

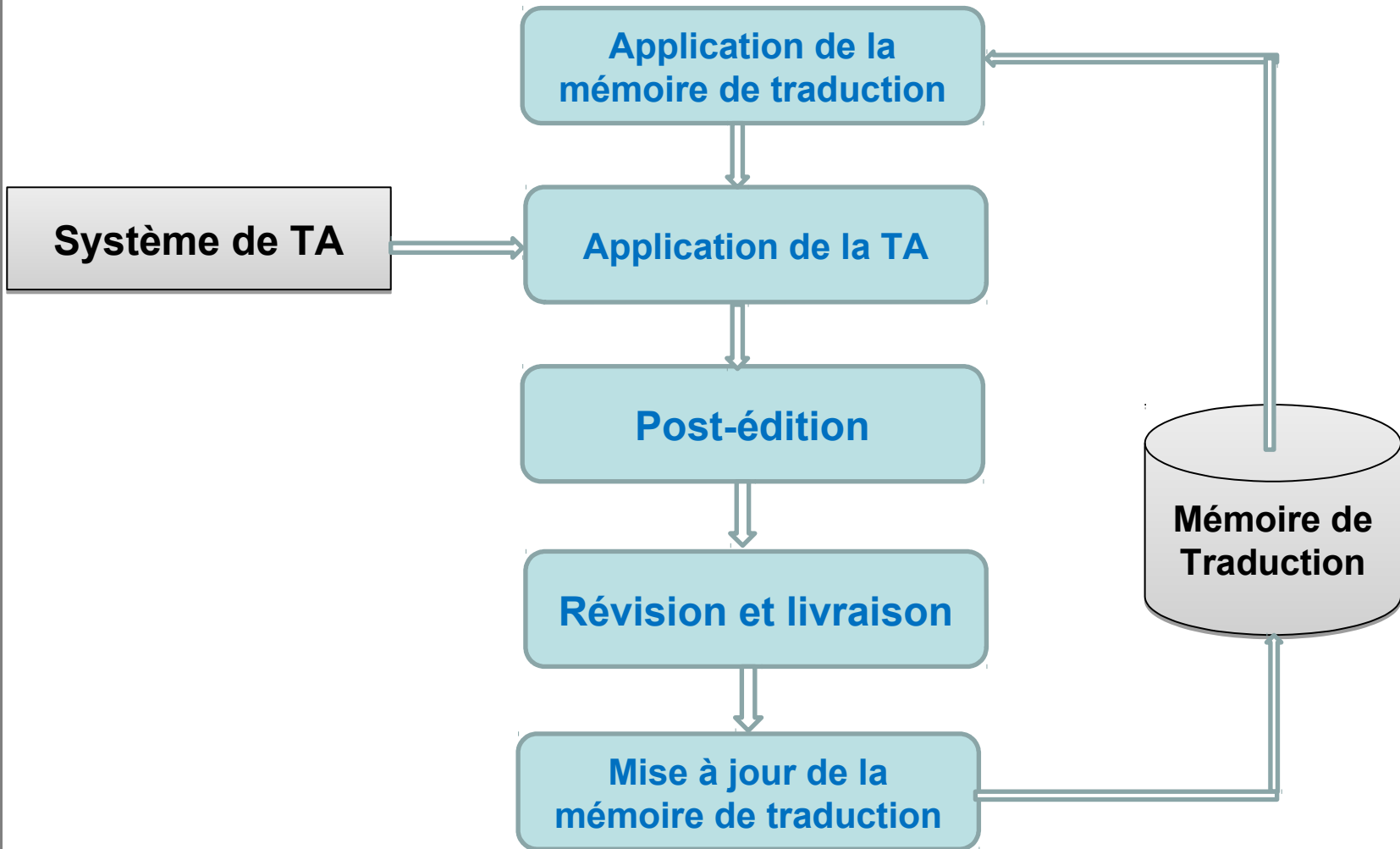
Outils d'aide à la gestion de ressources traductionnelles incrémentales : introduction à la **textométrie multilingue**

Maria ZIMINA-POIROT
mzimina@eila.univ-paris-diderot.fr

CLILLAC-ARP, Université Paris Diderot

Mémoires de Traduction (MT)

TA-TAO : principes



<input type="checkbox"/>	6 juil. 2018 18:48	La vente illégale de cigarettes à des mineurs peut-elle être évitée...	Can illegal cigarette sales to underage youth be prevented, and ...	
<input type="checkbox"/>	6 juil. 2018 18:48	Les interventions de prévention du tabagisme ciblant les jeunes ...	Can smoking prevention interventions targeted at Indigenous yo...	
<input type="checkbox"/>	6 juil. 2018 18:48	Les vaccins antinicotines peuvent-ils aider les personnes à arrêt...	Can nicotine vaccines help people stop smoking or help stop rec...	
<input type="checkbox"/>	6 juil. 2018 18:48	L'hypnothérapie peut-elle aider les personnes qui essaient d'arrê...	Does hypnotherapy help people who are trying to stop smoking	
<input type="checkbox"/>	6 juil. 2018 18:48	Les interventions de sevrage tabagique destinées aux population...	Can smoking cessation interventions targeted at Indigenous pop...	
<input type="checkbox"/>	6 juil. 2018 18:48	La publicité et la promotion du tabac augmentent-elles la probab...	Does tobacco advertising and promotion make it more likely that...	
<input type="checkbox"/>	6 juil. 2018 18:48	Les interventions communautaires peuvent-elles dissuader les je...	Can community interventions deter young people from starting t...	
<input checked="" type="checkbox"/>	6 juil. 2018 18:48	Les interventions basées sur les étapes sont-elles plus efficaces ...	Are stage-based interventions more effective than non-stage-ba...	
<input type="checkbox"/>	6 juil. 2018 18:48	La publicité et la promotion du tabac augmentent-elles la probab...	Does tobacco advertising and promotion make it more likely that...	
<input type="checkbox"/>	6 juil. 2018 18:48	Les interventions communautaires peuvent-elles dissuader les je...	Can community interventions deter young people from starting t...	

Affichage de l'élément 1 à 10 sur 12 éléments (filtré de 19 éléments au total)

Afficher 10 éléments

← Précédent

1

2

Suivant →

← Précédent

Suivant →

Correspondances exactes

Phrase source :

Les interventions basées sur les étapes sont-elles plus efficaces que les autres pour aider au sevrage des fumeurs ?



Phrase traduite : Correspondance avec la Mémoire de Traduction

Are stage-based interventions more effective than non-stage-based ones in helping smokers to quit?



Terminologie :

Autres traductions :

TM

1

MT

1



Correspondance exacte

Are stage-based interventions more effective than non-stage-based ones in helping smokers to quit?

Éditeur de traduction

SYSTRAN Enterprise Server

Nom du fichier : TITRES PLS.docx

Langue source :

Langue cible : Anglais

Profil :Cochrane PLS

📄 Télécharger

✅ Valider

🔄 Gérer la mémoire de traduction

🔍 Voir en contexte

<input type="checkbox"/>	Date	Phrase source	Traduction	Fuzzy Match
<input checked="" type="checkbox"/>	6 juil. 2018 18:48	Le lieu de travail est-il un cadre efficace pour aider les gens à arrêter d...	Is the workplace an effective framework to help people quit smoking?	0

Affichage de l'élément 1 à 1 sur 1 éléments (filtré de 19 éléments au total)

Afficher 10 éléments ← Précédent 1 Suivant →

← Précédent Suivant →

correspondances partielles

Phrase source :

Le lieu de travail est-il un cadre efficace pour aider les gens à arrêter de fumer ?

Phrase traduite : ⚙ Traduction automatique (SYSTRAN)

Is the workplace an effective framework to help people quit smoking? ✅

Terminologie :

Autres traductions : TM 1 MT 1

Fuzzy Match : 76.47%

Source Le lieu de travail est-il un cadre efficace pour que les personnes arrêtent de fumer ?

Cible Is the workplace an effective setting for people to stop smoking

Réponse industrielle : Translation Memory eXchange (TMX)

```
<tu>
<tuv xml:lang="fr">
<seg>Enseignement</seg></tuv>
<tuv xml:lang="en">
<seg>Courses</seg></tuv>
</tu>
<tu>
<tuv xml:lang="fr">
<seg>Emploi du temps:</seg></tuv>
<tuv xml:lang="en">
<seg>Timetable:</seg></tuv>
</tu>
<tu>
<tuv xml:lang="fr">
<seg>Bibliothèques</seg></tuv>
<tuv xml:lang="en">
<seg>Libraries</seg></tuv>
</tu>
<tu>
```

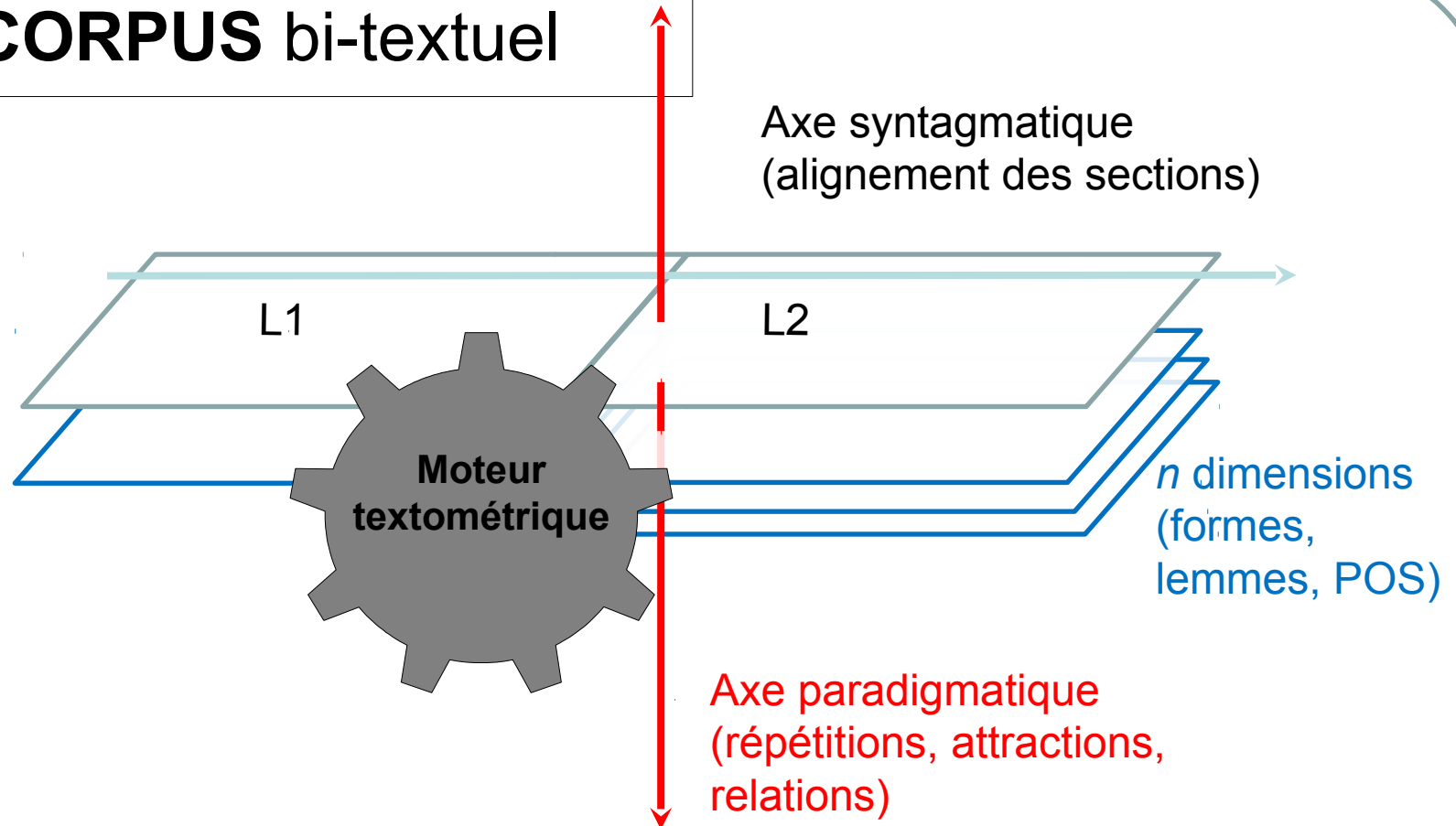
Quelle dynamique ?

FR	EN
Enseignement	Courses
Emploi du temps	Timetable
Bibliothèque	Libraries

La **textométrie** regroupe
l'ensemble des **méthodes**
quantitatives permettant d'opérer
des réorganisations formelles de la
séquence textuelle et des analyses
statistiques portant sur l'ensemble
des unités textuelles d'un **corpus**

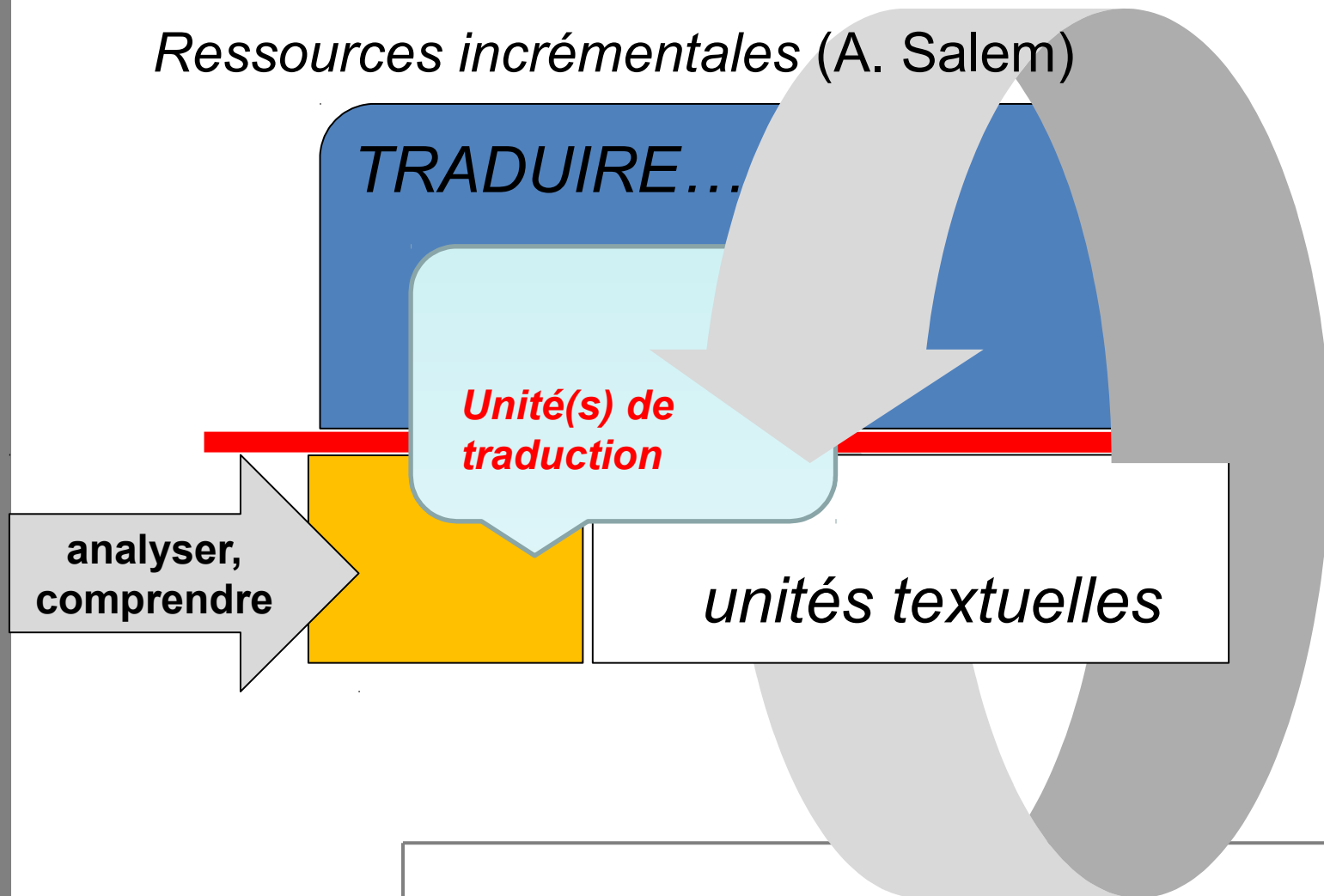
Mémoire de Traduction (MT)

CORPUS bi-textuel



Textométrie et unité(s) de traduction

Ressources incrémentales (A. Salem)

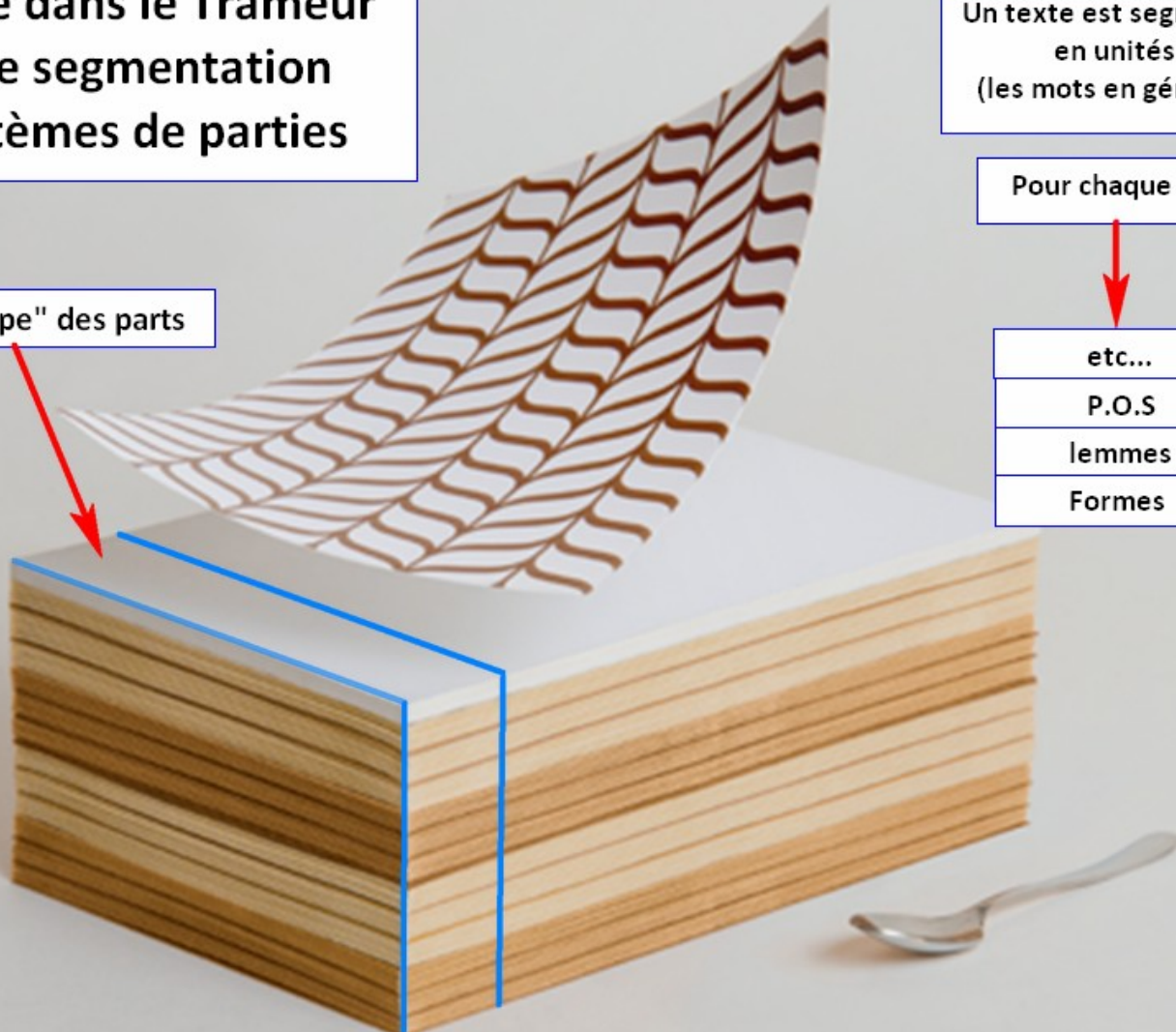


Architecture *Trame* / *Cadre*

Le texte dans le Trameur

1. Une segmentation
2. Systèmes de parties

on "découpe" des parts



Un texte est segmenté
en unités
(les mots en général)

Pour chaque mot

etc...

P.O.S

lemmes

Formes

On compte des « contenus » dans des « contenants »

Méthodologie

- Constructions de plusieurs bi-textes/partitions pour les analyses contrastives
- Etiquetage morpho-syntaxique (*TreeTagger*)
- Bases textométriques multi-annotées
- Analyse textométrique : segments répétés, cooccurrences, patrons, spécificités sur annotations, résonances textuelles
- Retours aux contextes, concordances

Outils : *iTrameur* (S. Fleury)

<http://www.tal.univ-paris3.fr/trameur/iTrameur/>

The screenshot displays the iTrameur web application interface. At the top, the browser address bar shows the URL www.tal.univ-paris3.fr/trameur/iTrameur/. The page title is "iTrameur : Analyse textométrique de données".

The main content area is titled "Paramètres" and contains a grid of configuration options:

PARTITION volet	PARTIES PLS-en	LG CONTEXTE 10	Fq MAX 5
GRAPHE H 400	SEUIL 5	CO-FREQ 2	INDSPMIN 5
GRAPHE L 800	SR LGMAX 12	SR FQMIN 10	NB SÉLECTION SECTION 1
ANNOTATION* 1	ANNOTATION SORTIE 1	RELATION REL	ANNOTATION RELATION 4

Below the grid, there are two input fields: **PÔLE SOURCE** (entrez la forme pôle) and **PÔLE CIBLE** (entrez la forme pôle).

The navigation bar includes the following tabs: Chargement, Trame, Cadre, SR/Patron, Section, Coocs, **Bi-Texte**, Dépendance, Sélection, Export, Aide.

The "Opérations sur un Bitexte" section contains the following options:

Dictionnaire Source	Dictionnaire Cible	Coocs Source*	Coocs Cible*	Spécifs SOURCE*	Spécifs CIBLE*
Sélection <input checked="" type="checkbox"/>	Sélection <input checked="" type="checkbox"/> Source	Sélection <input checked="" type="checkbox"/> Cible	Sélection <input type="checkbox"/>	Sélection <input type="checkbox"/> Source	Sélection <input type="checkbox"/> Cible
Bi-Concordance*	Bi-Concordance* <input type="checkbox"/>	Bi-Concordance* <input checked="" type="checkbox"/> Source	Bi-Concordance* <input checked="" type="checkbox"/> Cible		

Base textométrique *Cochrane*: contexte

- *Cochrane* met à disposition des spécialistes de la médecine des articles de synthèse (*Systematic reviews*)
 - résumés scientifiques destinés aux professionnels
 - résumés réécrits pour le grand public (*Plain Language Summaries: PLS*)
- S'il existe des consignes détaillées concernant les textes pour spécialistes, les consignes pour les PLS et les titres simplifiés sont moins formelles : ils doivent être « compréhensibles pour un locuteur sans formation universitaire » (Higgins et Green, 2011)

Base textométrique *Cochrane* (Gledhill *et al.*, 2017)

- Corpus parallèle anglais-français
- composé des titres de résumés d'articles scientifiques en anglais (*Cochrane Systematic Reviews*) réécrits pour le grand public dans *Plain Language Summaries* (PLS) et traduits en français (2010-2015)

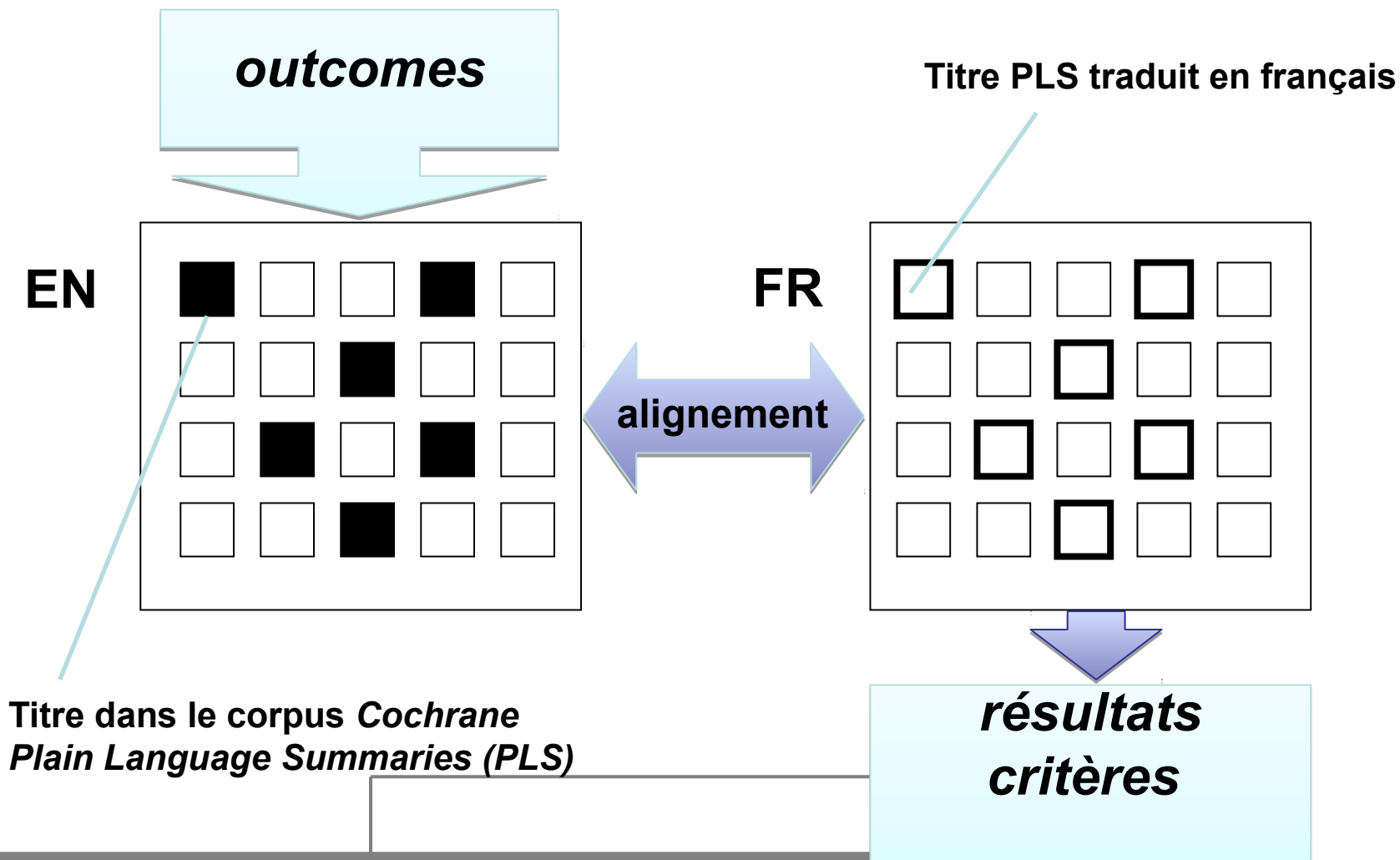
Base Cochrane

- 3 911 titres PLS (anglais/français)

Exemple :

T_ABS_en	T_PLS_en	T_ABS_fr	T_PLS_fr
Enhancing partner support to improve smoking cessation	Are there ways to help partners to give more effective support to people who are trying to quit smoking?	Améliorer le soutien du partenaire pour favoriser le sevrage tabagique	<i>Existe-t-il des manières d'aider les partenaires à apporter un soutien plus efficace aux fumeurs qui tentent d'arrêter de fumer ?</i>

Resonances textuelles (*formes, lemmes, POS*)



La méthode des spécificités

(Lebart et Salem, 1994. *Statistique textuelle.*)

PARTIES

Unités textuelles			
		K_{ij}	F_i
		t_j	

Si l'effectif K_{ij} ne se situe pas dans les limites de ce que le **calcul probabiliste** permettait d'espérer, on calcule un *indice de spécificité* : **sur-emploi** ou **sous-emploi** de l'unité (spec. **$\pm XX$**)

Tableau lexical :

K_{ij} : fréquence de l'unité j dans la partie i

F_i : fréquence de i dans le corpus

t_j : taille de la partie j

Principes de navigation textuelle

résonances



SOURCE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>outcomes</u>	56	56	97
<u>improving</u>	54	17	17
<u>improve</u>	78	15	11
<u>maternal</u>	18	8	10
<u>health</u>	93	14	9

CIBLE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>résultats</u>	46	33	46
<u>critères</u>	12	12	21
<u>améliorer</u>	111	25	19
<u>mère</u>	31	11	12
<u>sur</u>	163	20	11

N°	Contexte Source	Contexte Cible
31	Training, experience and volumes on <u>outcomes</u> for persons living with HIV/AIDS: A systematic review §	Formation, expérience et volumes <u>sur les résultats</u> pour les personnes vivant avec le VIH/SIDA : Une revue systématique §
32	Trials of interventions for pregnant women who are obese to lose weight and <u>improve</u> pregnancy <u>outcomes</u> . §	Essais consistant en des interventions dont l'objectif est de faire perdre du poids et d' <u>améliorer les résultats</u> de la grossesse de femmes enceintes obèses. §
33	Vitamin A supplementation during pregnancy for <u>maternal</u> and newborn <u>health outcomes</u> §	Supplémentation en vitamine A pendant la grossesse pour l'amélioration des <u>résultats</u> chez la <u>mère</u> et son nouveau-né §

Principes de navigation textuelle

résonances



SOURCE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>outcomes</u>	56	56	97
<u>improving</u>	54	17	17
<u>improve</u>	78	15	11
<u>maternal</u>	18	8	10
<u>health</u>	93	14	9

CIBLE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>résultats</u>	46	33	46
<u>critères</u>	12	12	21
<u>améliorer</u>	111	25	19
<u>mère</u>	31	11	12
<u>sur</u>	163	20	11

N°	Contexte Source	Contexte Cible
1	Antithrombotic therapy for <u>improving maternal</u> or infant <u>health outcomes</u> in women considered at risk of placental dysfunction §	Traitement antithrombotique pour <u>améliorer les critères</u> de <u>santé</u> chez la <u>mère</u> ou le nourrisson pour <u>les</u> femmes présentant un risque de dysfonctionnement placentaire §
2	Different magnesium sulphate regimens given to women at risk of preterm birth to help protect the baby's brain and <u>improve</u> long-term <u>outcomes</u> §	Différentes posologies de sulfate de magnésium administrées à la femme présentant un risque d'accouchement prématuré pour protéger le cerveau du bébé et <u>améliorer les critères</u> d'évaluation à long terme §
3	Does an aminophylline injection in addition to bronchodilators for an asthma attack <u>improve</u> lung function and other <u>outcomes</u> or cause harm? §	En cas de crise d'asthme, est-ce qu'une injection d'aminophylline en supplément aux bronchodilatateurs améliore la fonction pulmonaire et d'autres <u>critères</u> de <u>résultat</u> ou bien cause des préjudices ? §

Principes de navigation textuelle



SOURCE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>results</u>	7	7	21
<u>testing</u>	11	2	5

CIBLE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>résultats</u>	46	6	11
<u>examens</u>	2	2	7
<u>communication</u>	7	2	5
<u>téléphone</u>	10	2	5

N° ▲	Contexte Source	Contexte Cible
1	Mobile phone messaging for communicating <u>results</u> of medical investigations §	La messagerie par <u>téléphone</u> portable pour la <u>communication</u> de <u>résultats</u> d' <u>examens</u> médicaux §
2	The effects of using email to send test <u>results</u> to patients §	Les effets de l'utilisation du courrier électronique pour l'envoi aux patients de leurs <u>résultats</u> d' <u>examens</u> §

Induction sur annotations

résonance

Titres PLS en anglais


Titres PLS en français

SOURCE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>MD</u>	122	122	157
<u>VB</u>	56	41	40
<u>VV</u>	755	127	31
<u>RB</u>	389	52	10
<u>RBR</u>	29	12	9

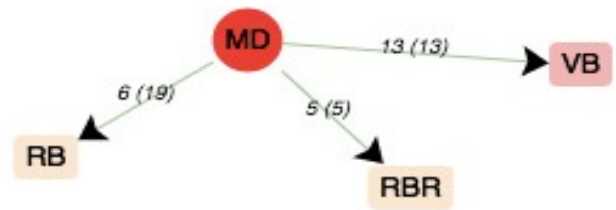
CIBLE			
Cooc	FqCooc total	FqCooc contexte	IndSP
<u>VER:pres</u>	746	148	47
<u>VER:infi</u>	1187	175	37
<u>VER:cond</u>	33	27	30
<u>ADV</u>	745	99	19
<u>DET:POS</u>	78	13	5

N° ▲	Contexte Source	Contexte Cible
1	A combination of cyclophosphamide or azathioprine with steroids can improve kidney function in people with severe lupus nephritis §	Une combinaison de cyclophosphamide ou d'azathioprine avec des stéroïdes peut améliorer la fonction rénale chez les personnes souffrant d'une néphrite lupique sévère §
2	A long course (five days or more) of a low dose of corticosteroids may be considered as an adjunct therapy in patients with septic shock. §	Une longue cure (cinq jours ou plus) d'une faible dose de corticostéroïdes peut être considérée comme un traitement d'appoint chez les patients souffrant de choc septique. §

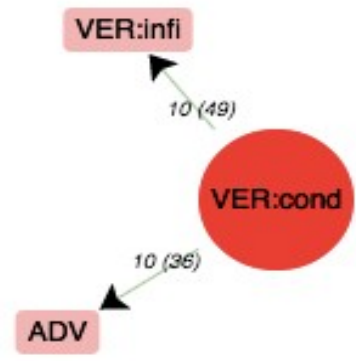
Cooccurrences

Couleur Noeud-Cooc : IndicSp < 10 < IndicSp < 20 < IndicSp < 30 < IndicSp < 50 < IndicSp
 Arc Label : POLE - IndiceSpécif(Co-Freq) → COOC


Titres PLS en anglais



Titres PLS en français



3	Atypical antipsychotics benefit people with dementia but the risks of adverse events may outweigh the benefits, particularly with long term treatment §	Les antipsychotiques atypiques sont bénéfiques pour les personnes atteintes de démence, mais les risques d'événements indésirables pourraient être plus importants que les bénéfices, en particulier avec un traitement à long terme §
4	Catheter ablation would be an alternative to inhibit recurrence of paroxysmal or persistent atrial fibrillation §	L'ablation par cathéter pourrait constituer une alternative pour l'inhibition de la récurrence de la fibrillation auriculaire paroxystique ou persistante §
5	Centralisation of care may prolong survival in women with ovarian cancer, and possibly more generally , gynaecological cancer §	La centralisation des soins pourrait prolonger la survie des femmes atteintes d'un cancer de l'ovaire et peut-être, plus généralement , d'un cancer gynécologique §

...varier le paradigme (les éléments plus variables du schéma), afin d'obtenir des effets de style ou de mieux intégrer la construction dans le discours environnant. (Gledhill, 2008).

Titres PLS en anglais (extrait)

Item	FQ	fq	Sp
VV	957	755	60
DT	2032	1432	56
MD	123	122	32
IN/that	32	32	10

Titres PLS en français (extrait)

Item	FQ	fq	Sp
VER:pres	915	746	73
VER:infi	1810	1187	27
PRO:REL	149	133	22

Corrélation de formes => routine discursive ou collocation syntagmatique

DT

DT

IN/that

MD VV

DT

PRO:REL

VER:pres VER:infi

A review of the ways that healthcare professionals can improve the use of suitable medicines for older people §

Revue des méthodes par lesquelles les professionnels de santé peuvent améliorer l'usage des médicaments appropriés aux personnes âgées §

Ressources incrémentales

« En faisant en sorte que les traitements textométriques appliqués successivement à une ressource textuelle /.../ enrichissent le cadre textométrique /.../, on obtient un schéma de ressources textuelles incrémentales. »

Keyser Söze-Duval, 2008. *Pour une textométrie opérationnelle*. En ligne <http://www.tal.univ-paris3.fr/trameur/RTI6provisoire.doc>

Traduction et qualité

“Quality is, in a sense, the return of the repressed translation detail. The careful, detailed attention to text, language, and meaning that is implicit in the act of translation re-emerges in the context of automation in the debates about the extent and role of post-editing, and about how to achieve acceptable quality in translation output.”

Michael Cronin (2013)

Conclusion et perspectives

- Evolution des Mémoires de Traduction avec prise en compte des ressources textuelles incrémentales
- Approche paradigmatique au delà des représentations tabulaires « figées »
- Unité(s) de traduction : à la frontière du quantitatif et du qualitatif
- Dynamique de l'exploration (modélisation conceptuelle – prototypes – implémentation – ergonomie)

Merci

mzimina@eila.univ-paris-diderot.fr

Bibliographie

CRONIN, M., *Translation in the Digital Age*, Milton Park, Abingdon, Oxon, Routledge, 2013.

GLEDHILL, C., 2008, « Le signe et le syntagme : entre phraséologie et lexico-grammaire. Une synthèse du modèle systémique fonctionnel de Michael Halliday et de la théorie sémiotique de Charles S. Peirce ». *Habilitation à diriger des recherches*. Université de Bretagne Occidentale.

GLEDHILL, C., MARTIKAINEN, H., MESTIVIER, A. & ZIMINA, M. (2017), *Applying Textometric Analysis to a Description of Cochrane Medical Abstracts and their Plain Language versions*, CILC 2017.

GLEDHILL, C., PATIN, S., ZIMINA, M., 2017, « Lexico-grammaire et textométrie : identification et visualisation de schémas lexico-grammaticaux caractéristiques dans deux corpus juridiques comparables en français », *Corpus*, vol. 17 « segments et séquences textuelles : méthodologie et caractérisation ».

HARRIS, B., "Bi-text: A New Concept in Translation Theory." *Language Monthly*, no. 54, 1988.

ROSSI, C., 2017, De quoi la traduction automatique est-elle le nom ? *En attendant Nadeau. Journal de la littérature, des idées et des arts*. En ligne :

<https://www.en-attendant-nadeau.fr/2017/07/25/traduction-automatique/>

SALEM, A., « Introduction à la résonance textuelle », Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles, 10-12 mars 2004, Louvain-la-Neuve, Université catholique de Louvain, éd. Purnelle, Gérard, Fairon, Cédric, Dister, Anne, Louvain-la-Neuve, Presses Universitaires de Louvain, vol. 2, 2004, p. 986-992.

ZIMINA M., FLEURY, S., « Approche systémique de la résonance textuelle multilingue », Actes des 12es Journées Internationales d'Analyse Statistiques des Données Textuelles, Paris, INaLCO, 3-6 juin 2014, éd. Née, Émilie, Daube Jean-Michel, Valette, Mathieu, Fleury, Serge, Paris, JADT.org, 2014, p. 717-728.

ZIMINA, M., FLEURY, S., « Perspectives de l'architecture Trame/Cadre pour les alignements multilingues ». *Nouvelles perspectives en sciences sociales : revue internationale de systémique complexe et d'études relationnelles*, vol. 11, no 1, 2015.