

La recherche SHS sur le vin : le point de vue de la sémantique et de la communication

19 octobre 2018 – Université de Craiova, Roumanie

Pr Laurent Gautier, Centre Interlangues Texte Image Langage
(UBFC, EA 4182)





1. Objectifs

- Présenter l'**écosystème de recherche et de la valorisation** sur Vigne et Vin sur le grand campus dijonnais avec un focus sur les SHS
 - Discuter quelques **problématiques de recherche** possibles à l'intersection entre **sciences du langage** et **info-com**
 - Présenter un tour d'horizon des **ressources numériques** constituées et de leurs enjeux pour la **recherche sémantique** en termes **méthodologiques**
- 
- 

2. Éléments de contexte

- Un écosystème local et régional au service et / ou en appui à la recherche sur Vigne et Vin incluant les SHS :



Donnons goût à l'innovation!



- Plusieurs projets achevés ou en cours autour des **discours** professionnels de la filière viti-vinicole reposant sur des **corpus numériques** :
 - Projet européen **VinoLingua** à visée didactique (D, E, F, I) (Lavric 2013, Gautier 2014)
 - Projet régional **OenoLex Bourgogne** à visée lexicographique et cognitive – corpus oral (F, extensions plurilingues en cours de négociation) (Gautier/Leroyer 2016)
 - Projet Interreg IV F-CH : construction sémantique et sensorielle de la **minéralité** (F) – triple corpus : écrit, oral, expérimental (Deneulin et al. 2014)

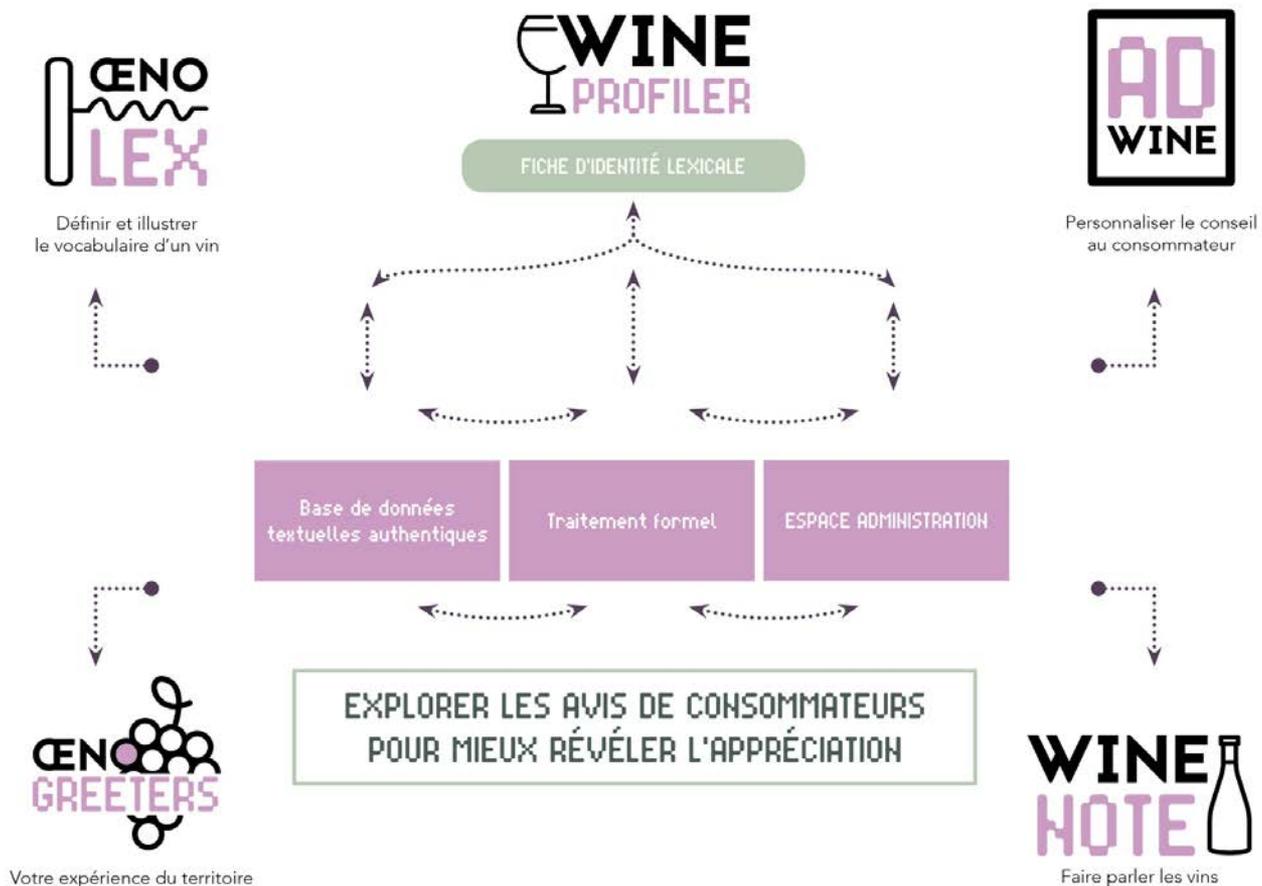
- 
- Projet BSN autour du **Bulletin de l'Organisation Internationale de la Vigne et du Vin** (Gautier 2015) (D, E, F, GB, I)
 - Projet **Bibliothèque numérique** de l'OIV (F)
 - Projet **Crémant de Bourgogne** en lien avec UPECB (F) – triple corpus (Mancebo-Humbert/Gautier 2015)
 - Projet **Mémoires du vin** avec la communauté de communes de Gevrey-Chambertin (F)
 - Projet **Adwine** de recommandation consos (lauréat CVT-Athena) (F)

=> Synthèse dans une offre globale de service
Words4Wine



Words4Wine

Quand le numérique révolutionne la communication autour du vin



3. Typologie des ressources « langagières »

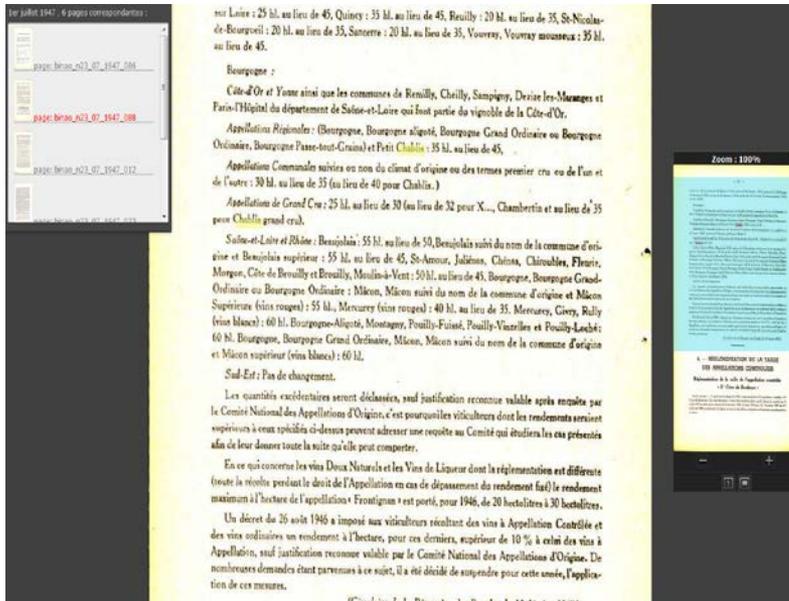
- Différents modes de typologisation possibles
- **Par type de données :**
 - Textes
 - Images fixes + données géo-spatialisées
 - Données audio(-visuelles)
 - Corpus de « paroles » (*cf. infra*)
- **Par origine :**
 - Données existantes (bibliothèques, archives, revues)
 - Données produites
 - Sans stimulus : questionnaires, entretiens
 - Avec stimulus : données de dégustation

Quelques focus sur des données à partager :

– Le Bulletin de l'OIV depuis 1928 :

- Volume : 98 000 pages / 5.4 To
- Numérisation => fichiers images + pdf

production de fichiers normalisés XML-ALTO et XML-METS permettant une recherche en plein texte depuis portail PANDOR
création de fichiers TXT pour travail sur corpus



- Des corpus audio(-visuels) de **discours professionnels**
 - Les enregistrements de **dégustation** (OenoLex, Adwine) :
 - Authentiques
 - Différentes situations de dégustation
 - Flux audio + transcription alignée + annotations sous @SONAL

Vue d'ensemble du corpus

Situation	Nombre d'enregistrements	Durée
Didactique	4	4h33min20s
Marketing direct	3	2h9min5s
Marketing promotionnel	5	1h33min19s
	12	8h15min44s

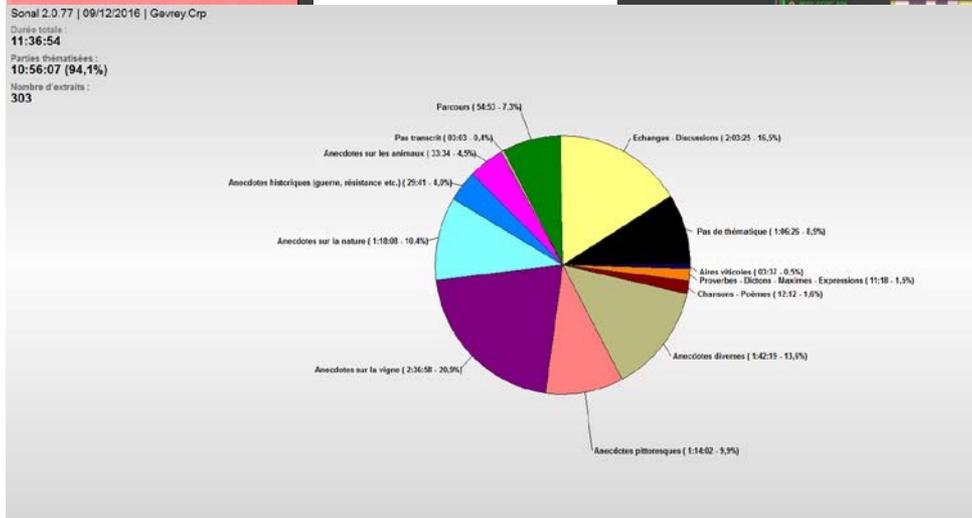
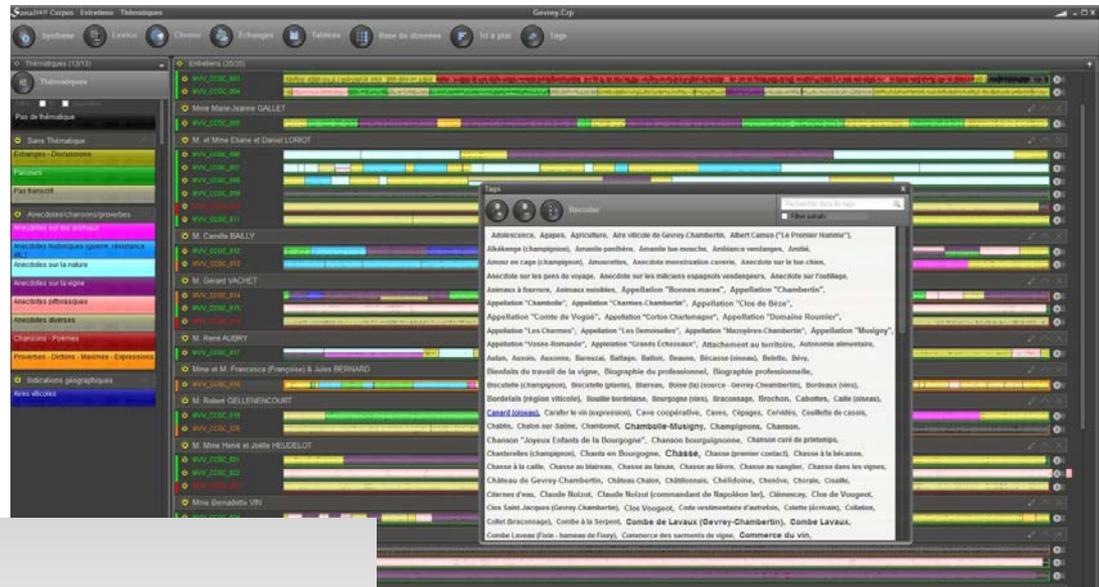
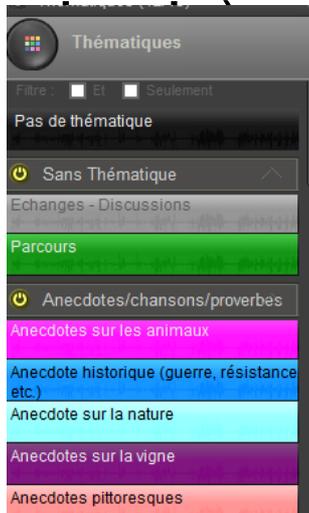
- 
- Des corpus audio(-visuels) de discours professionnels
 - Les « **mémoires du vin** » de Gevrey Chambertin (12 heures) :
 - Même mode de traitement
 - Visée documentaire et valorisation socio-économique du territoire via une application

[414,8] [>Témoin]: alors quand mon père a vu ça, il dit on peut pas rester là c'était la guerre et puis il savait pas si si il repartirait, si il repartirait pas mais là nan il a dit on restera en France. Alors on a acheté une maison en France mes parents ont achetés une maison en France y z'ont y z'ont voulu s'installer là du coup...

- 
- Des **corpus de « paroles »** produites en appui à des thèses :
 - Ludivine DOMONT-LEVEQUE : Minéral/minéralité : étude diachronique de la construction d'un descripteur sensoriel dans les textes réglementaires, prescriptifs et descriptifs de la filière vitivinicole (français, allemand, anglais)
 - Corpus trilingue
 - Guides + revues + blogs + forums
 - Marielle MANCEBO-HUMBERT : La construction discursive et terminologique dans la filière viti-vinicole : mieux communiquer auprès des consommateurs et des experts sur le crémant de Bourgogne.
 - Guides + revues + comm. Interne
 - Questionnaires sans stimulus
 - Paroles produites en dégustation

4. Traitement et pistes méthodologiques d'exploitation

- Exemple de traitement documentaire oral (mémoires)



Un exemple de traitement sémantique

- Une proposition de tripartition du lexique impliquant un traitement différencié (Gautier 2018, 323-328)
- Termes susceptibles d'une description objectiviste, dénommant des propriétés définitoires du vin en tant qu'objet physique, exemple typique des cépages

Afin de réaliser une **caractérisation** précise et rapide des variétés de vigne, une « Liste des **descripteurs primaires prioritaires** » a été élaborée dans le cadre du programme GenRes 96 n°81. L'utilisation de cette liste comprenant 14 descripteurs est recommandée par l'OIV. Les descripteurs retenus ont un **fort pouvoir discriminant entre les variétés**. La plupart sont faciles à **noter**. [...] Comme pour la gestion des collections de ressources génétiques, il se révèle utile de compléter la description des variétés de vigne par des données supplémentaires. (OIV 2009, s.p., souligné par nous, LG)

Il y a mots et mots de la dégustation...

- Termes sensoriels, dont le sens se construit à la fois dans l'expérience de dégustation *et* par le discours produit, à saisir sur un mode constructiviste et en termes de prototypes

Acquise par Louis Latour, la maison Simmonet-Febvre ***champagnise*** le chablis depuis le milieu du XIXe. C'est aujourd'hui l'unique élaborateur d'effervescents dans ce vignoble, travaillant aussi pour divers viticulteurs . (Guide Hachette 2005)

- Les descripteurs hédoniques à dimension évaluative :

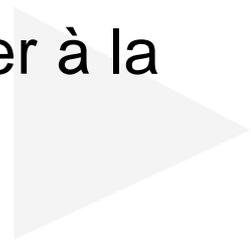
Voilà un vin simple, mais terriblement bien fait et ***follement buvable***. Dominé par les fruits blancs, avec suffisamment de tension pour trouver un bel équilibre.

=> Difficulté : peuvent se superposer aux deux autres catégories

- Approches méthodologiques en cours
 - **Sémantique** lexicale et cognitive : cartographie (interculturelle) des champs de la sensorialité, représentations mentales
 - Traditions **discursives** : émergence et développement des genres propres à la filière selon les langues-cultures
 - Fouille de textes (*text mining*) et **analyse de sentiments** appliquée au vin (*sentiment analysis*)
 - Approches **multimodales** : présence du vin sur les réseaux sociaux



5. Bilan et perspectives

- Ere des données (y compris en SHS)
 - Interactions bienvenues avec les différentes disciplines
 - Savoir passer de la ressource statique au corpus « intelligent »
 - Vraie dynamique offerte par les Humanités Numériques en Europe pour des interconnexions plurilingues et interculturelles (Darjah, Europeana,...)
 - Bases méthodologiques éprouvées pour passer à la comparaison interlangue et interculturelle
- 
- 



Merci de votre attention !

Pr. Laurent Gautier (laurent.gautier@u-bourgogne.fr)