

# Description sémantique des familles dérivationnelles dans un cadre paradigmatique

Daniele Sanacore <sup>1</sup>

<sup>1</sup>CLLE, CNRS & University of Toulouse

21 octobre

Séminaire “Thématiques actuelles de la recherche en TAL”



## Cadre du travail

Le travail ici présenté s'inscrit dans le cadre de ma thèse codirigée par

- **Nabil Hathout** (CLLE, CNRS & Université de Toulouse)
- **Fiammetta Namer** (ATILF, CNRS & Université de Lorraine).

Il est destiné à alimenter *Démonette*, une ressource lexicale du français, et s'inscrit dans le cadre du projet “Démonext” (ANR-17-CE23-0005).

## Problématique

Dans le cadre de la **morphologie dérivationnelle paradigmatique**, comment est-ce que on peut décrire les **relations morphosémantiques** contenues dans les **familles dérivationnelles**?

## Méthode

Nous allons voir dans quelle mesure des structures inspirées de la **Sémantique des Frames** pourraient être utilisées pour représenter le sens des familles dérivationnelles et comment ces représentations s'organisent en paradigmes.

## Exemple de famille dérivationnelle

**laver**<sub>v</sub>; **lavage**<sub>n</sub>; **lavoir**<sub>n</sub>; **laverie**<sub>n</sub>; **laveur**<sub>n</sub>; **laveuse**<sub>n</sub>; **lavette**<sub>n</sub>; **lavable**<sub>a</sub>;  
**lavement**<sub>n</sub>...

## Exemple de frame

An **Avenger** performs a **Punishment** on a **Offender** as a consequence of an earlier action by the **Offender**, the **Injury**. The **Avenger** inflicting the **Punishment** needs not be the same as the **Injured\_party** who suffered the **Injury**, but the **Avenger** does have to share the judgment that the **Offender's** action was wrong. The judgment that the **Offender** had inflicted an **Injury** is made without regard to the law.

# Pourquoi la Sémantique des Frames?

## Les frames contiennent des participants

La sémantique des frames, comme l'on verra dans cette présentation, représente des situations dans des objets nommés **frames**. Ces frames contiennent plusieurs **participants** qui sont impliqués dans les situations décrites par les frames et réunis dans un seul objet.

## Éléments des familles dérivationnelles comme éléments d'un frame

Ce que nous cherchons c'est concevoir une représentation des relations sémantiques ou les membres d'une famille soient réunis au sein de frames.

## Ressources modernes en morphologie

- Les ressources modernes en morphologie dérivationnelle ont été caractérisés par l'adoption du lexème comme unité minimale (Fradin, 2003)
- structure paradigmatique du lexique en morphologie dérivationnelle, avec la famille comme élément central
- double fonction des constructions dérivationnelles: créer des nouveaux lexèmes et établir des relations de motivation sémantiques et formelles entre eux (Namer and Hathout, 2019).

## Evolution dans la 'morphologie dérivationnelle'

Deux facteurs ont contribué à la récente évolution dans l'étude des formations des mots:

- adoption du lexème comme unité minimale au lieu du morphème
- structuration du lexique morphologique en paradigmes

## Abandon du morphème comme unité minimale

Les désavantages les plus significatifs concernent la rigidité du morphème, qui ne peut pas être facilement adapté aux derivations qui présentent un décalage entre forme et sens, comme la conversion (1) ou des phénomènes parasynthétiques (2):

(1) *nurse*<sub>N</sub> - *nurse*<sub>V</sub> (anglais)

(2) *banque*<sub>N</sub> - *interbancaire*<sub>A</sub> (français)

## Famille dérivationnelle

Une famille dérivationnelle est un ensemble de lexèmes connectés par des relations morphologiques dérivationnelles (Hathout, 2011). Un exemple de famille dérivationnelle pour le Français c'est la famille (partielle) construite autour du verbe *laver*:

## Famille partielle de *laver*

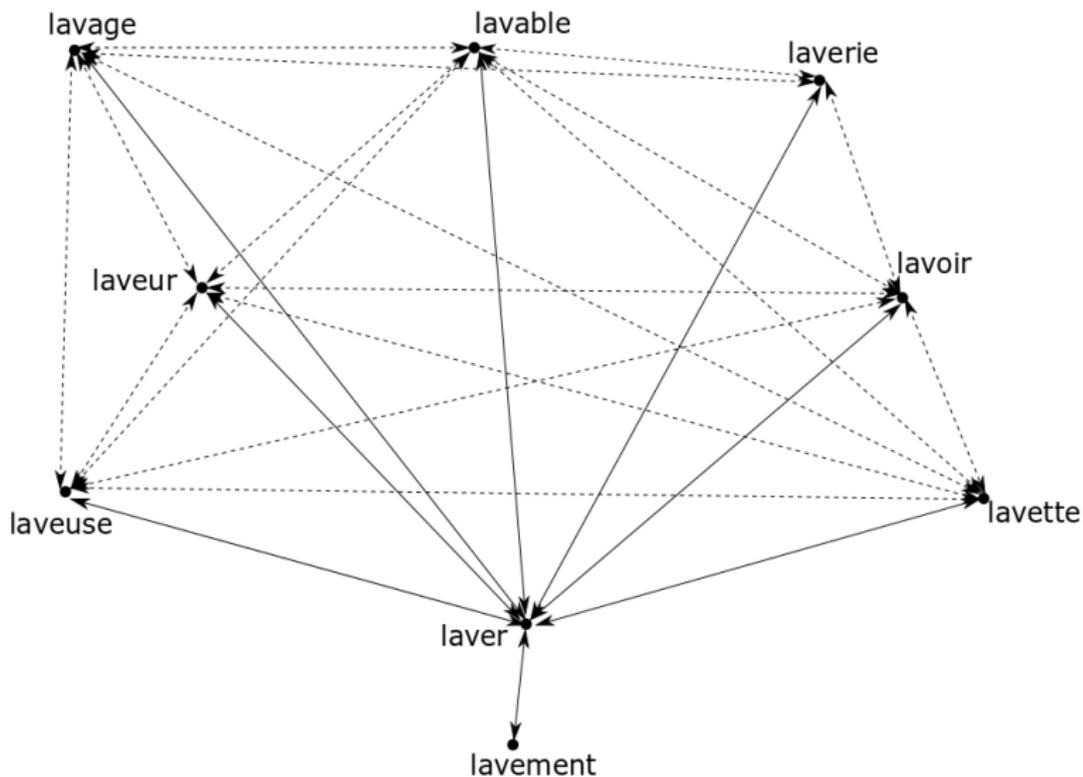
**laver**<sub>v</sub>; **lavage**<sub>n</sub>; **lavoir**<sub>n</sub>; **laverie**<sub>n</sub>; **laveur**<sub>n</sub>; **laveuse**<sub>n</sub>; **lavette**<sub>n</sub>; **lavable**<sub>a</sub>;  
**lavement**<sub>n...</sub>

## Relations dérivationnelles directes et indirectes

Dans une famille, on peut avoir deux types de relations dérivationnelles: relations directes et indirectes.

- Une **relation dérivationnelle directe** connecte un lexème avec un de ces ascendants ou descendants.  
laver<sub>v.</sub> → laveur<sub>n.</sub>
- En revanche, une **relation dérivationnelle indirecte** connecte des éléments plus distants de la famille.  
laveur<sub>n.</sub> → lavage<sub>n.</sub>

# Relations directes et indirectes dans la famille (partielle) de *laver*



## Paradigmes et séries dérivationnelles

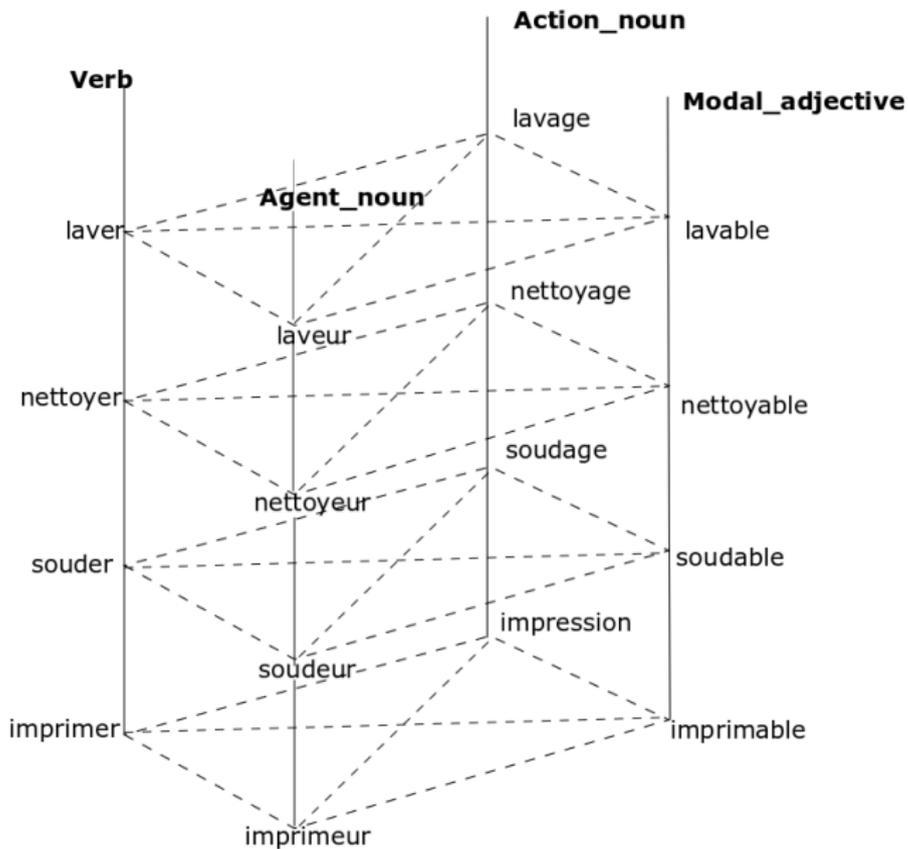
- Un **système paradigmatic** est une collection de familles dérivationnelles partielles qui se trouvent alignées pour ce qui concerne les relations de contenu qui existent parmi ses membres (Bonami and Strnadová, 2018)
- le **contenu** c'est la spécification des propriétés syntaxiques et sémantiques d'un mot, alors que la **forme** est la spécification de son orthographe et phonologie
- plusieurs relations de contenu parmi les membres de famille différentes qui se trouvent alignées forment des **séries dérivationnelles**

## Exemple de paradigme et de série

Prenons par exemple les quatre sous-familles construites autour des verbes *imprimer*, *souder*, *laver* et *nettoyer*:

<b>verbe</b>	<b>agent_m</b>	<b>adj</b>	<b>nom d'action</b>
imprimer	imprimeur	imprimable	impression
souder	soudeur	soudable	soudage
laver	laveur	lavable	lavage
nettoyer	nettoyeur	nettoyable	nettoyage

# Système paradigmatique



*Démonette* (Hathout et al., 2017; Hathout and Namer, 2014) est une ressource lexicale consacrée à la description de la formation des mots en Français. *Démonette* est basée sur l'hypothèse que la morphologie est relationnelle et chaque relation qui implique un lexème donné contribue à son sens

- *Démonette* cherche une description complète, redondante et explicite des propriétés des relations dérivationnelles
- Les entrées dans *Démonette* **ne décrivent pas le propriétés des lexèmes**, mais ils décrivent les propriétés des relations dérivationnelles qui connectent des paires de lexèmes
- *Démonette* est donc un graphe où une relation ( $w_1 \leftarrow w_2$ ) décrit la motivation morphologique de  $w_1$  par rapport à  $w_2$ . La majorité des lexèmes sont connectés entre eux dans les deux directions (Hathout and Namer, 2016).

## Description sémantique dans *Démonette*

La description **sémantique des relations dérivationnelles** est fournie en utilisant quatre champs dans la base des données. Deux champs sont dédiés aux types sémantiques de  $w_1$  and  $w_2$ , le troisième contient la définition concrète de  $w_1$  relativement à  $w_2$  et le quatrième contient la définition abstraite de  $w_1$  relativement à  $w_2$ . Dans la définition abstraite,  $w_2$  est remplacé par son type sémantique.

### Types sémantiques et définitions concrètes et abstraites dans *Démonette*

W1	W2	Type W1	Type W2	Concrete definition	Abstract definition
laveuse	laver	@AGF	@	"she who performs the action of laver"	"she who performs the action of @"
nettoyeuse	nettoyer	@AGF	@	"she who performs the action of nettoyer"	"she who performs the action of @"
imprimeuse	imprimer	@AGF	@	"she who performs the action of imprimer"	"she who performs the action of @"

## Problème

La **catégorie ontologique** et le **rôle sémantique** sont mélangés dans une seule étiquette (e.g. @AGF for *laveuse*, *nettoyeuse* et *imprimeuse*). Comme la catégorie ontologique associée aux lexèmes est indépendante de son rôle sémantique, il faudrait avoir deux étiquettes distinguées.

## Par exemple...

Avec le système de typage actuel, un nom d'instrument comme *aspirateur* et un agent humain comme *observateur* seront typés avec la même étiquette @AGM.

La **Frame Semantics** se base sur l'hypothèse que nous comprenons le langage au moyen de situations (des représentations du monde réel) qui sont évoquées dans notre esprit par des mots. Ces représentations des situations évoquées dans notre esprit sont nommées frames (Fillmore et al., 1976).

## REVENGE frame

An **Avenger** performs a **Punishment** on a **Offender** as a consequence of an earlier action by the **Offender**, the **Injury**. The **Avenger** inflicting the **Punishment** needs not be the same as the **Injured\_party** who suffered the **Injury**, but the **Avenger** does have to share the judgment that the **Offender's** action was wrong. The judgment that the **Offender** had inflicted an **Injury** is made without regard to the law.

**FrameNet** (Ruppenhofer et al., 2006), une ressource lexicale pour l'anglais, **est une implementation de la Sémantique des Frames**. *FrameNet* met en relation les mots avec leur sens à travers des frames qu'ils activent. Par ailleurs, *FrameNet* enregistre comme les phrases et les syntagmes sont structurés autour des mots.

## Objectifs de *FrameNet*

- caractériser les frames et trouver les unités lexicales qui les évoquent
- développer une terminologie descriptive pour chaque frame (éléments core et non core, etc.)
- extraire des phrases qui réalisent chaque frame (pour pouvoir après utiliser la ressource pour annoter des corpus)

# Frames dans FrameNet

Les frames représentent des fragments d'une histoire/situation caractérisés par un nombre donné de participants qu'y sont impliqués:

## REVENGE frame

An **Avenger** performs a **Punishment** on a **Offender** as a consequence of an earlier action by the **Offender**, the **Injury**. The **Avenger** inflicting the **Punishment** needs not be the same as the **Injured\_party** who suffered the **Injury**, but the **Avenger** does have to share the judgment that the **Offender's** action was wrong. The judgment that the **Offender** had inflicted an **Injury** is made without regard to the law.

Phrases qui réalisent le REVENGE frame:

1. **They** took *revenge* for the deaths of two loyalist prisoners.
2. **Lachlan** went out to *avenge* **them**.
3. The next day, **the Roman forces** took *revenge* on **their enemies**.

**Les sens d'un lexème polysémique sont associés à un frame différent** (Ruppenhofer et al., 2006). Par exemple, le lexème *bake* évoque trois frames différents:

## Trois sens différents pour *bake*

- APPLY\_HEAT: Michelle *baked* the potatoes for 45 minutes.
- COOKING\_CREATION: Michelle *baked* her mother a cake for her birthday.
- ABSORB\_HEAT: The potatoes have to *bake* for more than 30 minutes.

## Core vs Non-core

**Core frame elements:** éléments qui sont essentiels pour la compréhension du frame (e.g. AVENGER, PUNISHMENT, OFFENDER, INJURY, INJURED\_PARTY)

**Non-core frame elements:** éléments qui sont généralement présents sur plusieurs frames (e.g. DEGREE, DURATION, MANNER, PLACE, TIME)

Par ailleurs, chaque frame est associé à un ensemble donné d'**unités lexicales** qui l'évoquent. Pour REVENGE, ces unités sont:

*avenge.v, avenger.n, get back (at).v, get even.v, payback.n, retaliate.v, retaliation.n, retribution.n, retributive.a, retributory.a, revenge.n, revenge.v, revengeful.a, revenger.n, sanction.n, vengeance.n, vengeful.a, vindictive.a*

# Frames dans FrameNet

La représentation du frame comprend des définitions **redondantes** pour tous les éléments core:

## Core frame elements de REVENGE

AVENGER <sub>[SENTIENT]</sub>: The **Avenger** exacts revenge from the **Offender** for the **Injury**.

INJURED\_PARTY <sub>[SENTIENT]</sub>: This frame element identifies the constituent that encodes who or what suffered the **Injury** at the hands of the **Offender**.

INJURY: The **Injury** is the injurious action committed by the **Offender** against the **Injured\_Party**. This Frame Element needs not always to be realized, although it is conceptually necessary.

OFFENDER<sub>[SENTIENT]</sub>: The **Offender** has committed the earlier **Injury** for which the **Avenger** seeks **Revenge**.

PUNISHMENT: The **Avenger** carries out a **Punishment** in order to exact a **Revenge** on the **Offender**.

# Typage sémantique pour les éléments du frame

Au moins une partie des éléments qui composent chaque frame est associée à une catégorie ontologique. Par exemple, dans le REVENGE frame, les étiquettes ontologiques associées aux éléments core et non-core sont:

## Étiquettes ontologiques dans le REVENGE frame

AVENGER	sentient
OFFENDER	sentient
INJURED_PARTY	sentient
INSTRUMENT	physical_entity
PURPOSE	state_of_affairs
...	...

## *FrameNet vs Démonette*

- *FrameNet* est une ressource pour l'anglais, alors que *Démonette* décrit le lexique français
- *FrameNet* caractérise les situations évoquées dans l'esprit du locuteur à travers des mots stockés dans le lexique, alors que *Démonette* représente les relations morphologiques dérivationnelles dans chaque famille
- *FrameNet* est utilisé aussi pour l'annotation des corpus alors que *Démonette* décrit la formation des mots.

## Structures à la *FrameNet* pour *Démonette*

Malgré les différences, les frames pourraient être adaptés pour décrire le sens des relations dérivationnelles dans une ressource comme *Démonette*.

On peut interpréter les éléments d'une famille dérivationnelle comme des éléments d'un frame et les insérer dans une structure à la *FrameNet*

### Familles dérivationnelles structurés comme des frames

*FrameNet*:

An **Avenger** performs a **Punishment** on an **Offender** as a consequence...

*Démonette*:

Un **laveur** **lave** quelque chose dans un **lavoir**...

Des gloses qui impliquent tous les éléments d'un frame comme celle qu'on a vu dans *FrameNet* ne peuvent pas être repliqués . Cependant, des gloses partielles et redondantes comme celles qu'on trouve toujours dans *FrameNet* peuvent l'être...

### Définition redondante de *lavage*

Quand on **lave** quelque chose on fait un **lavage**.

Un **laveur** fait le **lavage** de quelque chose.

On fait le **lavage** de quelque chose dans une **laverie**.

On fait le **lavage** de quelque chose avec une **lavette**.

## Famille partielle de *laver*

laver	nettoyer avec de l'eau
laveur,laveuse	personne qui lave
lavoir,laverie	lieu public ou les gens lave le linge
lavette	morceau de tissu qui sert à nettoyer
lavable	qui peut être lavé
lavement	injection d'un liquide dans le gros intestin
lavage	action ou resultat de laver

Nous proposons de structurer la description sur trois niveaux de représentation: **ontologique**, **relationnel** et **argumental**.

# Niveau ontologique

L'ontologie choisie comme référence pour le typage sémantique est l'ontologie des **unique beginners for nouns** utilisés par *Wordnet*, une base de données lexicale qui groupe noms, verbes et adjectifs dans des synsets (Miller, 1995).

## Unique Beginners for nouns

act, activity	communication	motivation, motive	process
animal, fauna	event, happening	natural object	quantity, amount
artifact	feeling, emotion	natural phenomenon	relation
attribute	food	person, human being	shape
body	group, grouping	plant, flora	state
cognition, knowledge	location	possession	substance
time			

Chaque lexème nominal de la sous-famille considérée est associé avec un de ces 25 unique beginners...

## Types ontologiques pour la sous-famille de *laver*

<b>laver</b>	<b>activity</b>
<b>lavage</b>	activity
<b>laveur, laveuse</b>	human
<b>lavoir, laverie</b>	artifact
<b>lavable</b>	<b>attribute</b>
<b>lavette</b>	artifact
<b>lavement</b>	activity

## Niveau relationnel

L'information sur le niveau relationnel montre le type de relation qui connecte les membres de la famille au moyen de gloses comme celles utilisées dans *FrameNet*. **La condition importante c'est que tous les éléments de la sous-famille considérée soient impliqués dans au moins une relation dérivationnelle, donc dans au moins une glose:**

Un **laveur** **lave** quelque chose.

Une **laveuse** **lave** quelque chose.

Quelque chose est **lavable** si on peut la **laver**.

On **lave** quelque chose dans une **laverie**.

On **lave** quelque chose dans un **lavoir**.

## Niveau relationnel

Un **lavement** **lave** l'intestin.

On réalise le **lavage** de quelque chose avec une **lavette**.

Un **laveur** procède au **lavage** de quelque chose.

Une **laveuse** procède au **lavage** de quelque chose.

On réalise un **lavage** quand on **lave** quelque chose.

On pratique un **lavage** sur quelque chose qui est **lavable**.

# Niveau relationnel

Les gloses peuvent mettre en relation deux, trois, quatre ou plus membres de la famille:

Quelque chose est **lavable** si un **laveur** peut la **laver**.

Quelque chose est **lavable** si une **laveuse** peut la **laver**.

Un **laveur** fait le **lavage** de quelque chose avec une **lavette**.

Une **laveuse** fait le **lavage** de quelque chose avec une **lavette**.

Un **laveur** **lave** quelque chose dans un **lavoir** avec une **lavette**.

Une **laveuse** **lave** quelque chose dans un **lavoir** avec une **lavette**.

# Rôles sémantiques

La représentation fournit aussi le rôle sémantique de chaque membre de la famille. Le rôle du lexème dans la structure argumentale est déduit par la catégorie de relation où il est inscrit:

<b>laver</b>	predicate	verbe de support
<b>lavage</b>	predicate	
<b>laveur, laveuse</b>	agent	
<b>lavoir, laverie</b>	place	
<b>lavette</b>	instrument	
<b>lavable</b>	modifier	

La relation entre *laver* et *lavement* est décrite séparément.

<b>laver</b>	predicate	verbe de support
<b>lavement</b>	procedure	

# Généralisations paradigmatiques (familles)

Les frames morphosémantiques peuvent rendre compte de l'organisation paradigmatique du lexique dérivationnel.

## Familles dérivationnelles partielles pour *observer*, *imprimer* et *nettoyer*

observer	observateur	observatrice	observation	observable	observatoire
imprimer	imprimeur	imprimeuse	impression	imprimable	imprimerie
nettoyer	nettoyeur	nettoyeuse	nettoyage	nettoyable	–

# Généralisations paradigmatiques (relations)

## Agent masculin humain et activité

Un	<b>laveur</b>	<b>lave</b>	quelque chose
Un	<b>nettoyeur</b>	<b>nettoie</b>	quelque chose
Un	<b>observateur</b>	<b>observe</b>	quelque chose
Un	<b>imprimeur</b>	<b>imprime</b>	quelque chose
—	<b>[agent; human ]</b>	<b>[predicate; activity]</b>	—

## Activité et artefact

On	<b>lave</b>	quelque chose dans une	<b>laverie</b> <b>lavoir</b>
On	<b>imprime</b>	quelque chose dans une	<b>imprimerie</b>
On	<b>observe</b>	quelque chose dans un	<b>observatoire</b>
On	<b>nettoie</b>	quelque chose dans un	-
—	<b>[predicate;activity]</b>	—	<b>[place; artifact]</b>

# Généralisations paradigmatiques (relations)

## Attribut et activité

Quelque chose est	<b>lavable</b>	si on peut	la <b>laver</b>
Quelque chose est	<b>imprimable</b>	si on peut	l' <b>imprimer</b>
Quelque chose est	<b>observable</b>	si on peut	l' <b>observer</b>
Quelque chose est	<b>nettoyable</b>	si on peut	la <b>nettoyer</b>
—	[modifier; attribute]	—	[predicate; activity]

## Activité et nom d'action

On fait un	<b>lavage</b>	quand on	<b>lave</b>	quelque chose
On fait une	<b>impression</b>	quand on	<b>imprime</b>	quelque chose
On fait une	<b>observation</b>	quand on	<b>observe</b>	quelque chose
On fait un	<b>nettoyage</b>	quand on	<b>nettoye</b>	quelque chose
—	[predicate; activity]	—	[predicate; activity]	

# Généralisations paradigmatiques (relations)

## Agent humain féminin, activité et artefact

Une	<b>imprimeuse</b>	<b>imprime</b>	quelque chose dans une	<b>imprimerie</b>
Une	<b>observatrice</b>	<b>observe</b>	quelque chose dans un	<b>observatoire</b>
Une	<b>laveuse</b>	<b>lave</b>	quelque chose dans une	<b>laverie</b>
Une	<b>nettoyeuse</b>	<b>nettoie</b>	quelque chose dans une	-
—	<b>[agent; human ]</b>	<b>[predicate; activity]</b>	—	<b>[place; artifact]</b>

## Potentialité, agent humain et activité

Quelque chose est	<b>imprimable</b>	si un	<b>imprimeur</b>	peut l'	<b>imprimer</b>
Quelque chose est	<b>observable</b>	si un	<b>observateur</b>	peut l'	<b>observer</b>
Quelque chose est	<b>lavable</b>	si un	<b>laveur</b>	peut la	<b>laver</b>
Quelque chose est	<b>nettoyable</b>	si un	<b>nettoyeur</b>	peut la	<b>nettoyer</b>
—	<b>[modifier; potentiality]</b>	—	<b>[agent; human m.]</b>	—	<b>[predicate; activity]</b>

Considérons la sous-famille construite autour du nom *science*:

**science** <sub>n.</sub>, **scientifique** <sub>n.</sub>, **scientifique** <sub>a.</sub>, **scientifiser** <sub>v.</sub>, **scientificité** <sub>n.</sub> ...

### Étiquettes ontologiques pour *science*

science	knowledge
scientifique	human
scientifique	attribute
scientifiser	activity
scientificité	attribute

Un **scientifique** est un spécialiste de la **science**

Une chose **scientifique** est pertinente pour la **science**

On **scientifise** quelque chose quand on l'analyse du point de vue de la **science**

Un **scientifique** **scientifise** quelque chose

Un **scientifique** analyse quelque chose d'une manière **scientifique**

Un **scientifique** **scientifise** quelque chose en l'analysant du point de vue de la **science**

### Rôles argumentals pour *science*

science	domain
scientifique	agent
scientifique	modifier
scientifiser	predicate
scientificité	property

# Généralisations paradigmatiques

## Familles dérivationnelles pour *biologie*, *économie*, *mathématique*

biologie	biologiste	biologique	biologiser	—
économie	économiste	économique	économiser	économicité
mathématique	mathématicien	mathématique	mathématiser	mathématicité

## Agent humain et domaine

Une chose	<b>scientifique</b>	est pertinente à la	<b>science</b>
Une chose	<b>biologique</b>	est pertinente à la	<b>biologie</b>
Une chose	<b>économique</b>	est pertinente à l'	<b>économie</b>
Une chose	<b>mathématique</b>	est pertinente à la	<b>mathématique</b>
—	[attribute;modifier]	—	[knowledge;domain]

## Agent humain, activité et domaine

Un	<b>scientifique</b>	<b>scientifise</b>	qqc en l'analysant d'un point de vue	<b>scientifique</b>
Un	<b>biologue</b>	<b>biologise</b>	qqc en l'analysant d'un point de vue	<b>biologique</b>
Un	<b>économiste</b>	<b>économise</b>	qqc en l'analysant d'un point de vue	<b>économique</b>
Un	<b>mathématicien</b>	<b>mathématise</b>	qqc en l'analysant d'un point de vue	<b>mathématique</b>
—	[human;agent]	[activity;predicate]	—	[attribute;modifier]

Considérons la sous-famille construite autour de l'adjectif *riche*:

**riche** <sub>a.</sub>, **riche** <sub>n.</sub>, **richesse** <sub>n.</sub>, **enrichir** <sub>v.</sub>, **enrichissement** <sub>n.</sub>

### Étiquettes ontologiques pour *riche*

riche	attribute
riche	human
richesse	attribute
enrichir	activity
enrichissement	activity

Un **riche** est une personne **riche**

La **richesse** est la qualité de qui est **riche**

On **enrichit** quelq'un ou quelque chose en le rendant **riche**

On fait un **enrichissement** quand on **enrichit** quelque chose

On fait un **enrichissement** quand on rend quelque chose plus **riche**

On fait un **enrichissement** quand on **enrichit** quelque chose en contribuant à sa **richesse**

## Rôles argumentals pour *riche*

riche	modifier
riche	agent
richesse	property
enrichir	predicate
enrichissement	predicate

# Généralisations paradigmatiques

## Familles dérivationnelles pour *pauvre*, *blanc*, *simple*

pauvre	pauvre	pauvreté	appauvrir	appauvrissement
beau	beau	beauté	embellir	embellissement
simple	—	simplicité	simplifier	simplification

## Agent humain et domaine

La	<b>richesse</b>	est la qualité de qui est	<b>riche</b>
La	<b>pauvreté</b>	est la qualité de qui est	<b>pauvre</b>
La	<b>beauté</b>	est la qualité de qui est	<b>beau</b>
La	<b>simplicité</b>	est la qualité de qui est	<b>simple</b>
—	[property;modifier]	—	[attribute;modifier]

## Agent humain, activité et domaine

On fait un	<b>enrichissement</b>	quand on	<b>enrichit</b>	qqc/qqn en contribuant à sa	<b>richesse</b>
On fait un	<b>appauvrissement</b>	quand on	<b>appauvre</b>	qqc/qqn en contribuant à sa	<b>pauvreté</b>
On fait un	<b>embellissement</b>	quand on	<b>embellit</b>	qqc/qqn en contribuant à sa	<b>beauté</b>
On fait une	<b>simplification</b>	quand on	<b>simplifie</b>	qqc en contribuant à sa	<b>simplicité</b>
—	[predicate;activity]	—	[predicate;activity]	—	[property;attribute]

# Automatiser la construction des frames morphosémantiques (en cours)

## Quelles données utiliser?

- Données lexicographiques
  - extraire des regularités par les définitions
  - extraire des autres informations utiles à alimenter les trois niveau de représentation
- Corpus
  - grande quantité de données
  - pourraient être utiles pour extraire des voisins distributionnels ou pour extraire des informations sur la structure argumentale

Informations présents dans les définitions:

## Définitions extraites par GLAWI (Sajous and Hathout, 2015)

**BIOLOGISTE:** **Personne** qui étudie la **biologie**, c'est-à-dire les **organismes vivants** ainsi que leurs constituants et leurs interactions avec leur environnement

**LAVERIE:** **Endroit** où l'on **lave** industriellement le **linge**

**OBSERVATOIRE:** **Lieu** d'où l'on peut **observer**, surveiller **l'ennemi**, d'un **endroit** aménagé à cet effet

**LAVEMENT:** **Liquide** que l'on introduit par l'**anus** pour **laver** les intestins, les **soigner** ou pour procéder à des examens radiologiques ou visuels. On dit plutôt aujourd'hui '**lavage** d'intestin', et il peut être alors remplacé par une solution à boire la veille d'un examen comme la coloscopie par exemple

Etymologie, domaine d'utilisation et la classe syntaxique:

## Informations de domaine, étymologiques, et classes syntaxiques extraites par GLAWI

BIOLOGISTE	<b>biologie</b>	composé de <b>biologie</b> et -iste.	N
LAVERIE	vêtements	voir <b>laver</b> et -erie	N
OBSERVATOIRE	militaire	dérivé savant sur le radical de <b>observateur</b> et -oire	N
LAVEMENT	médecine	voir <b>laver</b> et -ment	N

Voisins distributionnels obtenus avec un modèle entraîné sur FrWac  
(Baroni et al., 2009)

SCIENCE	LAVER	RICHE
philosophie: 0.716	blanchisserie: 0.747	pauvre: 0.658
sociologie: 0.696	eicerie: 0.703	<b>richesse</b> : 0.634
<b>scientifique</b> : 0.684	buanderie: 0.673	fertile: 0.627
anthropologie: 0.681	<b>lave-linge</b> : 0.644	modeste: 0.618
biologie: 0.640	self-service: 0.642	abondant: 0.587
cognition: 0.633	blanchissage: 0.626	attrayant: 0.575
ethnologie: 0.628	toilette: 0.614	opulent: 0.564
neuro-physiologie: 0.623	vaisselle: 0.608	florissant: 0.562
<b>neuro-sciences</b> : 0.608	seche-linge: 0.601	fabuleusement: 0.548
astronomie: 0.597	<b>lave-vaisselle</b> : 0.578	dense: 0.548

## Conclusions

- Les structures utilisées pour les frames dans *FrameNet* peuvent être efficacement adaptées pour représenter les relations entre des lexèmes de la même famille
- L'empilage de plusieurs sous-familles sous les mêmes structures nous permet de représenter la nature paradigmatique du lexique

## Pistes à suivre

- Étendre l'ontologie (notamment pour verbes et adjectifs)
- Approfondir dans le détail la question de la génération automatique des frames dérivationnels
- Tester la possibilité d'utiliser des autres données à côté des dictionnaires (structures syntaxiques, word embeddings)
- Trouver une manière efficace d'implémenter physiquement les frames morphosémantiques *Démonette*

Merci à votre attention!!!

- Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). The wacky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3):209–226.
- Bonami, O. and Strnadová, J. (2018). Paradigm structure and predictability in derivational morphology. *Morphology*, pages 1–31.
- Fillmore, C. J. et al. (1976). Frame semantics and the nature of language. In *Annals of the New York Academy of Sciences: Conference on the origin and development of language and speech*, volume 280, pages 20–32.
- Fradin, B. (2003). *Nouvelles approches en morphologie*. Presses universitaires de France.
- Hathout, N. (2011). Morphonette: a paradigm-based morphological network. *Lingue e linguaggio*, 10(2):245–264.

- Hathout, N. and Namer, F. (2014). Démonette, a french derivational morpho-semantic network. *LiLT (Linguistic Issues in Language Technology)*, 11.
- Hathout, N. and Namer, F. (2016). Giving lexical resources a second life: Démonette, a multi-sourced morpho-semantic network for french. In *Language Ressources and Evaluation Conference*.
- Hathout, N., Namer, F., and Lignon, S. (2017). Adding morpho-phonological features to a french morphosemantic resource: the demonette derivational database. In *Workshop on Resources and Tools for Derivational Morphology (DeriMo)*.
- Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

- Namer, F. and Hathout, N. (2019). Paradis and démonette: from theory to resources for derivational paradigms. In *In Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology*.
- Ruppenhofer, J., Ellsworth, M., Schwarzer-Petruck, M., Johnson, C. R., and Scheffczyk, J. (2006). Framenet ii: Extended theory and practice.
- Sajous, F. and Hathout, N. (2015). Glawi, a free xml-encoded machine-readable dictionary built from the french wiktionary.