne font pas frissonner sa narine ;</LIGNE>
<LIGNE type="VERS13">Il dort dans le soleil, la main sur sa poitrine,</LIGNE>
<LIGNE type="VERS14">Tranquille. Il a deux trous rouges au côté droit.</LIGNE>
</GROUPEVERS>
</STRUCTURE>
<STRUCTURE niveau="AUTEUR">Arthur Rimbaud</STRUCTURE>
</poeme>
```

Exemple n°4 :

```
<?xml version="1.0" encoding="iso-8859-1"?>
<TEI>
<teiHeader>
<fileDesc>
```

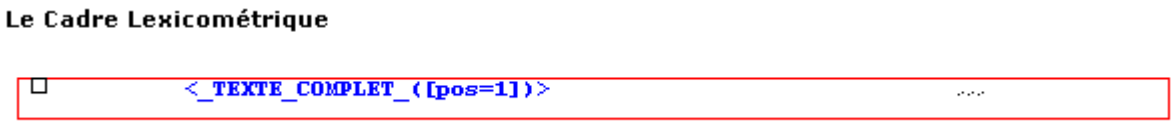
```
<titleStmt>
<title>Poésies</title>
<author>Arthur Rimbaud</author>
</titleStmt>
</fileDesc>
</teiHeader>
<text>
<group>
<text>
<front>
<head>Le Dormeur du Val</head>
<dateline>7 octobre 1870</dateline>
</front>
<body>
<lg>
<lg>
<l>C'est un trou de verdure où chante une rivière</l>
<l>Accrochant follement aux herbes des haillons</l>
<l>D'argent ; où le soleil, de la montagne fière,</l>
<l>Luit : c'est un petit val qui mousse de rayons.</l>
</lg>
<lg>
<l>Un soldat jeune, bouche ouverte, tête nue,</l>
<l>Et la nuque baignant dans le frais cresson bleu,</l>
<l>Dort ; il est étendu dans l'herbe, sous la nue,</l>
<l>Pâle dans son lit vert où la lumière pleut.</l>
</lg>
<lg>
<l>Les pieds dans les glaïeuls, il dort. Souriant comme</l>
<l>Sourirait un enfant malade, il fait un somme :</l>
<l>Nature, berce-le chaudement : il a froid.</l>
</lg>
<lg>
<l>Les parfums ne font pas frissonner sa narine ;</l>
<l>Il dort dans le soleil, la main sur sa poitrine</l>
<l>Tranquille. Il a deux trous rouges au côté droit.</l>
</lg>
</lg>
</body>
</text>
</group>
</text>
</TEI>
```

Dans les deux exemples précédents, le texte contient une ***Trame*** et un ***Cadre*** (chaque partie est parfaitement délimitée : balises de début et de fin explicites).

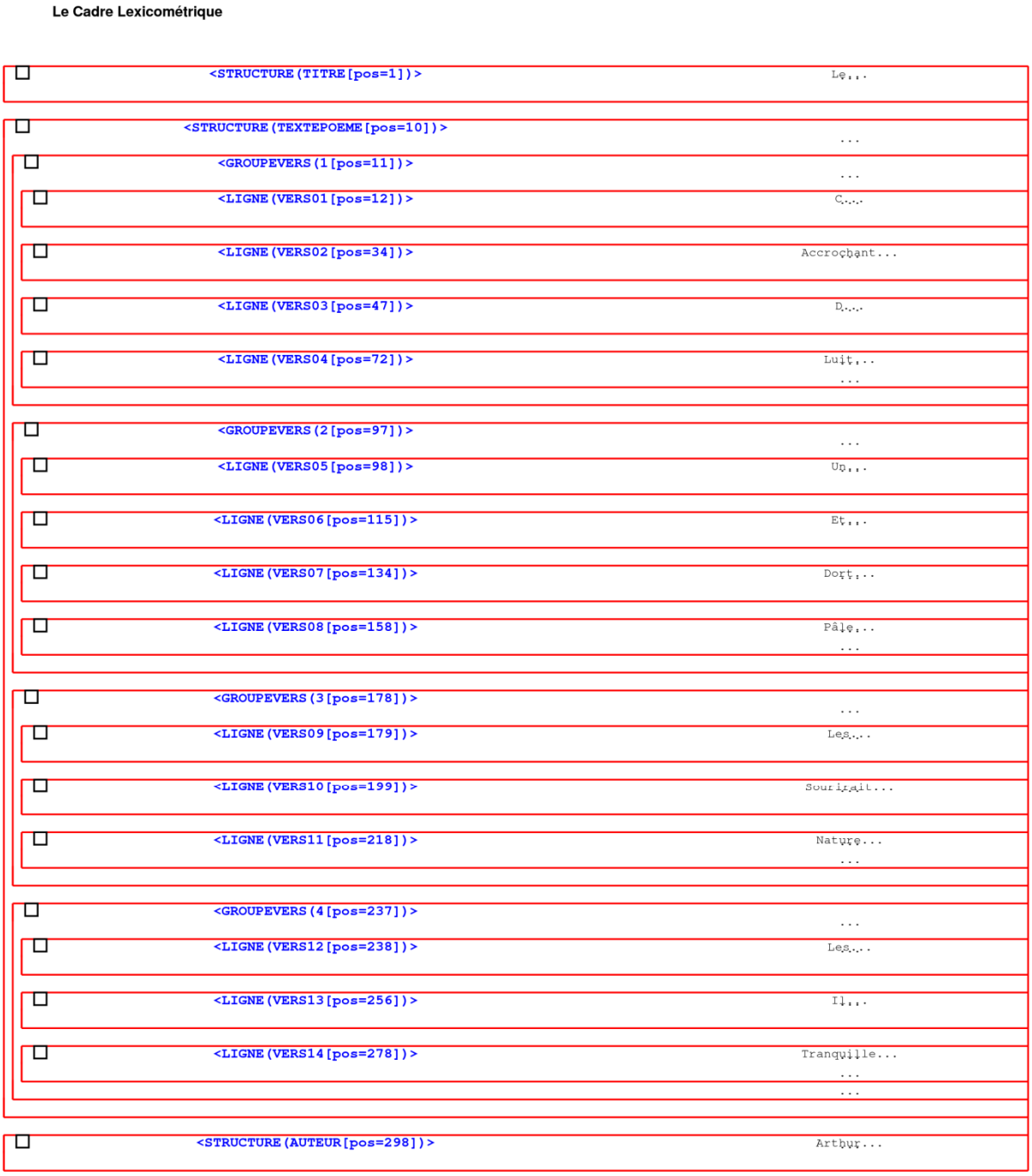
2.2.4 Matérialisation graphique du cadre

Pour les différents exemples précédents, les figures suivantes donnent à voir la représentation graphique construite par le ***Trameur*** du ***Cadre*** (complet) de chacun de ces textes :

Exemple n°1 :

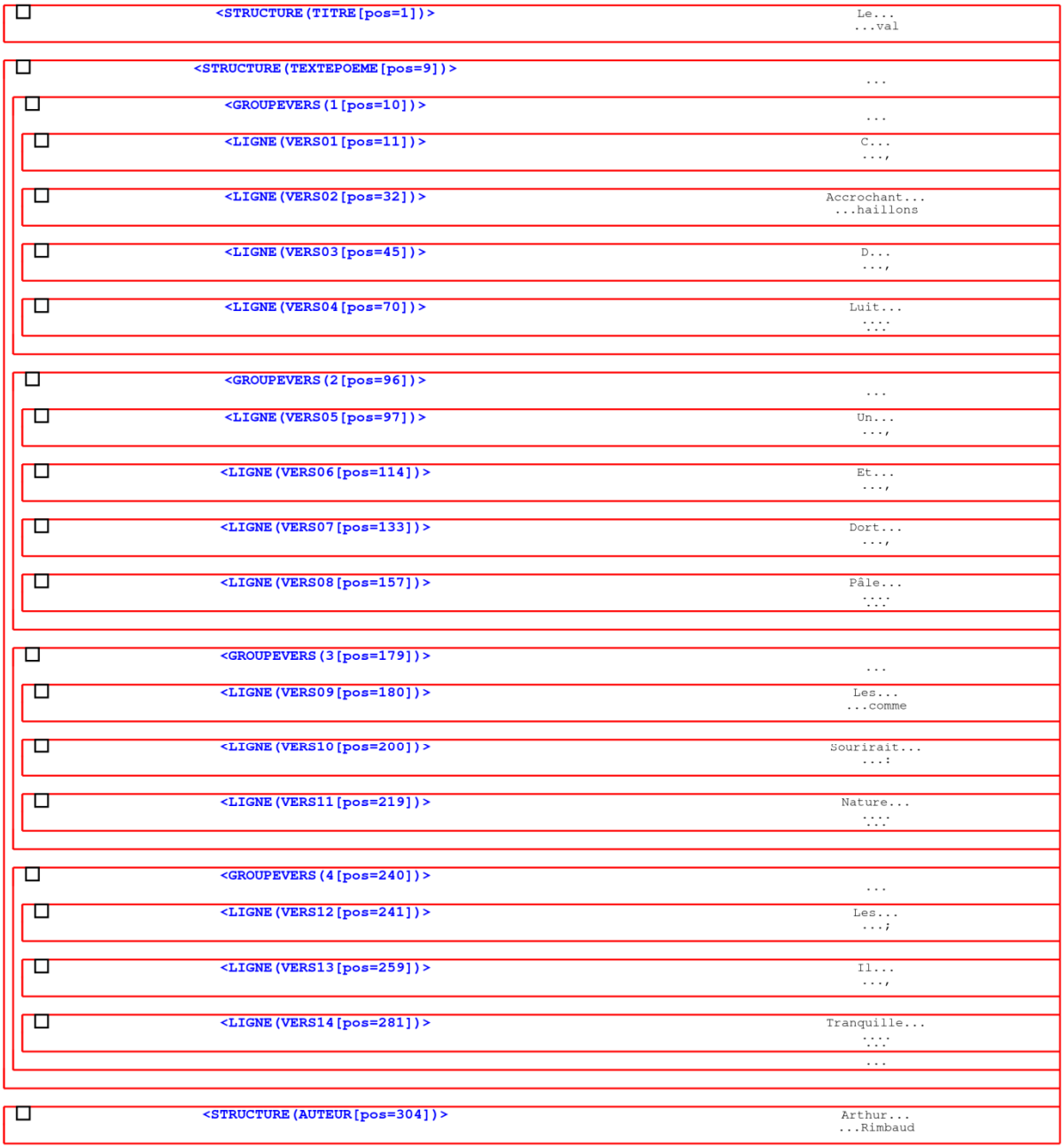


Exemple n°2 :



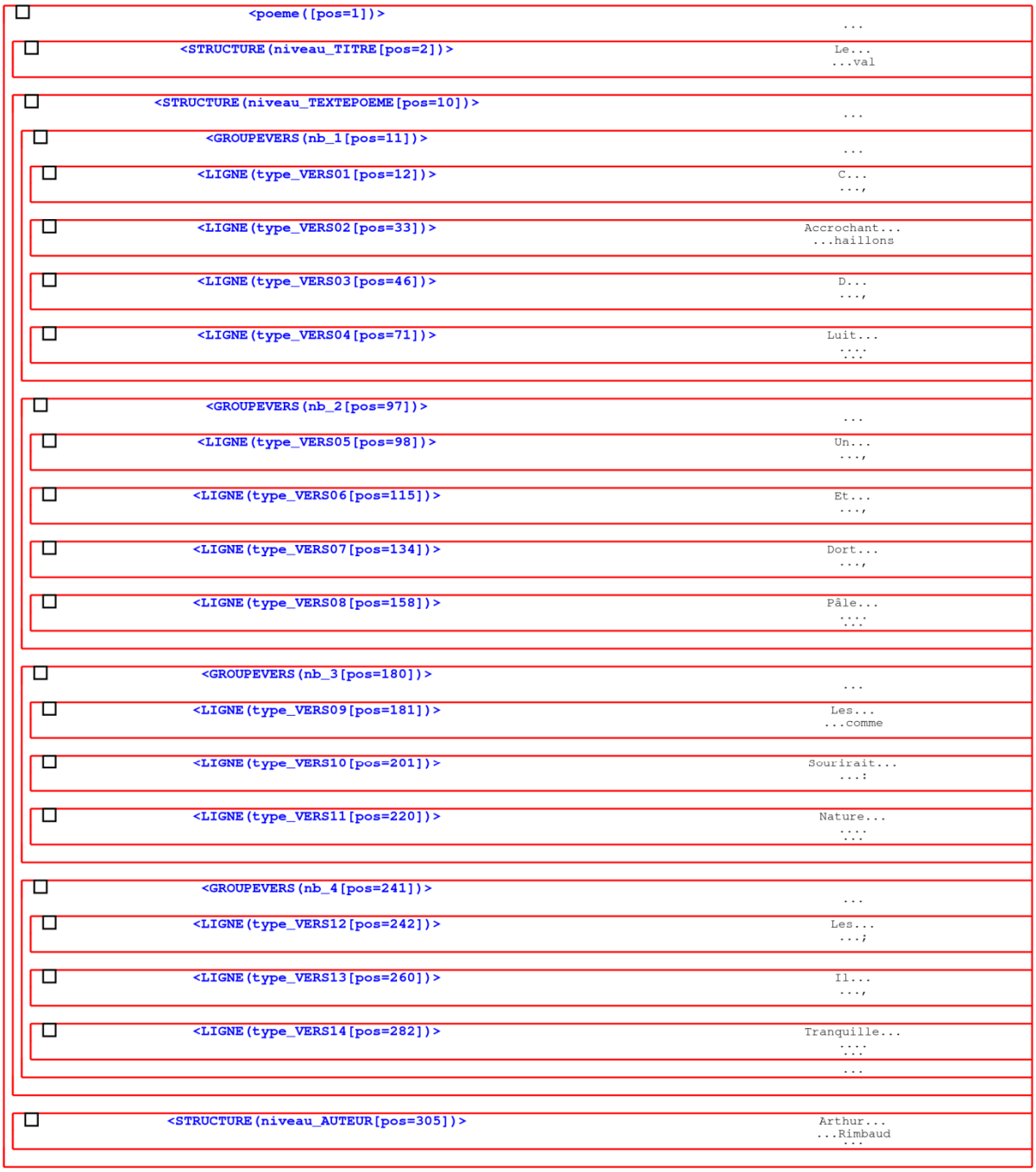
Exemple n°2 avec "balisage explicite des parties"

Le Cadre Lexicométrique



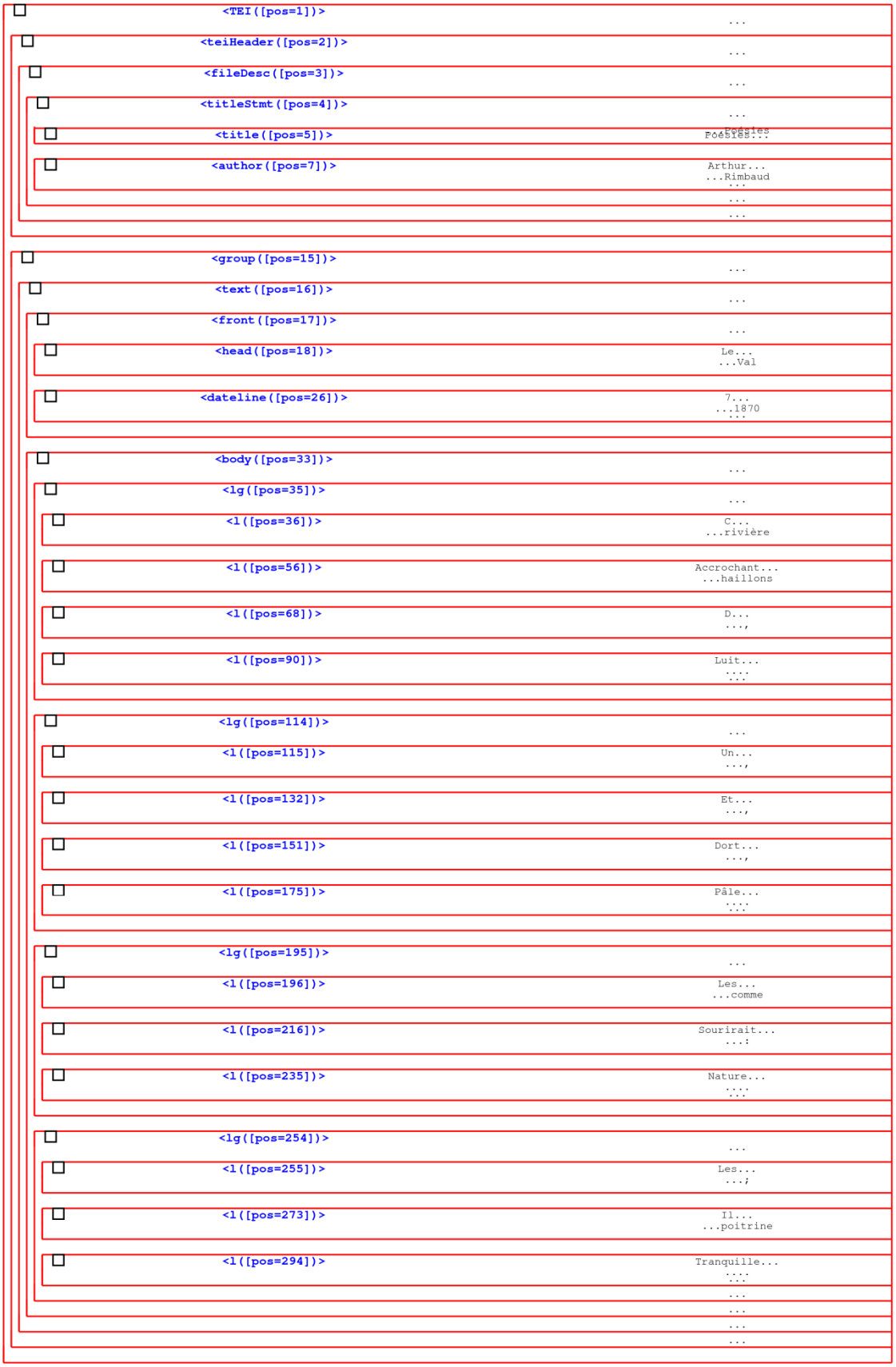
Exemple n°3 :

Le Cadre Lexicométrique



Exemple n°4 :

Le Cadre Lexicométrique



Pour le ***Trameur*** , le système de parties constituant le ***Cadre*** est au final :

- soit une suite de boîtes (une ou plusieurs boîtes) encapsulant d'autres boîtes les unes dans les autres (exemples 1 et 2).

- soit une seule boîte encapsulant d'autres boîtes les unes dans les autres (exemples 2, 3). L'élément racine dans un document au format XML encapsule par définition tous les autres éléments de l'arborescence des éléments dans ce type de format.

Les marqueurs de parties insérés (sous la forme de balise) peuvent respecter le format adopté dans le logiciel Lexico3 ou celui que l'on utilise dans le langage XML (balise+attribut(s)).

3. Les accès aux textes en machine

Les exemples de textes vus dans la partie précédente permettent de mettre en oeuvre différents accès aux différentes zones de contenu textuel disponibles dans chacun des différents formats utilisés pour représenter le même texte.

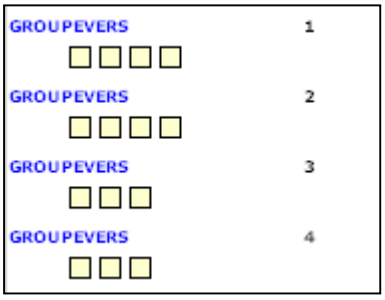
Dans l'exemple 1, l'absence de partitionnement explicite offre cependant la possibilité de promouvoir un caractère délimiteur (le caractère "retour à la ligne") au rang de délimiteur de section. On peut aussi coder le texte en introduisant explicitement un caractère dédié à ce marquage de sections comme dans l'exemple suivant :

```
          Le dormeur du val §
C'est un trou de verdure où chante une rivière, §
Accrochant follement aux herbes des haillons §
D'argent ; où le soleil, de la montagne fière, §
Luit : c'est un petit val qui mousse de rayons. §

Un soldat jeune, bouche ouverte, tête nue, §
Et la nuque baignant dans le frais cresson bleu, §
Dort ; il est étendu dans l'herbe, sous la nue, §
Pâle dans son lit vert où la lumière pleut. §

Les pieds dans les glaïeuls, il dort. Souriant comme §
Sourirait un enfant malade, il fait un somme : §
Nature, berce-le chaudement : il a froid. §

Les parfums ne font pas frissonner sa narine ; §
Il dort dans le soleil, la main sur sa poitrine, §
Tranquille. Il a deux trous rouges au côté droit. §

           Arthur Rimbaud §
```

Ce marquage de sections permet de différentier les contenus textuels de chacune des zones induites ; on peut en construire une représentation graphique sous la forme d'une carte de sections :

Dans ce graphique chaque carré correspond à une zone textuelle se terminant par le caractère délimiteur de section choisi (§). Ce mode d'accès au texte est ici rendu possible par un choix fait en amont d'introduire un caractère particulier dans le codage du texte.

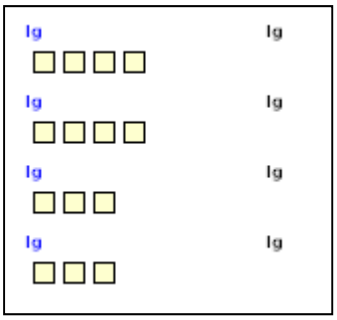
L'introduction de partitions logiques dans le texte permet d'aller plus loin dans le repérage des accès au texte. Dans le graphique suivant (cf exemple 2 supra) :



Les balises LIGNE agissent aussi comme des délimiteurs de sections (le contenu textuel compris entre la balise d'ouverture LIGNE et celle de fermeture). On peut en outre positionner ces sections

dans le partitionnement logique mis en œuvre : ces sections (vers) sont situées dans une partie les contenant qualifiant le groupe de vers.

Idem ici avec l'exemple 4 :



Dans les deux cartes de sections présentées ici, il est important de souligner que les accès mis au jour permettent de focaliser sur une *sélection* de zones textuelles : celles associées aux délimiteurs de sections choisis. Les autres zones textuelles sont inhibées et non prises en compte dans les calculs que l'on pourrait faire sur les sections mises en avant.

L'enrichissement du corpus en amont de son traitement permet de coder l'organisation logique des contenus textuels. Il permet aussi d'y définir des sections particulières du texte.

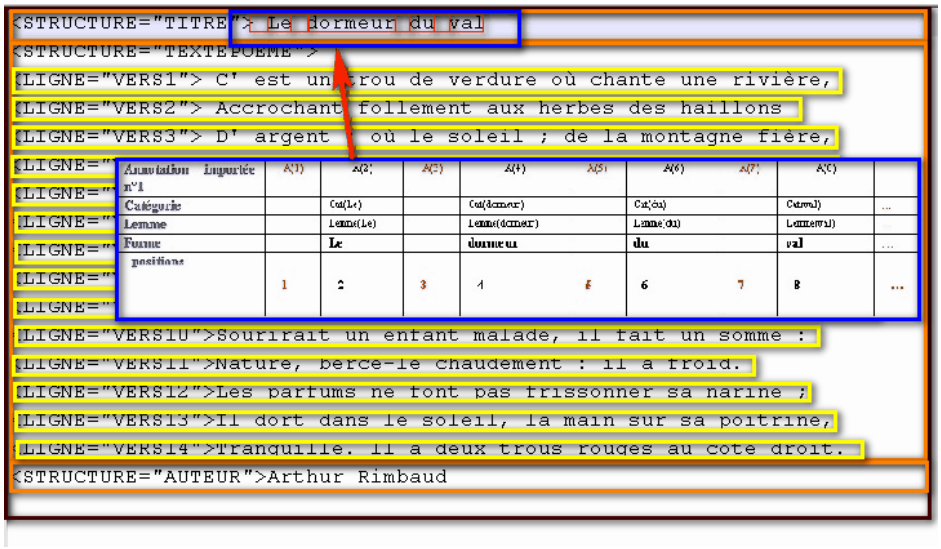
Ces opérations de codage du texte, permettant de mettre en place des points d'accès utilisables ensuite pour analyser l'ensemble du texte, nécessitent un travail en amont au cours duquel il est nécessaire de définir et de marquer précisément les zones textuelles sur lesquelles des accès devront être disponibles pour l'analyse.

4. Les objets textométriques

4.1 Trame et *Cadre* textométriques

Source : [Söze-Duval, 2008]

*A partir d'un texte segmenté, la numérotation des items découpés dans le texte de départ permet de constituer un système de coordonnées sur le texte dans lequel chaque item est repéré par son numéro d'ordre. Nous appelons ce système de coordonnées sur la séquence textuelle : la **Trame** textométrique. Ce même système de coordonnées permet de définir et de localiser, au sein du corpus, des zones textuelles (zones formées par une suite d'items consécutifs, entre la position x_1 la position x_2, réunion d'un certain nombre de zones de ce type, etc.). La définition d'une **Trame** textométrique sur un corpus de textes permet de décrire les systèmes de zones qui correspondent aux contenants de l'analyse textométrique (parties, paragraphes, phrases, sections, chapitres etc.). On peut rassembler les descriptions relatives aux systèmes de contenants dans une structure de données particulière le **Cadre** textométrique.*



*Nous appelons Trameur l'outil informatique qui permet de construire une ressource textométrique **Trame/Cadre**, à partir d'un ensemble de textes rassemblés en corpus, selon les principes définis plus haut. La partie **Trame** de la ressource textuelle produite par le trameur est constituée par la suite des items isolés lors de l'opération de segmentation. La partie **Cadre** rassemble les données relatives aux différents découpages réalisés sur le corpus ainsi que les annotations déjà projetées sur les contenants ou les contenus constitués lors des étapes précédentes.*

*La transmission d'une ressource textuelle constituée sous la forme **Trame/Cadre** constitue une solution suffisante pour servir de base à toute exploration textométrique ultérieure.*

La textométrie s'intéresse à des opérations de comptage sur les textes. Plus précisément, il s'agit de compter des éléments (des unités textuelles) présents dans des ensembles (zones regroupant ces unités textuelles) pour mettre au jour des phénomènes particuliers liés aux comportements de ces unités dans les différentes zones textuelles mises au jour dans les textes.

- on va compter des répartitions d'unités dans certaines parties
- contraster des parties suivant leur vocabulaire
- évaluer l'évolution de certaines unités au fil des textes
- etc.

Une première étape consiste donc à identifier ces unités. Ces unités se réalisent sous la forme d'un flux textuel qui est implicitement ou explicitement organisé en systèmes regroupant ces unités : les phrases, les paragraphes, des parties etc.

Une seconde étape consiste à mettre au jour l'organisation des unités textuelles (c'est-à-dire identifier les parties induites par le texte à analyser). Cette phase de travail est cruciale pour la démarche textométrique et elle peut se révéler « difficile » à mettre en œuvre.

Une fois les parties du texte identifiées, une étape de marquage des parties du texte est nécessaire en amont pour que le repérage en aval dans les outils de ces parties soit possible.

Proposition n°0 : le texte induit à la fois une trame et un cadre

Dans l'approche textométrique, le texte « codé » en machine résulte de la prise en compte de 2 notions :

> Le texte est « par essence » un flux de caractères ; la présence de certains caractères délimiteurs permet par exemple de dire que le texte est un flux séquentiel de mots : les unités traitées résultent du découpage (de la segmentation) du flux textuel sur la base des caractères délimiteurs définis et/ou disponibles.
> Le texte peut aussi induire de manière plus ou moins naturelle des parties qui structurent son organisation. La mise au jour implicite ou explicite des différentes zones du texte permet de découper le texte en système de parties : le texte peut ainsi être vu et traité suivant les différents accès que constituent les différentes partitions rendues disponibles.

Ces 2 aspects sont pris en compte dans notre approche textométrique sous la forme de deux objets : la ***Trame***, le ***Cadre***.

4.2 Sélection

Source : [Söze-Duval, 2008]

> ***La sélection est un objet textométrique qui permet de décrire de manière homogène la plupart des objets sur lesquels portent les opérations textométriques. Une sélection est un sous-ensemble d'items appartenant à la trame. Cette définition permet de rassembler dans une même classe d'objets, dont la description et la gestion informatique peuvent être analogues, plusieurs objets textométriques qui jouent des rôles différents dans l'analyse automatisée des textes.***

En adoptant cette définition, on peut tout à la fois considérer comme des sélections :

- *des sélections-types, unités qui correspondent plutôt à des contenus que l'on constitue en réunissant un certain nombre d'items de la trame (par exemple : l'ensemble des occurrences d'une même, forme graphique, d'un même lemme, les items qui correspondent à une expression régulière opérant sur un ou plusieurs des niveaux d'annotation, etc.).*

- *des sélections-topes, unités qui correspondent plutôt à la notion de contenants (zones, phrases, paragraphes, sections, chapitres, parties du texte, etc.) constituées de séquences d'items connexes, plutôt sélectionnés en fonction de leur position dans le texte.*

Mais cette définition générique permet également de considérer comme des sélections :

- *tout marquage d'un sous-ensemble d'items qui résulterait, par exemple, du surlignage introduit par un intervenant extérieur, sur la base de critères possiblement complexes connus de lui, voire impossibles à expliciter formellement ;*

- *toute zone textuelle mise en évidence au sein du corpus à la suite d'un calcul textométrique ;*

- *tout ensemble d'unités textuelles, sélectionnées à la suite d'un calcul textométrique visant à mettre en évidence la répartition particulière d'items lexicaux au sein du corpus ;*

(voir aussi notion de TGEN dans [Salem, 2006])

4.3 Annotations

Source : [Söze-Duval, 2008]

La définition d'une trame textométrique permet de définir et de relier entre elles les annotations sur les items qui constituent le texte. Les annotations peuvent concerner un item isolé ou ensemble d'items réunis dans une sélection. On appelle annotation systématique un système d'annotations dont chacune porte sur l'un des items de la trame textométrique d'un texte. L'annotation qui associe, par exemple, à chaque item un lemme ou une catégorie grammaticale fournit un exemple d'annotation systématique.

Une annotation est décrite par :

- ***son ancrage : le numéro de l'item auquel elle se rattache principalement ;***

- ***son empan : la définition exacte de la zone à laquelle elle se rattache ;***

- ***son type : qui permet de la distinguer d'autres annotations sur la même cible et de la relier à d'autres annotations du même type portant sur des cibles distinctes ;***

- ***son contenu : la valeur de l'annotation attachée au type d'annotation considéré.***

---

5. Propositions d'implémentation dans le Trameur

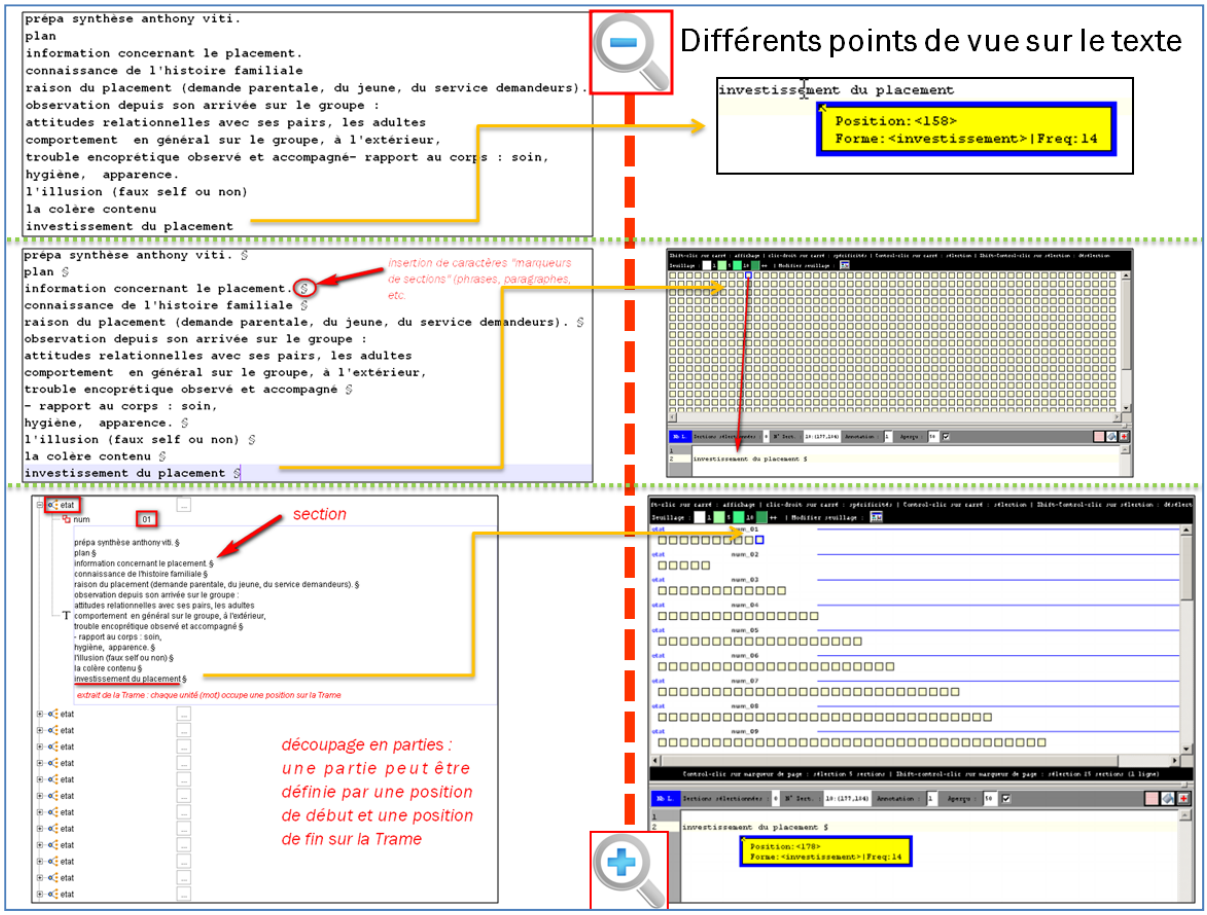
5.1 Préparer les textes aux explorations textométriques

Avant de traiter les données en machine, celles-ci doivent souvent être préparées pour éventuellement mettre au jour les éléments qui conduiront les analyses. Cela implique en amont :

- un codage du texte (choix de découpage des unités, de certains regroupement d'unités en section)
- un codage des parties éventuelles (partitionnements du texte en différents systèmes de parties)

Les traitements ultérieurs s'appuieront sur ces choix initiaux pour en construire une représentation informatique (des dictionnaires, des zones regroupant des unités en partie) et aussi des tableaux lexicaux traduisant des comptages de ces unités dans les différentes parties mises au jour. Il est donc crucial de prêter beaucoup d'attention à cette étape de travail.

La figure suivante illustre sur sa partie gauche 3 états de codage initial du même texte. Le premier état est constitué du seul texte brut.



Le second introduit un caractère particulier permettant de délimiter des sections (à définir).

Enfin le troisième introduit un système de parties (la partition ETAT).

Les opérations textométriques disponibles à chaque état ne sont pas de même nature : les objets textométriques induits sur chaque état n'étant pas de même type.

La partie droite de la figure met au jour certaines opérations textométriques rendues possibles par l'enrichissement des données en amont. On passe de la seule information disponible sur un item de la ***Trame***, à une représentation cartographique de la ***Trame*** (rendue possible par un délimiteur de sections) et enfin à une représentation cartographique de la ***Trame*** (carte des sections) sur laquelle on projette une partition disponible dans le ***Cadre***.

5.2 La *Trame* et le *Cadre* dans le Trameur

Proposition n°1 : le texte comme liste ordonnée de positions

Le texte est constitué d'une suite ordonnée d'unités textuelles (forme ou délimiteur) et éventuellement de marqueurs de parties (exprimés implicitement ou par un système de marques identifiables)
Le texte est représentable sous la forme d'une liste ordonnée de positions (celles des unités textuelles dans le texte)
Les marqueurs de parties définissent des sélections de positions (début et fin de partie)

Proposition n°2 : Trois objets de base (***Trame***, ***Cadre***, ***Sélection***)

Trame : vecteur de positions (de 1 à n), à chaque position on associe une liste d'annotations (forme, lemme, catégorie, etc.)

Sélection : sélections de position sur la ***Trame***
(A) une liste de positions
(B) une liste de listes de couples de positions

Exemple : représentation d'une partition sur le texte

Partition **MOIS** :

MOIS=1 ⟹ (1,100)

i.e la première partie de la partition MOIS couvre les positions de 1 à 100

MOIS=2 ⟹ (101,200)

i.e la seconde partie de la partition MOIS couvre les positions de 101 à 200

Etc.

Partition **SEMAINE :**

SEMAINE=1 ⟹ (1,25)

SEMAINE=2 ⟹ (26, 50)

Etc.

Au final, il est possible de disposer d'un ensemble de partitions sur le texte, chacune étant définie par des ensembles de couples de position :

MOIS->((1,100),(101,200)…)

QUINZAINE ->((1,50),(51,100)…)

SEMAINE->((1,25),(26,50)…),…)

Cadre : instance de sélection de type (B)

Proposition n°3 : le Tramage est un processus de découpage du texte réalisé à partir d'une liste finie de caractères dits « délimiteurs » d'unités

Le processus de ***Tramage*** (segmentation et repérage des systèmes de parties) a juste besoin de connaître la liste des caractères délimiteurs pour construire la ***Trame*** et le ***Cadre***.

Ce processus construit au démarrage une ***Trame*** (vecteur de positions initialement vide) et un objet ***Sélection*** (pour mémoriser le ***Cadre***) ; il a ensuite pour tâche d'enrichir ces 2 objets en mettant au jour au fil du texte les unités repérées : les unités textuelles de la ***Trame*** (forme ou délimiteur) et les marqueurs de parties (position de début ou position de fin d'une partie).

Chaque unité repérée est associé à un numéro d'apparition sur la ***Trame***: la ***Trame*** stocke pour chaque niveau d'annotation des valeurs numériques (numéros d'apparition des unités pour chaque annotation).
Des dictionnaires (pour chaque niveau d'annotation) permettent de stocker les équivalences : n° apparition <=> valeur de l'annotation (forme, lemme, catégories, …)

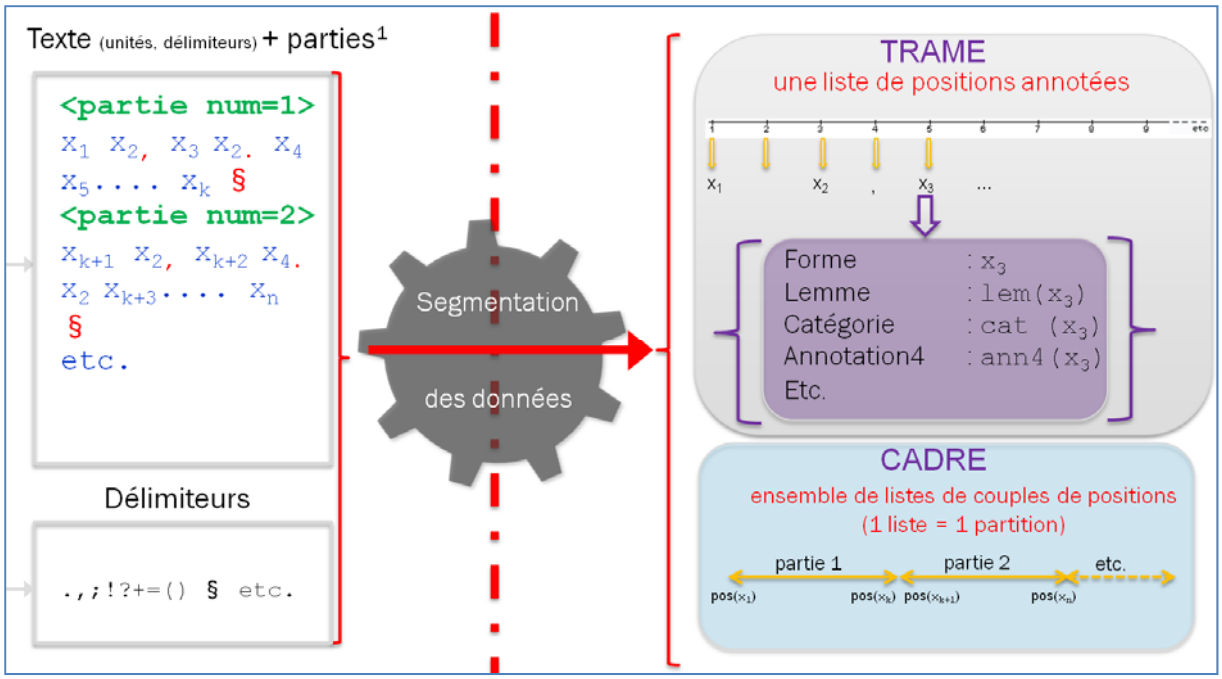
5.3 Les annotations dans le Trameur

Proposition n°4 : annoter les unités de la Trame

La textométrie s'intéresse traditionnellement aux mots du texte (tels qu'ils apparaissent dans les textes produits) voire à des versions lemmatisées de ces textes. L'approche textométrique mise en œuvre dans le logiciel Le ***Trameur*** étend cette situation en permettant d'annoter les unités textuelles d'annotations complémentaires : chaque unité peut être associée à un nombre indéfini d'annotations.

5.4 Implémentation des propositions : le moteur textométrique

La mise en œuvre des ces propositions dans le ***Trameur*** est décrite dans le schéma suivant :



A noter dans ce dessin une simplification dans la représentation du système d'annotation associé à chaque unité de la ***Trame*** : ici, ce sont les valeurs des annotations qui sont indiquées, l'implémentation met en fait en œuvre une numérisation de ces valeurs en leur associant leur numéro d'apparition, et ce sont ces valeurs numériques que portent chaque position de la Trame.

5.3 L'objet Sélection dans le Trameur

Source : [Söze-Duval, 2008]

> ***La sélection est un objet textométrique qui permet de décrire de manière homogène la plupart des objets sur lesquels portent les opérations textométriques. Une sélection est un sous-ensemble d'items appartenant à la Trame.***
>
> ***Pour construire des unités complexes et étudier leur répartition dans différentes zones d'un corpus de textes, un logiciel de textométrie doit disposer d'un gestionnaire de sélection permettant de constituer, de modifier, de gérer, les sélections à partir de toutes les vues possibles sur le texte et les sélections déjà construites. Le gestionnaire de sélections prend appui sur toutes les présentations possibles du texte ou des comptages d'unités réalisés à partir de ce dernier pour définir d'autres unités qui seront elles-mêmes soumises à des évaluations textométriques.***

Le ***Trameur*** dispose de fonctionnalités permettant de sélectionner des positions de la ***Trame***, de les mémoriser et de constituer des unités réutilisables dans des opérations textométriques. Un outil *gestionnaire de sélection* qui permet de travailler à partir de ces sélections.

Le *gestionnaire de sélection* permet de « sélectionner» des unités de la ***Trame*** (via leur position) ou de « sélectionner » des zones de la ***Trame*** (via les parties disponibles). Dans les deux cas, on peut ensuite lancer des opérations sur les sélections réalisées.

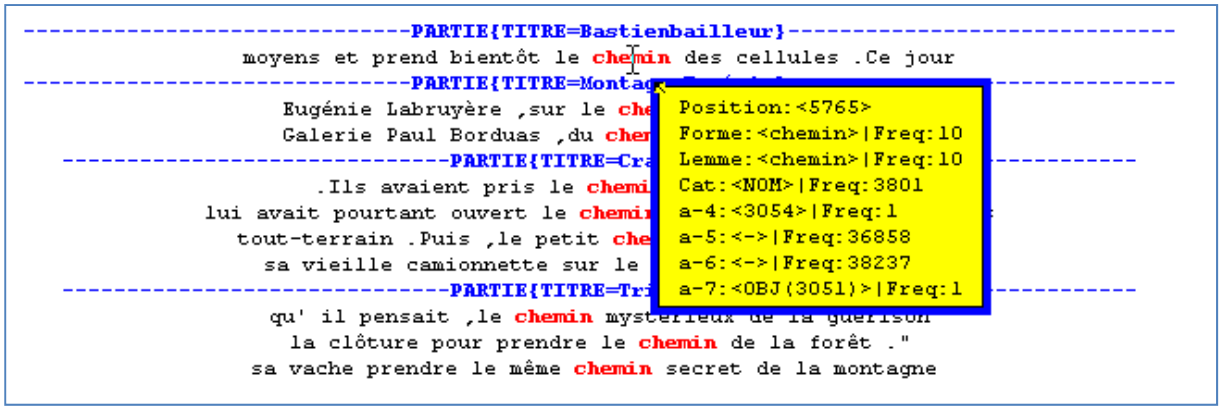
(voir aussi notion de TGEN dans [Salem, 2006])

6. Modifier des ressources textométriques (corriger les textes dans la machine)

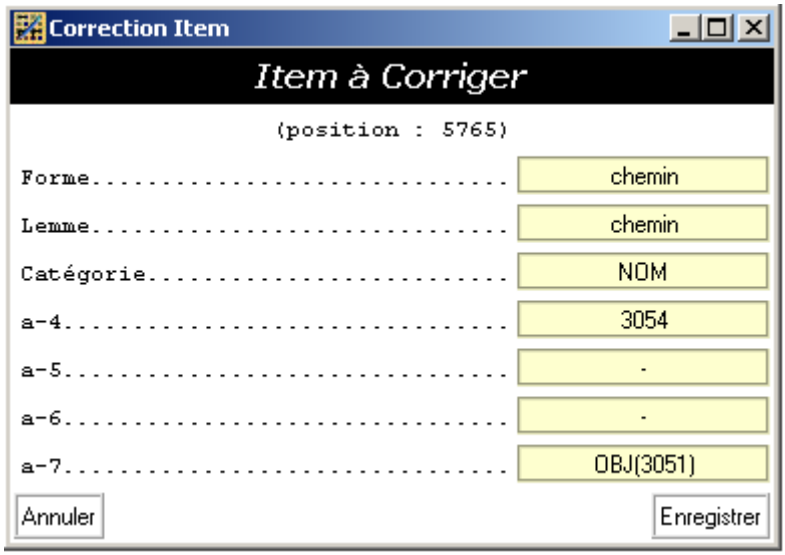
Dans le ***Trameur***, chaque position de la ***Trame*** peut être associée à plusieurs couches d'annotation. Chaque item peut être corrigé dynamiquement (mise à jour de ses annotations) ; on peut aussi insérer ou modifier des annotations en sélectionnant des formes de la ***Trame*** dans le dictionnaire des formes. On peut enfin importer un système d'annotation sur l'intégralité des positions de la ***Trame*** (cf manuel du ***Trameur***)

Mise à jour locale de la Trame

Dans le ***Trameur***, toutes les zones d'édition (partie, section, concordance) constituent des accès au texte i.e à un fragment de la ***Trame*** (de telle position à telle autre). Les contenus de chacune de ces positions sont modifiables par l'utilisateur (modification des différentes valeurs des annotations associées à chacune de ces positions).



La figure précédente met au jour l'édition d'une concordance de la forme graphique « chemin » ; cette concordance affiche de fait les zones de la ***Trame*** contenant la forme visée ; l'item à la position 5765 est décrit par une liste de 7 annotations. L'édition de cet item permet de modifier les valeurs de chacune d'elle :

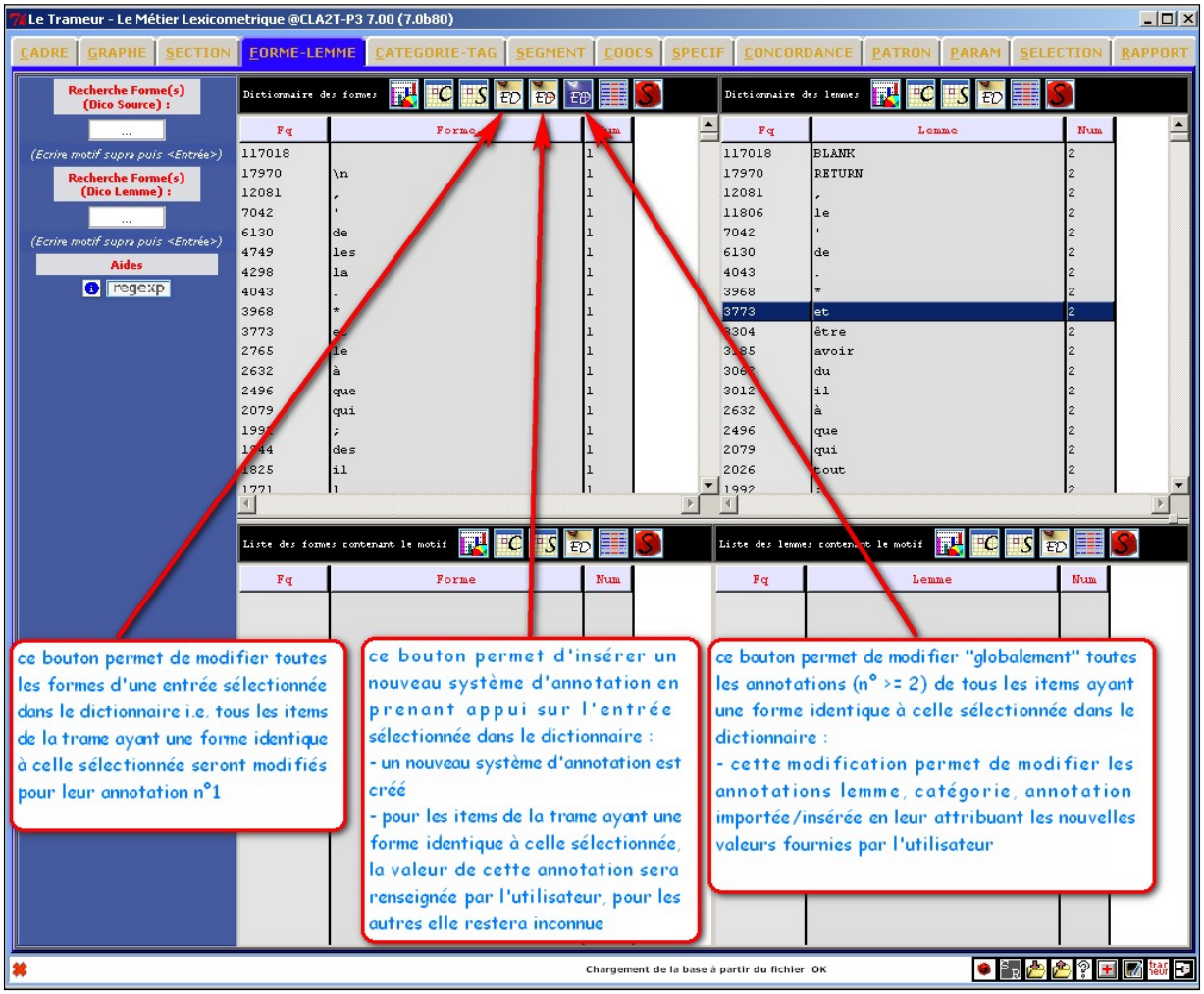


Mise à jour globale de la Trame

Les dictionnaires des différentes couches d'annotation disponibles dans une base textométrique permettent de recenser toutes les occurrences d'une même valeur d'annotation sur toute la ***Trame***.

Dans l'exemple qui suit, le ***Trameur*** a été paramétré pour avoir un étiquetage du texte de travail (3 couches d'annotation sur chaque item) : l'annotation n°1 correspondant à « Forme », l'annotation n°2 correspondant à « Lemme », l'annotation n°3 correspondant à « Catégorie ».

La figure suivante donne à voir les dictionnaires de deux premières couches d'annotation. Elle met aussi en avant les trois opérations disponibles sur le dictionnaire des formes pour globalement modifier ou ajouter des annotations sur les items sélectionnés dans le dictionnaire.

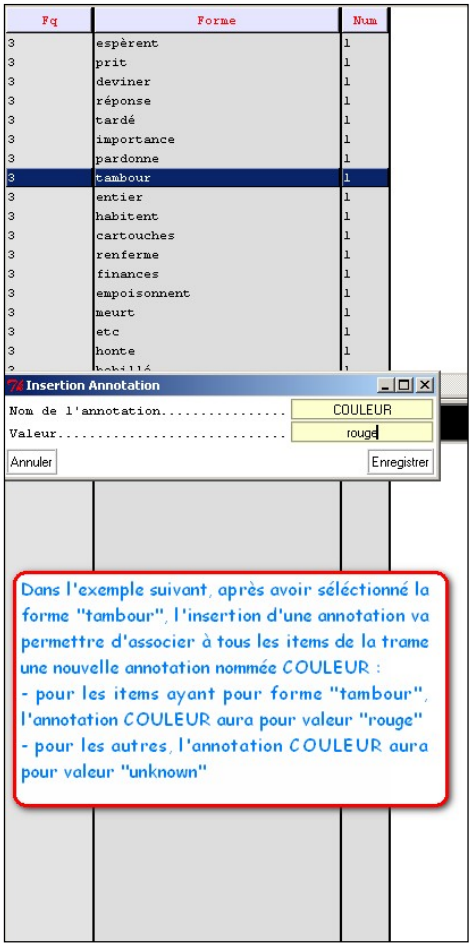


Un item étant sélectionné dans le dictionnaire, on peut donc :

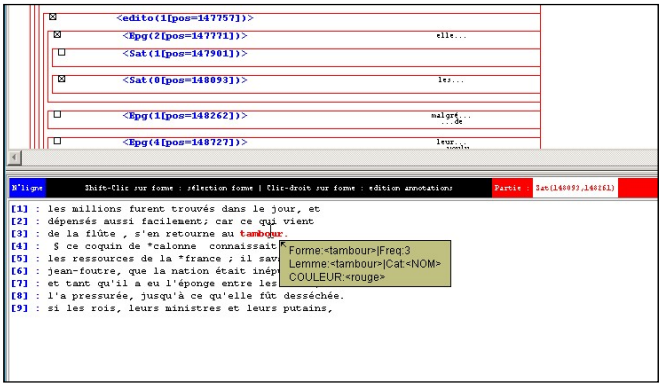
- Modifier toutes les occurrences de la forme sélectionnée
- Insérer une nouvelle annotation sur ces occurrences (*cf* figure suivante)
- Modifier les différentes couches d'annotations de toutes ces occurrences.

Ajouter des annotations sur la Trame

L'exemple qui suit illustre l'opération d'insertion d'une annotation sur la ***Trame***. Un item étant sélectionné dans le dictionnaire, l'opération d'insertion permet de définir le nom de l'annotation et sa valeur pour les items visés.



Phase 1 : ajout d'une annotation



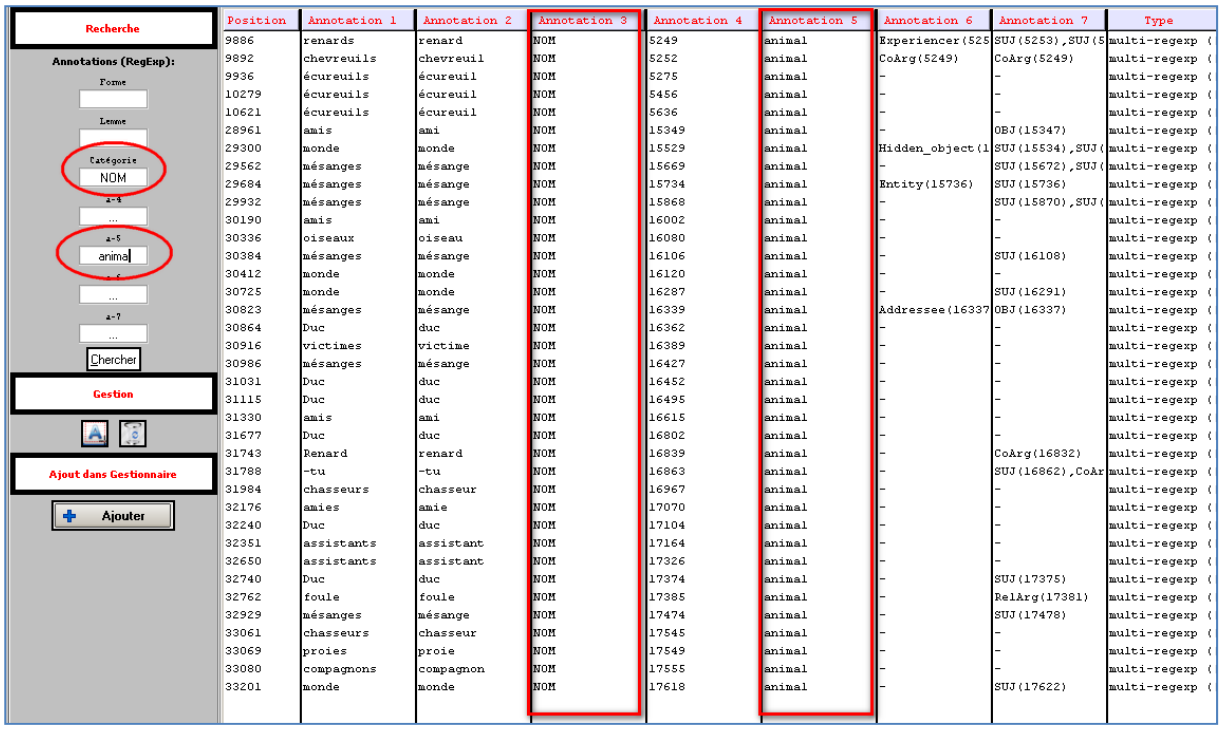
*Phase 2 : visualisation sur la **Trame***

Les différentes opérations ainsi définies permettent de remodeler la ***Trame*** en agissant localement ou globalement sur ses unités.

Mise à jour d'une sélection de positions sur la Trame

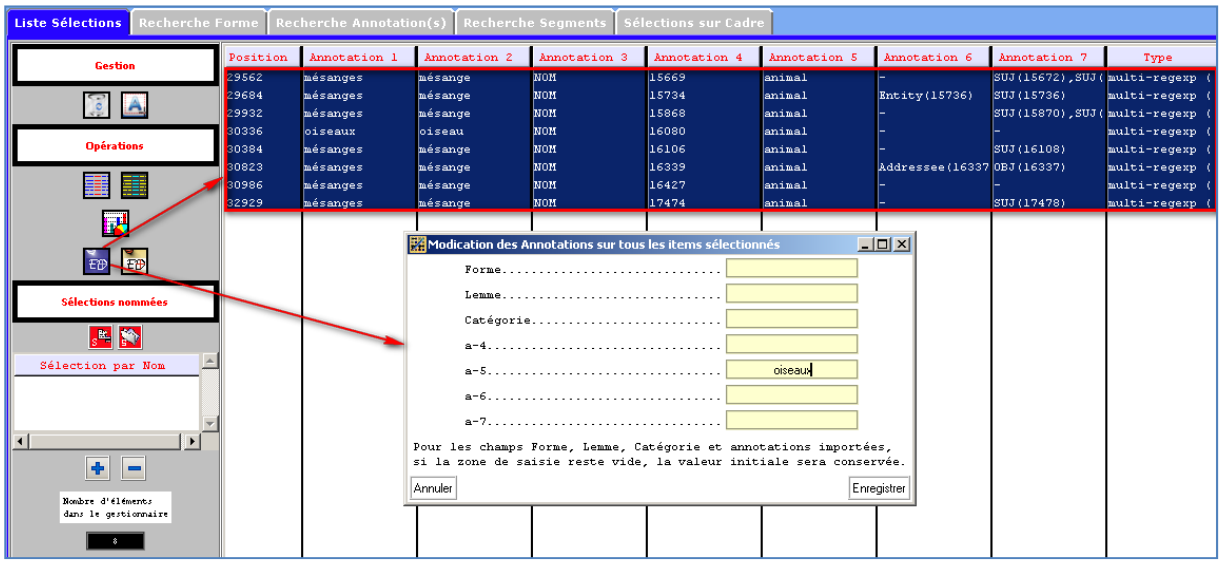
Enfin, le *gestionnaire de sélection* permet de « filtrer » des items de la ***Trame*** ayant telle ou telle propriété (tous les items ayant telle forme graphique et telle annotation morpho-syntaxique). Les items résultants peuvent ensuite être choisis pour des modifications éventuelles sur les valeurs associées à leurs différents niveaux d'annotation (*cf* manuel du ***Trameur***).

La figure suivante illustre l'étape de filtrage des items visés : recherche des items de type NOM (annotation catégorie) et décrivant des animaux (annotation a-4) :



Position	Annotation 1	Annotation 2	Annotation 3	Annotation 4	Annotation 5	Annotation 6	Annotation 7	Type
9886	renards	renard	NOM	5249	animal	Experiencer(525	SUJ(5253),SUJ(5	multi-regexp (
9892	chevreuils	chevreuil	NOM	5252	animal	CoArg(5249)	CoArg(5249)	multi-regexp (
9936	écureuils	écureuil	NOM	5275	animal	-	-	multi-regexp (
10279	écureuils	écureuil	NOM	5456	animal	-	-	multi-regexp (
10621	écureuils	écureuil	NOM	5636	animal	-	-	multi-regexp (
28961	amis	ami	NOM	15349	animal	-	OBJ(15347)	multi-regexp (
29300	monde	monde	NOM	15529	animal	Hidden_object(1	SUJ(15534),SUJ(	multi-regexp (
29562	mésanges	mésange	NOM	15669	animal	-	SUJ(15672),SUJ(	multi-regexp (
29684	mésanges	mésange	NOM	15734	animal	Entity(15736)	SUJ(15736)	multi-regexp (
29932	mésanges	mésange	NOM	15868	animal	-	SUJ(15870),SUJ(	multi-regexp (
30190	amis	ami	NOM	16002	animal	-	-	multi-regexp (
30336	oiseaux	oiseau	NOM	16080	animal	-	-	multi-regexp (
30384	mésanges	mésange	NOM	16106	animal	-	SUJ(16108)	multi-regexp (
30412	monde	monde	NOM	16120	animal	-	-	multi-regexp (
30725	monde	monde	NOM	16287	animal	-	SUJ(16291)	multi-regexp (
30823	mésanges	mésange	NOM	16339	animal	Addressee(16337	OBJ(16337)	multi-regexp (
30864	Duc	duc	NOM	16362	animal	-	-	multi-regexp (
30916	victimes	victime	NOM	16389	animal	-	-	multi-regexp (
30986	mésanges	mésange	NOM	16427	animal	-	-	multi-regexp (
31031	Duc	duc	NOM	16452	animal	-	-	multi-regexp (
31115	Duc	duc	NOM	16495	animal	-	-	multi-regexp (
31330	amis	ami	NOM	16615	animal	-	-	multi-regexp (
31677	Duc	duc	NOM	16802	animal	-	-	multi-regexp (
31743	Renard	renard	NOM	16839	animal	-	CoArg(16832)	multi-regexp (
31788	-tu	-tu	NOM	16863	animal	-	SUJ(16862),CoAr	multi-regexp (
31984	chasseurs	chasseur	NOM	16967	animal	-	-	multi-regexp (
32176	amies	amie	NOM	17070	animal	-	-	multi-regexp (
32240	Duc	duc	NOM	17104	animal	-	-	multi-regexp (
32351	assistants	assistant	NOM	17164	animal	-	-	multi-regexp (
32650	assistants	assistant	NOM	17326	animal	-	-	multi-regexp (
32740	Duc	duc	NOM	17374	animal	-	SUJ(17375)	multi-regexp (
32762	foule	foule	NOM	17385	animal	-	RelArg(17381)	multi-regexp (
32929	mésanges	mésange	NOM	17474	animal	-	SUJ(17478)	multi-regexp (
33061	chasseurs	chasseur	NOM	17545	animal	-	-	multi-regexp (
33069	proies	proie	NOM	17549	animal	-	-	multi-regexp (
33080	compagnons	compagnon	NOM	17555	animal	-	-	multi-regexp (
33201	monde	monde	NOM	17618	animal	-	SUJ(17622)	multi-regexp (

Les items résultants sont répartis sur la Trame (et repérables par leurs positions respectives). La modification de certains de ces items peut être mise en œuvre :



La modification visée ici consiste à modifier la valeur de l'annotation n°5 sur une sélection d'item (changer « animal » par « oiseaux) :

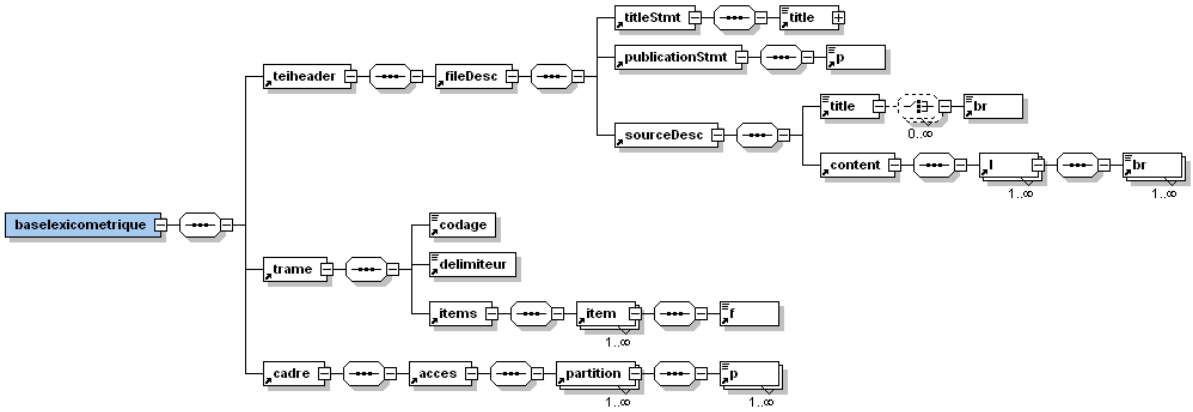
Position	Annotation 1	Annotation 2	Annotation 3	Annotation 4	Annotation 5	Annotation 6	Annotation 7	Type
29562	mésanges	mésange	NOM	15669	oiseaux	-	SUJ(15672),SUJ(	multi-regexp
29684	mésanges	mésange	NOM	15734	oiseaux	Entity(15736)	SUJ(15736)	multi-regexp
29932	mésanges	mésange	NOM	15868	oiseaux	-	SUJ(15870),SUJ(	multi-regexp
30336	oiseaux	oiseau	NOM	16080	oiseaux	-	-	multi-regexp
30384	mésanges	mésange	NOM	16106	oiseaux	-	SUJ(16108)	multi-regexp
30823	mésanges	mésange	NOM	16339	oiseaux	Addressee(16337	OBJ(16337)	multi-regexp
30986	mésanges	mésange	NOM	16427	oiseaux	-	-	multi-regexp
32929	mésanges	mésange	NOM	17474	oiseaux	-	SUJ(17478)	multi-regexp

Dans tous les cas, les modifications réalisées induisent aussi une mise à jour automatique des dictionnaires associées à chacun des niveaux d'annotation disponibles.

Le *gestionnaire de sélections* permet aussi de réaliser l'insertion de nouvelles annotations sur les items sélectionnés (cf manuel du ***Trameur***).

7. Echanges de Bases Textométriques

Le fichier d'export de l'ensemble ***Trame*** + ***Cadre*** sur la version TEI du *Dormeur du Val* a l'allure suivante (schéma XML)



La DTD (grammaire) de la version texte brut du *Dormeur du Val* a l'allure suivante :

```
<!ELEMENT baselexicometrique (teiheader, trame, cadre)>
<!ELEMENT teiheader (fileDesc)>
<!ELEMENT fileDesc (titleStmt, publicationStmt, sourceDesc)>
<!ELEMENT titleStmt (title)>
<!ELEMENT publicationStmt (p)>
<!ELEMENT p (#PCDATA)>
<!ELEMENT sourceDesc (title, content)>
<!ELEMENT content (l+)>
<!ELEMENT l (br+)>
<!ELEMENT br (#PCDATA)>
<!ELEMENT title (#PCDATA | br)*>
<!ELEMENT trame (codage, delimiteur, items)>
<!ELEMENT codage (#PCDATA)>
<!ELEMENT delimiteur (#PCDATA)>
<!ELEMENT items (item+)>
<!ELEMENT item (f)>
<!ELEMENT f (#PCDATA)>
<!ELEMENT cadre (acces)>
<!ELEMENT acces (#PCDATA)>
<!ATTLIST item
        type (delim | forme) #REQUIRED
        pos (1 | 10 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 11 | 110 |
111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 12 | 120 | 121 | 122 | 123 | 124 | 125 |
126 | 127 | 128 | 129 | 13 | 130 | 131 | 132 | 133 | 134 | 135 | 136 | 137 | 138 | 139 | 14 |
140 | 141 | 142 | 143 | 144 | 145 | 146 | 147 | 148 | 149 | 15 | 150 | 151 | 152 | 153 | 154 |
155 | 156 | 157 | 158 | 159 | 16 | 160 | 161 | 162 | 163 | 164 | 165 | 166 | 167 | 168 | 169 |
17 | 170 | 171 | 172 | 173 | 174 | 175 | 176 | 177 | 178 | 179 | 18 | 180 | 181 | 182 | 183 |
184 | 185 | 186 | 187 | 188 | 189 | 19 | 190 | 191 | 192 | 193 | 194 | 195 | 196 | 197 | 198 |
199 | 2 | 20 | 200 | 201 | 202 | 203 | 204 | 205 | 206 | 207 | 208 | 209 | 21 | 210 | 211 |
212 | 213 | 214 | 215 | 216 | 217 | 218 | 219 | 22 | 220 | 221 | 222 | 223 | 224 | 225 | 226 |
227 | 228 | 229 | 23 | 230 | 231 | 232 | 233 | 234 | 235 | 236 | 237 | 238 | 239 | 24 | 240 |
241 | 242 | 243 | 244 | 245 | 246 | 247 | 248 | 249 | 25 | 250 | 251 | 252 | 253 | 254 | 255 |
256 | 257 | 258 | 259 | 26 | 260 | 261 | 262 | 263 | 264 | 265 | 266 | 267 | 268 | 269 | 27 |
270 | 271 | 272 | 273 | 274 | 275 | 276 | 277 | 278 | 279 | 28 | 280 | 281 | 282 | 283 | 284 |
285 | 286 | 287 | 288 | 289 | 29 | 290 | 291 | 292 | 293 | 294 | 295 | 296 | 297 | 298 | 299 |
3 | 30 | 300 | 301 | 302 | 303 | 304 | 305 | 306 | 307 | 308 | 309 | 31 | 310 | 311 | 312 |
313 | 314 | 315 | 316 | 317 | 318 | 319 | 32 | 320 | 321 | 322 | 323 | 324 | 325 | 33 | 34 |
35 | 36 | 37 | 38 | 39 | 4 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 5 | 50 | 51 |
52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 6 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 |
7 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 8 | 80 | 81 | 82 | 83 | 84 | 85 | 86 |
87 | 88 | 89 | 9 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99) #REQUIRED
>
```

Dans ce dernier exemple, la ***Trame*** est clairement identifiable : elle est constituée d'une liste d'items (de type *delim* ou *forme*) associé à une position (*pos* : de 1 à 325) et constitué d'une seule annotation (f : forme graphique de l'unité).

On présente en annexe deux fichiers d'export produits sur 2 formats du texte « Le Dormeur du Val ».

La base exportée est composée d'un entête, de la ***Trame*** et du ***Cadre***.

Pour ces 2 derniers objets, il s'agit d'un reformatage XML des objets ***Trame*** et ***Cadre*** présenté *supra*.

La ***Trame*** est décrite par l'encodage des caractères utilisés, la liste des délimiteurs et la liste des unités (forme ou délimiteur).

Le ***Cadre*** est décrit par une liste de partition, chacune d'elle décrit les positions de chacune de ses parties.

Ce format de fichier permet de réimporter ce type de données dans le ***Trameur*** et de reconstituer une base textométrique. Si on reprend la figure du « moteur textométrique » *supra*, le format d'export devient le point d'entrée du moteur textométrique qui n'agit plus comme un segmenteur (et un détecteur de parties), il est simplement un lecteur d'une ***Trame*** et d'un ***Cadre*** déjà disponibles et décrits dans le fichier donné en entrée et il en reconstruit une représentation informatique identique à celle construite dans le cas général.

8. (Re)Modelage d'une Base Textométrique

Si l'on n'y prend pas garde, le choix de certains délimiteurs peut conduire à un problème de segmentation et à la mise au jour d'items inattendus. Si on considère les unités du type :

- grand-mère
- aujourd'hui
- en raison de
- de ce fait

Les délimiteurs par défaut intègrent les caractères – et ' (cf onglet PARAM) ; les 2 premières unités seront donc segmentées respectivement en `grand + mère` et en `aujourd + hui`. Pour les 2 dernières unités, le caractère espace étant un délimiteur par défaut, elles seront segmentées en 3 morceaux.

Modifier le paramétrage de segmentation avant le chargement d'une base :

Pour prendre en compte ce genre de phénomène, il est possible de modifier la liste des délimiteurs en ayant pris soin aussi de préparer le texte avant son chargement. Si on considère les unités précédentes, on peut par exemple :

- ne pas mettre la caractère ' dans la liste des délimiteurs (« aujourd'hui » sera ainsi considéré comme une forme)
- de même pour le caractère – (« grand-mère » sera une autre forme)

Pour les séquences comme « `de ce fait` » (en gros les locutions en plusieurs mots), il faut prévoir en amont un précodage du corpus dans lequel la séquence « `de ce fait` » pourra être codé par exemple sous cette forme : `de_ce_fait` (on prendra soin de supprimer le caractère _ de la liste des délimiteurs).

Modifier une base déjà segmentée :

Si une base déjà segmentée intègre des anomalies de segmentation comme celles présentées *supra*, il est aussi possible de modifier directement la base et de corriger les éventuelles erreurs de segmentation. Après exportation de la base chargée dans Le Trameur, ce travail de modification passe par l'édition de la base exportée (fichier XML) et l'intégration des correctifs visés en prenant soin de maintenir la cohérence des informations stockées dans cette base (articulation Trame et Cadre). Si on considère l'exemple suivant (extrait d'une base textométrique) :

```
<item type="delim" pos="19343"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="forme" pos="19344"><f>de</f><c>PRP</c><l>de</l></item>
<item type="delim" pos="19345"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="forme" pos="19346"><f>ce</f><c>PRO_DEM</c><l>ce</l></item>
<item type="delim" pos="19347"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="forme" pos="19348"><f>fait</f><c>CONNECTEUR_cons</c><l>CONNECTEUR_cons</l></item>
<item type="delim" pos="19349"><f>,</f><c>DELIM</c><l>,</l></item>
<item type="delim" pos="19350"><f> </f><c>DELIM</c><l>BLANK</l></item>
```

À la position 19344, commence la locution « de ce fait » que nous voudrions considérer comme une seule forme de la Trame. Pour cela, on peut procéder comme suit :

```
<item type="delim" pos="19343"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="forme" pos="19344"><f>de_ce_fait</f><c>CONNECTEUR_cons</c><l>de_ce_fait</l></item>
<item type="delim" pos="19345"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="delim" pos="19346"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="delim" pos="19347"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="delim" pos="19348"><f> </f><c>DELIM</c><l>BLANK</l></item>
<item type="delim" pos="19349"><f>,</f><c>DELIM</c><l>,</l></item>
<item type="delim" pos="19350"><f> </f><c>DELIM</c><l>BLANK</l></item>
```

Les différentes positions de la Trame sont maintenues. Les 3 positions initiales ont été fusionnées au niveau de la position 19344 (on a concaténé les 3 formes avec un tiret bas (ce qui nécessite de supprimer ce caractère de la liste des délimiteurs). On a aussi remplacé les 2 positions initiales, 19346 et 19348, par des caractères espace (changement de type, de forme, de catégorie et de lemme sur chacune de ces positions).

En procédant ainsi et après avoir rechargé la base, la forme visée est bien définie, 2 caractères délimiteurs de type blanc ont été introduits.

Il reste possible d'aller plus loin en corrigeant cet état d'édition ayant conduit à insérer des caractères d'espacement :

- la nouvelle base importée peut de nouveau être exportée au format texte brut (cf module Export Base)
- puis être corrigée (en supprimant les items non souhaités)
- et enfin être de nouveau chargée après reformatage au format XML.

D'autres modifications sur la base exportée sont envisageables, à condition de respecter la bonne numérotation de la Trame et sa synchronisation avec la représentation du Cadre.

9. Références bibliographiques

[Lamalle, 2002], Lamalle C., Salem A., "Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels", in Actes des 6èmes Journées d'analyse statistique des données textuelles, St Malo, Inria, 2002.
http://lexicometrica.univ-paris3.fr/jadt/jadt2002/PDF-2002/lamalle_salem.pdf

[Salem, 2006], André Salem, Cédric Lamalle, Serge Fleury (EA2290 SYLED/CLA2T). "*Vers une description formelle des traitements textométriques*", in Actes JADT 2006, Journées Internationales d'Analyse Statistiques des Données Textuelles, Besançon 2006.
http://lexicometrica.univ-paris3.fr/jadt/jadt2006/PDF/II-052.pdf

[Söze-Duval, 2008], Keyser Söze-Duval. *Pour une textométrie opérationnelle*.
http://www.tal.univ-paris3.fr/trameur/RTI6provisoire.doc

10. Annexes

10.1 Base lexicométrique exportée via Le Trameur (dormeur – texte brut)


```
<?xml version="1.0" encoding="utf-8"?>
<baselexicometrique>
<teiheader>
<fileDesc>
<titleStmt>
<title>Modelisation XML de la base textometrique (le metier = le cadre et la trame),
importable et echangeable avec d'autres plateformes</title>
</titleStmt>
<publicationStmt>
<p>Lundi 11 Fevrier 2013
 17:15:23
        Ce document n'est pas encore publie.</p>
</publicationStmt>
<sourceDesc>
<title>
<br>Le Trameur 10.70.6. Lundi 11 Fevrier 2013
 17:15:23 </br></title>
<content><l><br>Fichier traite </br><br> dormeur-brut.txt</br></l>
<l><br>Encodage </br><br> iso-8859-1</br></l>
<l><br>Nombre d'items </br><br> 324</br></l>
<l><br>Nombre de délimiteurs </br><br> 197</br></l>
<l><br>Nombre d'occurrences de forme </br><br> 127</br></l>
<l><br>Nombre de formes </br><br> 96</br></l>
<l><br>Nombre d'hapax </br><br> 80</br></l>
<l><br>Fréquence maximale </br><br> 5</br></l>
<l><br>Forme maximale </br><br> dans</br></l>
<l><br>Délimiteurs </br><br><![CDATA[. ,:;!?/_-"'()[]{}§$|*><=+
«»     ]]></br></l>
<l><br>Etiquetage Treetagger </br><br>NON</br></l>
</content></sourceDesc>
</fileDesc>
</teiheader>
<trame>
<codage>utf-8</codage>
<delimiteur><![CDATA[. ,:;!?/_-"'()[]{}§$|*><=+
«»     ]]></delimiteur>
<items>
<item type="delim" pos="1"><f> </f></item>
<item type="delim" pos="2"><f> </f></item>
<item type="delim" pos="3"><f> </f></item>
<item type="delim" pos="4"><f> </f></item>
<item type="delim" pos="5"><f> </f></item>
<item type="delim" pos="6"><f> </f></item>
<item type="delim" pos="7"><f> </f></item>
<item type="delim" pos="8"><f> </f></item>
<item type="delim" pos="9"><f> </f></item>
<item type="delim" pos="10"><f> </f></item>
<item type="delim" pos="11"><f> </f></item>
<item type="delim" pos="12"><f> </f></item>
<item type="delim" pos="13"><f> </f></item>
<item type="delim" pos="14"><f> </f></item>
<item type="delim" pos="15"><f> </f></item>
<item type="delim" pos="16"><f> </f></item>
<item type="delim" pos="17"><f> </f></item>
<item type="forme" pos="18"><f>Le</f></item>
<item type="delim" pos="19"><f> </f></item>
<item type="forme" pos="20"><f>dormeur</f></item>
<item type="delim" pos="21"><f> </f></item>
<item type="forme" pos="22"><f>du</f></item>
<item type="delim" pos="23"><f> </f></item>
<item type="forme" pos="24"><f>val</f></item>
<item type="delim" pos="25"><f> </f></item>
<item type="delim" pos="26"><f>RETURN</f></item>
<item type="delim" pos="27"><f>RETURN</f></item>
<item type="forme" pos="28"><f>C</f></item>
<item type="delim" pos="29"><f>'</f></item>
<item type="forme" pos="30"><f>est</f></item>
```

```
<item type="delim" pos="31"><f> </f></item>
<item type="forme" pos="32"><f>un</f></item>
<item type="delim" pos="33"><f> </f></item>
<item type="forme" pos="34"><f>trou</f></item>
<item type="delim" pos="35"><f> </f></item>
<item type="forme" pos="36"><f>de</f></item>
<item type="delim" pos="37"><f> </f></item>
<item type="forme" pos="38"><f>verdure</f></item>
<item type="delim" pos="39"><f> </f></item>
<item type="forme" pos="40"><f>où</f></item>
<item type="delim" pos="41"><f> </f></item>
<item type="forme" pos="42"><f>chante</f></item>
<item type="delim" pos="43"><f> </f></item>
<item type="forme" pos="44"><f>une</f></item>
<item type="delim" pos="45"><f> </f></item>
<item type="forme" pos="46"><f>rivière</f></item>
<item type="delim" pos="47"><f>,</f></item>
<item type="delim" pos="48"><f>RETURN</f></item>
<item type="forme" pos="49"><f>Accrochant</f></item>
<item type="delim" pos="50"><f> </f></item>
<item type="forme" pos="51"><f>follement</f></item>
<item type="delim" pos="52"><f> </f></item>
<item type="forme" pos="53"><f>aux</f></item>
<item type="delim" pos="54"><f> </f></item>
<item type="forme" pos="55"><f>herbes</f></item>
<item type="delim" pos="56"><f> </f></item>
<item type="forme" pos="57"><f>des</f></item>
<item type="delim" pos="58"><f> </f></item>
<item type="forme" pos="59"><f>haillons</f></item>
<item type="delim" pos="60"><f>RETURN</f></item>
<item type="forme" pos="61"><f>D</f></item>
<item type="delim" pos="62"><f>'</f></item>
<item type="forme" pos="63"><f>argent</f></item>
<item type="delim" pos="64"><f> </f></item>
<item type="delim" pos="65"><f>;</f></item>
<item type="delim" pos="66"><f> </f></item>
<item type="forme" pos="67"><f>où</f></item>
<item type="delim" pos="68"><f> </f></item>
<item type="forme" pos="69"><f>le</f></item>
<item type="delim" pos="70"><f> </f></item>
<item type="forme" pos="71"><f>soleil</f></item>
<item type="delim" pos="72"><f>,</f></item>
<item type="delim" pos="73"><f> </f></item>
<item type="forme" pos="74"><f>de</f></item>
<item type="delim" pos="75"><f> </f></item>
<item type="forme" pos="76"><f>la</f></item>
<item type="delim" pos="77"><f> </f></item>
<item type="forme" pos="78"><f>montagne</f></item>
<item type="delim" pos="79"><f> </f></item>
<item type="forme" pos="80"><f>fière</f></item>
<item type="delim" pos="81"><f>,</f></item>
<item type="delim" pos="82"><f>RETURN</f></item>
<item type="forme" pos="83"><f>Luit</f></item>
<item type="delim" pos="84"><f> </f></item>
<item type="delim" pos="85"><f>:</f></item>
<item type="delim" pos="86"><f> </f></item>
<item type="forme" pos="87"><f>c</f></item>
<item type="delim" pos="88"><f>'</f></item>
<item type="forme" pos="89"><f>est</f></item>
<item type="delim" pos="90"><f> </f></item>
<item type="forme" pos="91"><f>un</f></item>
<item type="delim" pos="92"><f> </f></item>
<item type="forme" pos="93"><f>petit</f></item>
<item type="delim" pos="94"><f> </f></item>
<item type="forme" pos="95"><f>val</f></item>
<item type="delim" pos="96"><f> </f></item>
<item type="forme" pos="97"><f>qui</f></item>
<item type="delim" pos="98"><f> </f></item>
<item type="forme" pos="99"><f>mousse</f></item>
<item type="delim" pos="100"><f> </f></item>
<item type="forme" pos="101"><f>de</f></item>
<item type="delim" pos="102"><f> </f></item>
<item type="forme" pos="103"><f>rayons</f></item>
<item type="delim" pos="104"><f>.</f></item>
<item type="delim" pos="105"><f> </f></item>
<item type="delim" pos="106"><f>RETURN</f></item>
```

```
<item type="delim" pos="107"><f>RETURN</f></item>
<item type="forme" pos="108"><f>Un</f></item>
<item type="delim" pos="109"><f> </f></item>
<item type="forme" pos="110"><f>soldat</f></item>
<item type="delim" pos="111"><f> </f></item>
<item type="forme" pos="112"><f>jeune</f></item>
<item type="delim" pos="113"><f>,</f></item>
<item type="delim" pos="114"><f> </f></item>
<item type="forme" pos="115"><f>bouche</f></item>
<item type="delim" pos="116"><f> </f></item>
<item type="forme" pos="117"><f>ouverte</f></item>
<item type="delim" pos="118"><f>,</f></item>
<item type="delim" pos="119"><f> </f></item>
<item type="forme" pos="120"><f>tête</f></item>
<item type="delim" pos="121"><f> </f></item>
<item type="forme" pos="122"><f>nue</f></item>
<item type="delim" pos="123"><f>,</f></item>
<item type="delim" pos="124"><f>RETURN</f></item>
<item type="forme" pos="125"><f>Et</f></item>
<item type="delim" pos="126"><f> </f></item>
<item type="forme" pos="127"><f>la</f></item>
<item type="delim" pos="128"><f> </f></item>
<item type="forme" pos="129"><f>nuque</f></item>
<item type="delim" pos="130"><f> </f></item>
<item type="forme" pos="131"><f>baignant</f></item>
<item type="delim" pos="132"><f> </f></item>
<item type="forme" pos="133"><f>dans</f></item>
<item type="delim" pos="134"><f> </f></item>
<item type="forme" pos="135"><f>le</f></item>
<item type="delim" pos="136"><f> </f></item>
<item type="forme" pos="137"><f>frais</f></item>
<item type="delim" pos="138"><f> </f></item>
<item type="forme" pos="139"><f>cresson</f></item>
<item type="delim" pos="140"><f> </f></item>
<item type="forme" pos="141"><f>bleu</f></item>
<item type="delim" pos="142"><f>,</f></item>
<item type="delim" pos="143"><f>RETURN</f></item>
<item type="forme" pos="144"><f>Dort</f></item>
<item type="delim" pos="145"><f> </f></item>
<item type="delim" pos="146"><f>;</f></item>
<item type="delim" pos="147"><f> </f></item>
<item type="forme" pos="148"><f>il</f></item>
<item type="delim" pos="149"><f> </f></item>
<item type="forme" pos="150"><f>est</f></item>
<item type="delim" pos="151"><f> </f></item>
<item type="forme" pos="152"><f>étendu</f></item>
<item type="delim" pos="153"><f> </f></item>
<item type="forme" pos="154"><f>dans</f></item>
<item type="delim" pos="155"><f> </f></item>
<item type="forme" pos="156"><f>l</f></item>
<item type="delim" pos="157"><f>'</f></item>
<item type="forme" pos="158"><f>herbe</f></item>
<item type="delim" pos="159"><f>,</f></item>
<item type="delim" pos="160"><f> </f></item>
<item type="forme" pos="161"><f>sous</f></item>
<item type="delim" pos="162"><f> </f></item>
<item type="forme" pos="163"><f>la</f></item>
<item type="delim" pos="164"><f> </f></item>
<item type="forme" pos="165"><f>nue</f></item>
<item type="delim" pos="166"><f>,</f></item>
<item type="delim" pos="167"><f>RETURN</f></item>
<item type="forme" pos="168"><f>Pâle</f></item>
<item type="delim" pos="169"><f> </f></item>
<item type="forme" pos="170"><f>dans</f></item>
<item type="delim" pos="171"><f> </f></item>
<item type="forme" pos="172"><f>son</f></item>
<item type="delim" pos="173"><f> </f></item>
<item type="forme" pos="174"><f>lit</f></item>
<item type="delim" pos="175"><f> </f></item>
<item type="forme" pos="176"><f>vert</f></item>
<item type="delim" pos="177"><f> </f></item>
<item type="forme" pos="178"><f>où</f></item>
<item type="delim" pos="179"><f> </f></item>
<item type="forme" pos="180"><f>la</f></item>
<item type="delim" pos="181"><f> </f></item>
<item type="forme" pos="182"><f>lumière</f></item>
```

```
<item type="delim" pos="183"><f> </f></item>
<item type="forme" pos="184"><f>pleut</f></item>
<item type="delim" pos="185"><f>.</f></item>
<item type="delim" pos="186"><f> </f></item>
<item type="delim" pos="187"><f>RETURN</f></item>
<item type="delim" pos="188"><f>RETURN</f></item>
<item type="forme" pos="189"><f>Les</f></item>
<item type="delim" pos="190"><f> </f></item>
<item type="forme" pos="191"><f>pieds</f></item>
<item type="delim" pos="192"><f> </f></item>
<item type="forme" pos="193"><f>dans</f></item>
<item type="delim" pos="194"><f> </f></item>
<item type="forme" pos="195"><f>les</f></item>
<item type="delim" pos="196"><f> </f></item>
<item type="forme" pos="197"><f>glaïeuls</f></item>
<item type="delim" pos="198"><f>,</f></item>
<item type="delim" pos="199"><f> </f></item>
<item type="forme" pos="200"><f>il</f></item>
<item type="delim" pos="201"><f> </f></item>
<item type="forme" pos="202"><f>dort</f></item>
<item type="delim" pos="203"><f>.</f></item>
<item type="delim" pos="204"><f> </f></item>
<item type="forme" pos="205"><f>Souriant</f></item>
<item type="delim" pos="206"><f> </f></item>
<item type="forme" pos="207"><f>comme</f></item>
<item type="delim" pos="208"><f>RETURN</f></item>
<item type="forme" pos="209"><f>Sourirait</f></item>
<item type="delim" pos="210"><f> </f></item>
<item type="forme" pos="211"><f>un</f></item>
<item type="delim" pos="212"><f> </f></item>
<item type="forme" pos="213"><f>enfant</f></item>
<item type="delim" pos="214"><f> </f></item>
<item type="forme" pos="215"><f>malade</f></item>
<item type="delim" pos="216"><f>,</f></item>
<item type="delim" pos="217"><f> </f></item>
<item type="forme" pos="218"><f>il</f></item>
<item type="delim" pos="219"><f> </f></item>
<item type="forme" pos="220"><f>fait</f></item>
<item type="delim" pos="221"><f> </f></item>
<item type="forme" pos="222"><f>un</f></item>
<item type="delim" pos="223"><f> </f></item>
<item type="forme" pos="224"><f>somme</f></item>
<item type="delim" pos="225"><f> </f></item>
<item type="delim" pos="226"><f>:</f></item>
<item type="delim" pos="227"><f>RETURN</f></item>
<item type="forme" pos="228"><f>Nature</f></item>
<item type="delim" pos="229"><f>,</f></item>
<item type="delim" pos="230"><f> </f></item>
<item type="forme" pos="231"><f>berce</f></item>
<item type="delim" pos="232"><f>-</f></item>
<item type="forme" pos="233"><f>le</f></item>
<item type="delim" pos="234"><f> </f></item>
<item type="forme" pos="235"><f>chaudement</f></item>
<item type="delim" pos="236"><f> </f></item>
<item type="delim" pos="237"><f>:</f></item>
<item type="delim" pos="238"><f> </f></item>
<item type="forme" pos="239"><f>il</f></item>
<item type="delim" pos="240"><f> </f></item>
<item type="forme" pos="241"><f>a</f></item>
<item type="delim" pos="242"><f> </f></item>
<item type="forme" pos="243"><f>froid</f></item>
<item type="delim" pos="244"><f>.</f></item>
<item type="delim" pos="245"><f> </f></item>
<item type="delim" pos="246"><f>RETURN</f></item>
<item type="delim" pos="247"><f>RETURN</f></item>
<item type="forme" pos="248"><f>Les</f></item>
<item type="delim" pos="249"><f> </f></item>
<item type="forme" pos="250"><f>parfums</f></item>
<item type="delim" pos="251"><f> </f></item>
<item type="forme" pos="252"><f>ne</f></item>
<item type="delim" pos="253"><f> </f></item>
<item type="forme" pos="254"><f>font</f></item>
<item type="delim" pos="255"><f> </f></item>
<item type="forme" pos="256"><f>pas</f></item>
<item type="delim" pos="257"><f> </f></item>
<item type="forme" pos="258"><f>frissonner</f></item>
```

```
<item type="delim" pos="259"><f> </f></item>
<item type="forme" pos="260"><f>sa</f></item>
<item type="delim" pos="261"><f> </f></item>
<item type="forme" pos="262"><f>narine</f></item>
<item type="delim" pos="263"><f> </f></item>
<item type="delim" pos="264"><f>;</f></item>
<item type="delim" pos="265"><f>RETURN</f></item>
<item type="forme" pos="266"><f>Il</f></item>
<item type="delim" pos="267"><f> </f></item>
<item type="forme" pos="268"><f>dort</f></item>
<item type="delim" pos="269"><f> </f></item>
<item type="forme" pos="270"><f>dans</f></item>
<item type="delim" pos="271"><f> </f></item>
<item type="forme" pos="272"><f>le</f></item>
<item type="delim" pos="273"><f> </f></item>
<item type="forme" pos="274"><f>soleil</f></item>
<item type="delim" pos="275"><f>,</f></item>
<item type="delim" pos="276"><f> </f></item>
<item type="forme" pos="277"><f>la</f></item>
<item type="delim" pos="278"><f> </f></item>
<item type="forme" pos="279"><f>main</f></item>
<item type="delim" pos="280"><f> </f></item>
<item type="forme" pos="281"><f>sur</f></item>
<item type="delim" pos="282"><f> </f></item>
<item type="forme" pos="283"><f>sa</f></item>
<item type="delim" pos="284"><f> </f></item>
<item type="forme" pos="285"><f>poitrine</f></item>
<item type="delim" pos="286"><f>,</f></item>
<item type="delim" pos="287"><f>RETURN</f></item>
<item type="forme" pos="288"><f>Tranquille</f></item>
<item type="delim" pos="289"><f>.</f></item>
<item type="delim" pos="290"><f> </f></item>
<item type="forme" pos="291"><f>Il</f></item>
<item type="delim" pos="292"><f> </f></item>
<item type="forme" pos="293"><f>a</f></item>
<item type="delim" pos="294"><f> </f></item>
<item type="forme" pos="295"><f>deux</f></item>
<item type="delim" pos="296"><f> </f></item>
<item type="forme" pos="297"><f>trous</f></item>
<item type="delim" pos="298"><f> </f></item>
<item type="forme" pos="299"><f>rouges</f></item>
<item type="delim" pos="300"><f> </f></item>
<item type="forme" pos="301"><f>au</f></item>
<item type="delim" pos="302"><f> </f></item>
<item type="forme" pos="303"><f>côté</f></item>
<item type="delim" pos="304"><f> </f></item>
<item type="forme" pos="305"><f>droit</f></item>
<item type="delim" pos="306"><f>.</f></item>
<item type="delim" pos="307"><f> </f></item>
<item type="delim" pos="308"><f>RETURN</f></item>
<item type="delim" pos="309"><f>RETURN</f></item>
<item type="delim" pos="310"><f> </f></item>
<item type="delim" pos="311"><f> </f></item>
<item type="delim" pos="312"><f> </f></item>
<item type="delim" pos="313"><f> </f></item>
<item type="delim" pos="314"><f> </f></item>
<item type="delim" pos="315"><f> </f></item>
<item type="delim" pos="316"><f> </f></item>
<item type="delim" pos="317"><f> </f></item>
<item type="delim" pos="318"><f> </f></item>
<item type="delim" pos="319"><f> </f></item>
<item type="delim" pos="320"><f> </f></item>
<item type="forme" pos="321"><f>Arthur</f></item>
<item type="delim" pos="322"><f> </f></item>
<item type="forme" pos="323"><f>Rimbaud</f></item>
<item type="delim" pos="324"><f> </f></item>
<item type="delim" pos="325"><f> </f></item>
</items>
</trame>
<cadre>
<acces>
</acces>
</cadre>
</baselexicometrique>
```

10.2 Base lexicométrique exportée via Le Trameur (dormeur – TEI)


```
<?xml version="1.0" encoding="utf-8"?>
<baselexicometrique>
<teiheader>
<fileDesc>
<titleStmt>
<title>Modelisation XML de la base textometrique (le metier = le cadre et la trame),
importable et echangeable avec d'autres plateformes</title>
</titleStmt>
<publicationStmt>
<p>Lundi 11 Fevrier 2013
 17:18:49
        Ce document n'est pas encore publie.</p>
</publicationStmt>
<sourceDesc>
<title>
<br>Le Trameur 10.70.6. Lundi 11 Fevrier 2013
 17:18:49 </br></title>
<content><l><br>Fichier traite </br><br> dormeur-tei.xml</br></l>
<l><br>Encodage </br><br> iso-8859-1</br></l>
<l><br>Nombre d'items </br><br> 534</br></l>
<l><br>Nombre de délimiteurs </br><br> 403</br></l>
<l><br>Nombre d'occurrences de forme </br><br> 131</br></l>
<l><br>Nombre de formes </br><br> 101</br></l>
<l><br>Nombre d'hapax </br><br> 86</br></l>
<l><br>Fréquence maximale </br><br> 5</br></l>
<l><br>Forme maximale </br><br> dans</br></l>
<l><br>Délimiteurs </br><br><![CDATA[. ,:;!?/_-"'()[]{}§$|*><=+
«»     ]]></br></l>
<l><br>Etiquetage Treetagger </br><br>NON</br></l>
</content></sourceDesc>
</fileDesc>
</teiheader>
<trame>
<codage>utf-8</codage>
<delimiteur><![CDATA[. ,:;!?/_-"'()[]{}§$|*><=+
«»     ]]></delimiteur>
<items>
<item type="delim" pos="1"><f>RETURN</f></item>
<item type="delim" pos="2"><f>       </f></item>
<item type="delim" pos="3"><f>RETURN</f></item>
<item type="delim" pos="4"><f>       </f></item>
<item type="delim" pos="5"><f>       </f></item>
<item type="delim" pos="6"><f>RETURN</f></item>
<item type="delim" pos="7"><f>       </f></item>
<item type="delim" pos="8"><f>       </f></item>
<item type="delim" pos="9"><f>       </f></item>
<item type="delim" pos="10"><f>RETURN</f></item>
<item type="delim" pos="11"><f>      </f></item>
<item type="delim" pos="12"><f>      </f></item>
<item type="delim" pos="13"><f>      </f></item>
<item type="delim" pos="14"><f>      </f></item>
<item type="forme" pos="15"><f>Poésies</f></item>
<item type="delim" pos="16"><f>RETURN</f></item>
<item type="delim" pos="17"><f>      </f></item>
<item type="delim" pos="18"><f>      </f></item>
<item type="delim" pos="19"><f>      </f></item>
<item type="delim" pos="20"><f>      </f></item>
<item type="forme" pos="21"><f>Arthur</f></item>
<item type="delim" pos="22"><f> </f></item>
<item type="forme" pos="23"><f>Rimbaud</f></item>
<item type="delim" pos="24"><f>RETURN</f></item>
<item type="delim" pos="25"><f>      </f></item>
<item type="delim" pos="26"><f>      </f></item>
<item type="delim" pos="27"><f>      </f></item>
<item type="delim" pos="28"><f>RETURN</f></item>
<item type="delim" pos="29"><f>      </f></item>
<item type="delim" pos="30"><f>      </f></item>
<item type="delim" pos="31"><f>RETURN</f></item>
<item type="delim" pos="32"><f>      </f></item>
<item type="delim" pos="33"><f>RETURN</f></item>
<item type="delim" pos="34"><f>      </f></item>
<item type="delim" pos="35"><f>RETURN</f></item>
```

```
<item type="delim" pos="36"><f>      </f></item>
<item type="delim" pos="37"><f>      </f></item>
<item type="delim" pos="38"><f>RETURN</f></item>
<item type="delim" pos="39"><f>      </f></item>
<item type="delim" pos="40"><f>      </f></item>
<item type="delim" pos="41"><f>      </f></item>
<item type="delim" pos="42"><f>RETURN</f></item>
<item type="delim" pos="43"><f>      </f></item>
<item type="delim" pos="44"><f>      </f></item>
<item type="delim" pos="45"><f>      </f></item>
<item type="delim" pos="46"><f>      </f></item>
<item type="delim" pos="47"><f>RETURN</f></item>
<item type="delim" pos="48"><f>      </f></item>
<item type="delim" pos="49"><f>      </f></item>
<item type="delim" pos="50"><f>      </f></item>
<item type="delim" pos="51"><f>      </f></item>
<item type="delim" pos="52"><f>      </f></item>
<item type="forme" pos="53"><f>Le</f></item>
<item type="delim" pos="54"><f> </f></item>
<item type="forme" pos="55"><f>Dormeur</f></item>
<item type="delim" pos="56"><f> </f></item>
<item type="forme" pos="57"><f>du</f></item>
<item type="delim" pos="58"><f> </f></item>
<item type="forme" pos="59"><f>Val</f></item>
<item type="delim" pos="60"><f>RETURN</f></item>
<item type="delim" pos="61"><f>      </f></item>
<item type="delim" pos="62"><f>      </f></item>
<item type="delim" pos="63"><f>      </f></item>
<item type="delim" pos="64"><f>      </f></item>
<item type="delim" pos="65"><f>      </f></item>
<item type="forme" pos="66"><f>7</f></item>
<item type="delim" pos="67"><f> </f></item>
<item type="forme" pos="68"><f>octobre</f></item>
<item type="delim" pos="69"><f> </f></item>
<item type="forme" pos="70"><f>1870</f></item>
<item type="delim" pos="71"><f>RETURN</f></item>
<item type="delim" pos="72"><f>      </f></item>
<item type="delim" pos="73"><f>      </f></item>
<item type="delim" pos="74"><f>      </f></item>
<item type="delim" pos="75"><f>      </f></item>
<item type="delim" pos="76"><f>RETURN</f></item>
<item type="delim" pos="77"><f>      </f></item>
<item type="delim" pos="78"><f>      </f></item>
<item type="delim" pos="79"><f>      </f></item>
<item type="delim" pos="80"><f>      </f></item>
<item type="delim" pos="81"><f>RETURN</f></item>
<item type="delim" pos="82"><f>      </f></item>
<item type="delim" pos="83"><f>      </f></item>
<item type="delim" pos="84"><f>      </f></item>
<item type="delim" pos="85"><f>      </f></item>
<item type="delim" pos="86"><f>      </f></item>
<item type="delim" pos="87"><f>RETURN</f></item>
<item type="delim" pos="88"><f>      </f></item>
<item type="delim" pos="89"><f>      </f></item>
<item type="delim" pos="90"><f>      </f></item>
<item type="delim" pos="91"><f>      </f></item>
<item type="delim" pos="92"><f>      </f></item>
<item type="delim" pos="93"><f>      </f></item>
<item type="delim" pos="94"><f>RETURN</f></item>
<item type="delim" pos="95"><f>      </f></item>
<item type="delim" pos="96"><f>      </f></item>
<item type="delim" pos="97"><f>      </f></item>
<item type="delim" pos="98"><f>      </f></item>
<item type="delim" pos="99"><f>      </f></item>
<item type="delim" pos="100"><f>     </f></item>
<item type="delim" pos="101"><f>     </f></item>
<item type="forme" pos="102"><f>C</f></item>
<item type="delim" pos="103"><f>'</f></item>
<item type="forme" pos="104"><f>est</f></item>
<item type="delim" pos="105"><f> </f></item>
<item type="forme" pos="106"><f>un</f></item>
<item type="delim" pos="107"><f> </f></item>
<item type="forme" pos="108"><f>trou</f></item>
<item type="delim" pos="109"><f> </f></item>
<item type="forme" pos="110"><f>de</f></item>
<item type="delim" pos="111"><f> </f></item>
```

```
<item type="forme" pos="112"><f>verdure</f></item>
<item type="delim" pos="113"><f> </f></item>
<item type="forme" pos="114"><f>où</f></item>
<item type="delim" pos="115"><f> </f></item>
<item type="forme" pos="116"><f>chante</f></item>
<item type="delim" pos="117"><f> </f></item>
<item type="forme" pos="118"><f>une</f></item>
<item type="delim" pos="119"><f> </f></item>
<item type="forme" pos="120"><f>rivière</f></item>
<item type="delim" pos="121"><f>RETURN</f></item>
<item type="delim" pos="122"><f>     </f></item>
<item type="delim" pos="123"><f>     </f></item>
<item type="delim" pos="124"><f>     </f></item>
<item type="delim" pos="125"><f>     </f></item>
<item type="delim" pos="126"><f>     </f></item>
<item type="delim" pos="127"><f>     </f></item>
<item type="delim" pos="128"><f>     </f></item>
<item type="forme" pos="129"><f>Accrochant</f></item>
<item type="delim" pos="130"><f> </f></item>
<item type="forme" pos="131"><f>follement</f></item>
<item type="delim" pos="132"><f> </f></item>
<item type="forme" pos="133"><f>aux</f></item>
<item type="delim" pos="134"><f> </f></item>
<item type="forme" pos="135"><f>herbes</f></item>
<item type="delim" pos="136"><f> </f></item>
<item type="forme" pos="137"><f>des</f></item>
<item type="delim" pos="138"><f> </f></item>
<item type="forme" pos="139"><f>haillons</f></item>
<item type="delim" pos="140"><f>RETURN</f></item>
<item type="delim" pos="141"><f>     </f></item>
<item type="delim" pos="142"><f>     </f></item>
<item type="delim" pos="143"><f>     </f></item>
<item type="delim" pos="144"><f>     </f></item>
<item type="delim" pos="145"><f>     </f></item>
<item type="delim" pos="146"><f>     </f></item>
<item type="delim" pos="147"><f>     </f></item>
<item type="forme" pos="148"><f>D</f></item>
<item type="delim" pos="149"><f>'</f></item>
<item type="forme" pos="150"><f>argent</f></item>
<item type="delim" pos="151"><f> </f></item>
<item type="delim" pos="152"><f>;</f></item>
<item type="delim" pos="153"><f> </f></item>
<item type="forme" pos="154"><f>où</f></item>
<item type="delim" pos="155"><f> </f></item>
<item type="forme" pos="156"><f>le</f></item>
<item type="delim" pos="157"><f> </f></item>
<item type="forme" pos="158"><f>soleil</f></item>
<item type="delim" pos="159"><f>,</f></item>
<item type="delim" pos="160"><f> </f></item>
<item type="forme" pos="161"><f>de</f></item>
<item type="delim" pos="162"><f> </f></item>
<item type="forme" pos="163"><f>la</f></item>
<item type="delim" pos="164"><f> </f></item>
<item type="forme" pos="165"><f>montagne</f></item>
<item type="delim" pos="166"><f> </f></item>
<item type="forme" pos="167"><f>fière</f></item>
<item type="delim" pos="168"><f>,</f></item>
<item type="delim" pos="169"><f>RETURN</f></item>
<item type="delim" pos="170"><f>     </f></item>
<item type="delim" pos="171"><f>     </f></item>
<item type="delim" pos="172"><f>     </f></item>
<item type="delim" pos="173"><f>     </f></item>
<item type="delim" pos="174"><f>     </f></item>
<item type="delim" pos="175"><f>     </f></item>
<item type="delim" pos="176"><f>     </f></item>
<item type="forme" pos="177"><f>Luit</f></item>
<item type="delim" pos="178"><f> </f></item>
<item type="delim" pos="179"><f>:</f></item>
<item type="delim" pos="180"><f> </f></item>
<item type="forme" pos="181"><f>c</f></item>
<item type="delim" pos="182"><f>'</f></item>
<item type="forme" pos="183"><f>est</f></item>
<item type="delim" pos="184"><f> </f></item>
<item type="forme" pos="185"><f>un</f></item>
<item type="delim" pos="186"><f> </f></item>
<item type="forme" pos="187"><f>petit</f></item>
```

```
<item type="delim" pos="188"><f> </f></item>
<item type="forme" pos="189"><f>val</f></item>
<item type="delim" pos="190"><f> </f></item>
<item type="forme" pos="191"><f>qui</f></item>
<item type="delim" pos="192"><f> </f></item>
<item type="forme" pos="193"><f>mousse</f></item>
<item type="delim" pos="194"><f> </f></item>
<item type="forme" pos="195"><f>de</f></item>
<item type="delim" pos="196"><f> </f></item>
<item type="forme" pos="197"><f>rayons</f></item>
<item type="delim" pos="198"><f>.</f></item>
<item type="delim" pos="199"><f>RETURN</f></item>
<item type="delim" pos="200"><f>     </f></item>
<item type="delim" pos="201"><f>     </f></item>
<item type="delim" pos="202"><f>     </f></item>
<item type="delim" pos="203"><f>     </f></item>
<item type="delim" pos="204"><f>     </f></item>
<item type="delim" pos="205"><f>     </f></item>
<item type="delim" pos="206"><f>RETURN</f></item>
<item type="delim" pos="207"><f>     </f></item>
<item type="delim" pos="208"><f>     </f></item>
<item type="delim" pos="209"><f>     </f></item>
<item type="delim" pos="210"><f>     </f></item>
<item type="delim" pos="211"><f>     </f></item>
<item type="delim" pos="212"><f>     </f></item>
<item type="delim" pos="213"><f>RETURN</f></item>
<item type="delim" pos="214"><f>     </f></item>
<item type="delim" pos="215"><f>     </f></item>
<item type="delim" pos="216"><f>     </f></item>
<item type="delim" pos="217"><f>     </f></item>
<item type="delim" pos="218"><f>     </f></item>
<item type="delim" pos="219"><f>     </f></item>
<item type="delim" pos="220"><f>     </f></item>
<item type="forme" pos="221"><f>Un</f></item>
<item type="delim" pos="222"><f> </f></item>
<item type="forme" pos="223"><f>soldat</f></item>
<item type="delim" pos="224"><f> </f></item>
<item type="forme" pos="225"><f>jeune</f></item>
<item type="delim" pos="226"><f>,</f></item>
<item type="delim" pos="227"><f> </f></item>
<item type="forme" pos="228"><f>bouche</f></item>
<item type="delim" pos="229"><f> </f></item>
<item type="forme" pos="230"><f>ouverte</f></item>
<item type="delim" pos="231"><f>,</f></item>
<item type="delim" pos="232"><f> </f></item>
<item type="forme" pos="233"><f>tête</f></item>
<item type="delim" pos="234"><f> </f></item>
<item type="forme" pos="235"><f>nue</f></item>
<item type="delim" pos="236"><f>,</f></item>
<item type="delim" pos="237"><f>RETURN</f></item>
<item type="delim" pos="238"><f>     </f></item>
<item type="delim" pos="239"><f>     </f></item>
<item type="delim" pos="240"><f>     </f></item>
<item type="delim" pos="241"><f>     </f></item>
<item type="delim" pos="242"><f>     </f></item>
<item type="delim" pos="243"><f>     </f></item>
<item type="delim" pos="244"><f>     </f></item>
<item type="forme" pos="245"><f>Et</f></item>
<item type="delim" pos="246"><f> </f></item>
<item type="forme" pos="247"><f>la</f></item>
<item type="delim" pos="248"><f> </f></item>
<item type="forme" pos="249"><f>nuque</f></item>
<item type="delim" pos="250"><f> </f></item>
<item type="forme" pos="251"><f>baignant</f></item>
<item type="delim" pos="252"><f> </f></item>
<item type="forme" pos="253"><f>dans</f></item>
<item type="delim" pos="254"><f> </f></item>
<item type="forme" pos="255"><f>le</f></item>
<item type="delim" pos="256"><f> </f></item>
<item type="forme" pos="257"><f>frais</f></item>
<item type="delim" pos="258"><f> </f></item>
<item type="forme" pos="259"><f>cresson</f></item>
<item type="delim" pos="260"><f> </f></item>
<item type="forme" pos="261"><f>bleu</f></item>
<item type="delim" pos="262"><f>,</f></item>
<item type="delim" pos="263"><f>RETURN</f></item>
```

```
<item type="delim" pos="264"><f>     </f></item>
<item type="delim" pos="265"><f>     </f></item>
<item type="delim" pos="266"><f>     </f></item>
<item type="delim" pos="267"><f>     </f></item>
<item type="delim" pos="268"><f>     </f></item>
<item type="delim" pos="269"><f>     </f></item>
<item type="delim" pos="270"><f>     </f></item>
<item type="forme" pos="271"><f>Dort</f></item>
<item type="delim" pos="272"><f> </f></item>
<item type="delim" pos="273"><f>;</f></item>
<item type="delim" pos="274"><f> </f></item>
<item type="forme" pos="275"><f>il</f></item>
<item type="delim" pos="276"><f> </f></item>
<item type="forme" pos="277"><f>est</f></item>
<item type="delim" pos="278"><f> </f></item>
<item type="forme" pos="279"><f>étendu</f></item>
<item type="delim" pos="280"><f> </f></item>
<item type="forme" pos="281"><f>dans</f></item>
<item type="delim" pos="282"><f> </f></item>
<item type="forme" pos="283"><f>l</f></item>
<item type="delim" pos="284"><f>'</f></item>
<item type="forme" pos="285"><f>herbe</f></item>
<item type="delim" pos="286"><f>,</f></item>
<item type="delim" pos="287"><f> </f></item>
<item type="forme" pos="288"><f>sous</f></item>
<item type="delim" pos="289"><f> </f></item>
<item type="forme" pos="290"><f>la</f></item>
<item type="delim" pos="291"><f> </f></item>
<item type="forme" pos="292"><f>nue</f></item>
<item type="delim" pos="293"><f>,</f></item>
<item type="delim" pos="294"><f>RETURN</f></item>
<item type="delim" pos="295"><f>     </f></item>
<item type="delim" pos="296"><f>     </f></item>
<item type="delim" pos="297"><f>     </f></item>
<item type="delim" pos="298"><f>     </f></item>
<item type="delim" pos="299"><f>     </f></item>
<item type="delim" pos="300"><f>     </f></item>
<item type="delim" pos="301"><f>     </f></item>
<item type="forme" pos="302"><f>Pâle</f></item>
<item type="delim" pos="303"><f> </f></item>
<item type="forme" pos="304"><f>dans</f></item>
<item type="delim" pos="305"><f> </f></item>
<item type="forme" pos="306"><f>son</f></item>
<item type="delim" pos="307"><f> </f></item>
<item type="forme" pos="308"><f>lit</f></item>
<item type="delim" pos="309"><f> </f></item>
<item type="forme" pos="310"><f>vert</f></item>
<item type="delim" pos="311"><f> </f></item>
<item type="forme" pos="312"><f>où</f></item>
<item type="delim" pos="313"><f> </f></item>
<item type="forme" pos="314"><f>la</f></item>
<item type="delim" pos="315"><f> </f></item>
<item type="forme" pos="316"><f>lumière</f></item>
<item type="delim" pos="317"><f> </f></item>
<item type="forme" pos="318"><f>pleut</f></item>
<item type="delim" pos="319"><f>.</f></item>
<item type="delim" pos="320"><f>RETURN</f></item>
<item type="delim" pos="321"><f>     </f></item>
<item type="delim" pos="322"><f>     </f></item>
<item type="delim" pos="323"><f>     </f></item>
<item type="delim" pos="324"><f>     </f></item>
<item type="delim" pos="325"><f>     </f></item>
<item type="delim" pos="326"><f>     </f></item>
<item type="delim" pos="327"><f>RETURN</f></item>
<item type="delim" pos="328"><f>     </f></item>
<item type="delim" pos="329"><f>     </f></item>
<item type="delim" pos="330"><f>     </f></item>
<item type="delim" pos="331"><f>     </f></item>
<item type="delim" pos="332"><f>     </f></item>
<item type="delim" pos="333"><f>     </f></item>
<item type="delim" pos="334"><f>RETURN</f></item>
<item type="delim" pos="335"><f>     </f></item>
<item type="delim" pos="336"><f>     </f></item>
<item type="delim" pos="337"><f>     </f></item>
<item type="delim" pos="338"><f>     </f></item>
<item type="delim" pos="339"><f>     </f></item>
```

```
<item type="delim" pos="340"><f>     </f></item>
<item type="delim" pos="341"><f>     </f></item>
<item type="forme" pos="342"><f>Les</f></item>
<item type="delim" pos="343"><f> </f></item>
<item type="forme" pos="344"><f>pieds</f></item>
<item type="delim" pos="345"><f> </f></item>
<item type="forme" pos="346"><f>dans</f></item>
<item type="delim" pos="347"><f> </f></item>
<item type="forme" pos="348"><f>les</f></item>
<item type="delim" pos="349"><f> </f></item>
<item type="forme" pos="350"><f>glaïeuls</f></item>
<item type="delim" pos="351"><f>,</f></item>
<item type="delim" pos="352"><f> </f></item>
<item type="forme" pos="353"><f>il</f></item>
<item type="delim" pos="354"><f> </f></item>
<item type="forme" pos="355"><f>dort</f></item>
<item type="delim" pos="356"><f>.</f></item>
<item type="delim" pos="357"><f> </f></item>
<item type="forme" pos="358"><f>Souriant</f></item>
<item type="delim" pos="359"><f> </f></item>
<item type="forme" pos="360"><f>comme</f></item>
<item type="delim" pos="361"><f>RETURN</f></item>
<item type="delim" pos="362"><f>     </f></item>
<item type="delim" pos="363"><f>     </f></item>
<item type="delim" pos="364"><f>     </f></item>
<item type="delim" pos="365"><f>     </f></item>
<item type="delim" pos="366"><f>     </f></item>
<item type="delim" pos="367"><f>     </f></item>
<item type="delim" pos="368"><f>     </f></item>
<item type="forme" pos="369"><f>Sourirait</f></item>
<item type="delim" pos="370"><f> </f></item>
<item type="forme" pos="371"><f>un</f></item>
<item type="delim" pos="372"><f> </f></item>
<item type="forme" pos="373"><f>enfant</f></item>
<item type="delim" pos="374"><f> </f></item>
<item type="forme" pos="375"><f>malade</f></item>
<item type="delim" pos="376"><f>,</f></item>
<item type="delim" pos="377"><f> </f></item>
<item type="forme" pos="378"><f>il</f></item>
<item type="delim" pos="379"><f> </f></item>
<item type="forme" pos="380"><f>fait</f></item>
<item type="delim" pos="381"><f> </f></item>
<item type="forme" pos="382"><f>un</f></item>
<item type="delim" pos="383"><f> </f></item>
<item type="forme" pos="384"><f>somme</f></item>
<item type="delim" pos="385"><f> </f></item>
<item type="delim" pos="386"><f>:</f></item>
<item type="delim" pos="387"><f>RETURN</f></item>
<item type="delim" pos="388"><f>     </f></item>
<item type="delim" pos="389"><f>     </f></item>
<item type="delim" pos="390"><f>     </f></item>
<item type="delim" pos="391"><f>     </f></item>
<item type="delim" pos="392"><f>     </f></item>
<item type="delim" pos="393"><f>     </f></item>
<item type="delim" pos="394"><f>     </f></item>
<item type="forme" pos="395"><f>Nature</f></item>
<item type="delim" pos="396"><f>,</f></item>
<item type="delim" pos="397"><f> </f></item>
<item type="forme" pos="398"><f>berce</f></item>
<item type="delim" pos="399"><f>-</f></item>
<item type="forme" pos="400"><f>le</f></item>
<item type="delim" pos="401"><f> </f></item>
<item type="forme" pos="402"><f>chaudement</f></item>
<item type="delim" pos="403"><f> </f></item>
<item type="delim" pos="404"><f>:</f></item>
<item type="delim" pos="405"><f> </f></item>
<item type="forme" pos="406"><f>il</f></item>
<item type="delim" pos="407"><f> </f></item>
<item type="forme" pos="408"><f>a</f></item>
<item type="delim" pos="409"><f> </f></item>
<item type="forme" pos="410"><f>froid</f></item>
<item type="delim" pos="411"><f>.</f></item>
<item type="delim" pos="412"><f>RETURN</f></item>
<item type="delim" pos="413"><f>     </f></item>
<item type="delim" pos="414"><f>     </f></item>
<item type="delim" pos="415"><f>     </f></item>
```

```
<item type="delim" pos="416"><f>     </f></item>
<item type="delim" pos="417"><f>     </f></item>
<item type="delim" pos="418"><f>     </f></item>
<item type="delim" pos="419"><f>RETURN</f></item>
<item type="delim" pos="420"><f>     </f></item>
<item type="delim" pos="421"><f>     </f></item>
<item type="delim" pos="422"><f>     </f></item>
<item type="delim" pos="423"><f>     </f></item>
<item type="delim" pos="424"><f>     </f></item>
<item type="delim" pos="425"><f>     </f></item>
<item type="delim" pos="426"><f>RETURN</f></item>
<item type="delim" pos="427"><f>     </f></item>
<item type="delim" pos="428"><f>     </f></item>
<item type="delim" pos="429"><f>     </f></item>
<item type="delim" pos="430"><f>     </f></item>
<item type="delim" pos="431"><f>     </f></item>
<item type="delim" pos="432"><f>     </f></item>
<item type="delim" pos="433"><f>     </f></item>
<item type="forme" pos="434"><f>Les</f></item>
<item type="delim" pos="435"><f> </f></item>
<item type="forme" pos="436"><f>parfums</f></item>
<item type="delim" pos="437"><f> </f></item>
<item type="forme" pos="438"><f>ne</f></item>
<item type="delim" pos="439"><f> </f></item>
<item type="forme" pos="440"><f>font</f></item>
<item type="delim" pos="441"><f> </f></item>
<item type="forme" pos="442"><f>pas</f></item>
<item type="delim" pos="443"><f> </f></item>
<item type="forme" pos="444"><f>frissonner</f></item>
<item type="delim" pos="445"><f> </f></item>
<item type="forme" pos="446"><f>sa</f></item>
<item type="delim" pos="447"><f> </f></item>
<item type="forme" pos="448"><f>narine</f></item>
<item type="delim" pos="449"><f> </f></item>
<item type="delim" pos="450"><f>;</f></item>
<item type="delim" pos="451"><f>RETURN</f></item>
<item type="delim" pos="452"><f>     </f></item>
<item type="delim" pos="453"><f>     </f></item>
<item type="delim" pos="454"><f>     </f></item>
<item type="delim" pos="455"><f>     </f></item>
<item type="delim" pos="456"><f>     </f></item>
<item type="delim" pos="457"><f>     </f></item>
<item type="delim" pos="458"><f>     </f></item>
<item type="forme" pos="459"><f>Il</f></item>
<item type="delim" pos="460"><f> </f></item>
<item type="forme" pos="461"><f>dort</f></item>
<item type="delim" pos="462"><f> </f></item>
<item type="forme" pos="463"><f>dans</f></item>
<item type="delim" pos="464"><f> </f></item>
<item type="forme" pos="465"><f>le</f></item>
<item type="delim" pos="466"><f> </f></item>
<item type="forme" pos="467"><f>soleil</f></item>
<item type="delim" pos="468"><f>,</f></item>
<item type="delim" pos="469"><f> </f></item>
<item type="forme" pos="470"><f>la</f></item>
<item type="delim" pos="471"><f> </f></item>
<item type="forme" pos="472"><f>main</f></item>
<item type="delim" pos="473"><f> </f></item>
<item type="forme" pos="474"><f>sur</f></item>
<item type="delim" pos="475"><f> </f></item>
<item type="forme" pos="476"><f>sa</f></item>
<item type="delim" pos="477"><f> </f></item>
<item type="forme" pos="478"><f>poitrine</f></item>
<item type="delim" pos="479"><f>RETURN</f></item>
<item type="delim" pos="480"><f>     </f></item>
<item type="delim" pos="481"><f>     </f></item>
<item type="delim" pos="482"><f>     </f></item>
<item type="delim" pos="483"><f>     </f></item>
<item type="delim" pos="484"><f>     </f></item>
<item type="delim" pos="485"><f>     </f></item>
<item type="delim" pos="486"><f>     </f></item>
<item type="forme" pos="487"><f>Tranquille</f></item>
<item type="delim" pos="488"><f>.</f></item>
<item type="delim" pos="489"><f> </f></item>
<item type="forme" pos="490"><f>Il</f></item>
<item type="delim" pos="491"><f> </f></item>
```

```
<item type="forme" pos="492"><f>a</f></item>
<item type="delim" pos="493"><f> </f></item>
<item type="forme" pos="494"><f>deux</f></item>
<item type="delim" pos="495"><f> </f></item>
<item type="forme" pos="496"><f>trous</f></item>
<item type="delim" pos="497"><f> </f></item>
<item type="forme" pos="498"><f>rouges</f></item>
<item type="delim" pos="499"><f> </f></item>
<item type="forme" pos="500"><f>au</f></item>
<item type="delim" pos="501"><f> </f></item>
<item type="forme" pos="502"><f>côté</f></item>
<item type="delim" pos="503"><f> </f></item>
<item type="forme" pos="504"><f>droit</f></item>
<item type="delim" pos="505"><f>.</f></item>
<item type="delim" pos="506"><f>RETURN</f></item>
<item type="delim" pos="507"><f>     </f></item>
<item type="delim" pos="508"><f>     </f></item>
<item type="delim" pos="509"><f>     </f></item>
<item type="delim" pos="510"><f>     </f></item>
<item type="delim" pos="511"><f>     </f></item>
<item type="delim" pos="512"><f>     </f></item>
<item type="delim" pos="513"><f>RETURN</f></item>
<item type="delim" pos="514"><f>     </f></item>
<item type="delim" pos="515"><f>     </f></item>
<item type="delim" pos="516"><f>     </f></item>
<item type="delim" pos="517"><f>     </f></item>
<item type="delim" pos="518"><f>     </f></item>
<item type="delim" pos="519"><f>RETURN</f></item>
<item type="delim" pos="520"><f>     </f></item>
<item type="delim" pos="521"><f>     </f></item>
<item type="delim" pos="522"><f>     </f></item>
<item type="delim" pos="523"><f>     </f></item>
<item type="delim" pos="524"><f>RETURN</f></item>
<item type="delim" pos="525"><f>     </f></item>
<item type="delim" pos="526"><f>     </f></item>
<item type="delim" pos="527"><f>     </f></item>
<item type="delim" pos="528"><f>RETURN</f></item>
<item type="delim" pos="529"><f>     </f></item>
<item type="delim" pos="530"><f>     </f></item>
<item type="delim" pos="531"><f>RETURN</f></item>
<item type="delim" pos="532"><f>     </f></item>
<item type="delim" pos="533"><f>RETURN</f></item>
<item type="delim" pos="534"><f>RETURN</f></item>
<item type="delim" pos="535"><f> </f></item>
</items>
</trame>
<cadre>
<acces>
<partition nom="head">
<p n="head" d="53" f="60" nd="16" nf="17"/>
</partition>
<partition nom="tei">
<p n="TEI" d="1" f="534" nd="1" nf="62"/>
</partition>
<partition nom="front">
<p n="front" d="47" f="76" nd="15" nf="20"/>
</partition>
<partition nom="titlestmt">
<p n="titleStmt" d="10" f="28" nd="4" nf="9"/>
</partition>
<partition nom="teiheader">
<p n="teiHeader" d="3" f="33" nd="2" nf="11"/>
</partition>
<partition nom="author">
<p n="author" d="21" f="24" nd="7" nf="8"/>
</partition>
<partition nom="lg">
<p n="lg" d="94" f="206" nd="23" nf="32"/>
<p n="lg" d="213" f="327" nd="33" nf="42"/>
<p n="lg" d="334" f="419" nd="43" nf="50"/>
<p n="lg" d="426" f="513" nd="51" nf="58"/>
</partition>
<partition nom="filedesc">
<p n="fileDesc" d="6" f="31" nd="3" nf="10"/>
</partition>
<partition nom="body">
```

```
<p n="body" d="81" f="524" nd="21" nf="59"/>
</partition>
<partition nom="l">
<p n="l" d="102" f="121" nd="24" nf="25"/>
<p n="l" d="129" f="140" nd="26" nf="27"/>
<p n="l" d="148" f="169" nd="28" nf="29"/>
<p n="l" d="177" f="199" nd="30" nf="31"/>
<p n="l" d="221" f="237" nd="34" nf="35"/>
<p n="l" d="245" f="263" nd="36" nf="37"/>
<p n="l" d="271" f="294" nd="38" nf="39"/>
<p n="l" d="302" f="320" nd="40" nf="41"/>
<p n="l" d="342" f="361" nd="44" nf="45"/>
<p n="l" d="369" f="387" nd="46" nf="47"/>
<p n="l" d="395" f="412" nd="48" nf="49"/>
<p n="l" d="434" f="451" nd="52" nf="53"/>
<p n="l" d="459" f="479" nd="54" nf="55"/>
<p n="l" d="487" f="506" nd="56" nf="57"/>
</partition>
<partition nom="group">
<p n="group" d="38" f="531" nd="13" nf="61"/>
</partition>
<partition nom="dateline">
<p n="dateline" d="66" f="71" nd="18" nf="19"/>
</partition>
<partition nom="text">
<p n="text" d="42" f="528" nd="14" nf="60"/>
</partition>
<partition nom="title">
<p n="title" d="15" f="16" nd="6" nf="6"/>
</partition>
</acces>
</cadre>
</baselexicometrique>
```