Semantic descriptions of French derivational families in a families-and-paradigms framework

Daniele Sanacore [1] Nabil Hathout [1] Fiammetta Namer [2]

[1]CLLE, CNRS & University of Toulouse
[2]ATILF, CNRS & University of Lorraine

DeriMo 2019, September 19 - 20

Introduction

Objective

In the context of paradigmatic derivational morphology, the objective is to **describe morphosemantic relations** contained in the French lexicon. This description, contrary to what most morphological resources provide, must concern complete or partial **derivational families** rather than couples of lexemes.

Approach

We will show how structures inspired by frames in **Frame Semantics** could be used to represent derivational families and paradigms.

Theoretical background

- modern derivational resources have been characterised by the adoption of the lexeme as minimal unit
- paradigmatic structure of the derivational lexicon, with derivational families as central elements
- double function of derivational constructions: create new lexemes and establish semantic and formal relations of motivation between them. (Hathout and Namer, 2019)

Why Frame Semantics?

Frame Semantics represents conceptual situations in objects called **frames**. These frames contain many **participants** involved in the situation they represent.

Even though differences between the two tasks need to be considered, what we are seeking is a **representation of derivational relations within a derivational family, as if its lexemes were frame elements in a frame**.

Definitions

Derivational families

A derivational family is a set of lexemes connected by morphological derivational relations (Hathout, 2011). An example of derivational family for French is the partial family built around the verb *laver* 'to wash' in (1):

(1) **laver** 'to wash'; **lavage** 'washing'; **lavoir** 'wash house'; **laverie** 'laundromat'; **laveur** 'washer (male)'; **laveuse** 'washer (female)'; **lavette** 'dishcloth'; **lavable** 'washable'; **lavement** 'enema'...

Definitions

Direct and indirect derivational relations

The derivational relations between lexemes in a family may either be direct or indirect. A **direct derivational relation** connects a lexeme with one of its direct ascendants or descendants (2), while an **indirect derivational relation** connects more distant elements of the family (3).

(2) laver$_v$. $\rightarrow$ laveur$_n$.

(3) laveur$_n$. $\rightarrow$ lavage$_n$.

Definitions
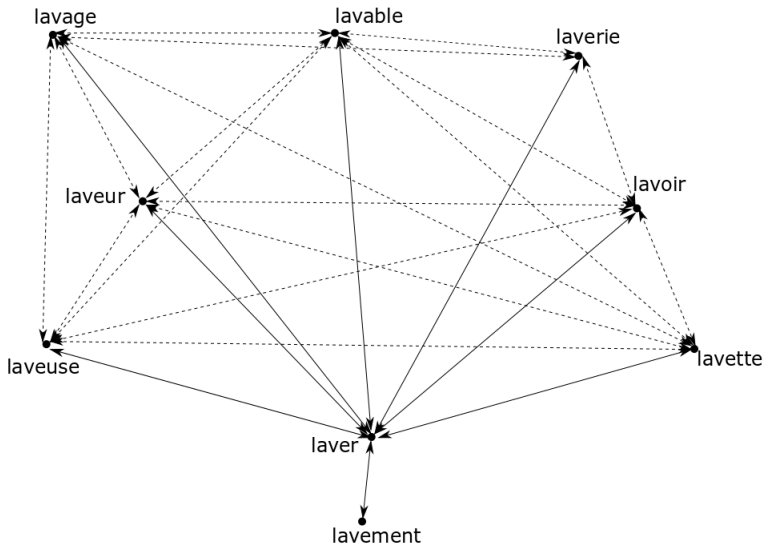
Paradigmatic systems and derivational series

A **paradigmatic system** is a collection of (partial) derivational families that are aligned in terms of the content-based relations that their members entertain (Bonami and Strnadová, 2018). The CONTENT is the specification of syntactic/semantic properties of a word, while the FORM is the specification of its phonology and or orthography. Aligned derivational relations in a paradigm form a **derivational series**. Let us take four subfamilies built around the verbs *imprimer* 'to print', *souder* 'to weld', *laver* 'to wash' and nettoyer 'to clean'.

verb	agent_m	adj	action noun
imprimer	imprimeur	imprimable	impression
souder	soudeur	soudable	soudage
laver	laveur	lavable	lavage
nettoyer	nettoyeur	nettoyable	nettoyage
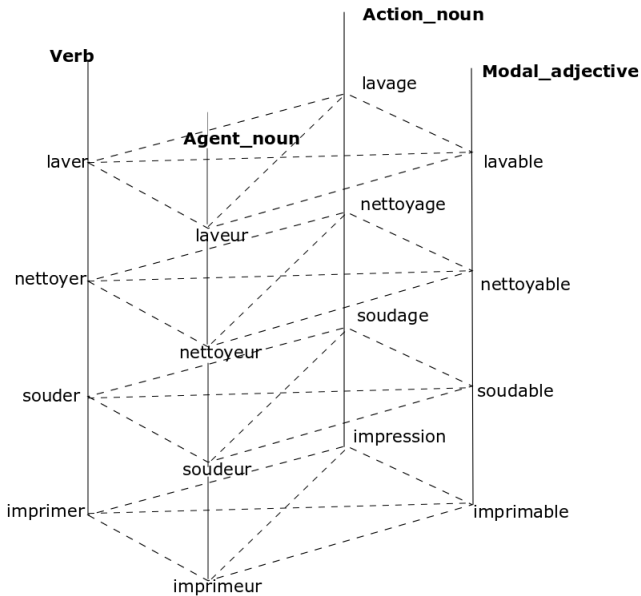
Démonette

Démonette (Hathout et al., 2017; Hathout and Namer, 2014) is a lexical resource designed for the description of word formation in French. It is based on the fundamental assumption that morphology is relational and each relation where a given lexeme is involved contributes to its meaning.

- *Démonette* seeks a complete, redundant and explicit description of all the properties of a derivational relation

- entries in *Démonette* do not describe the properties of the derivatives, they describe properties of the relations connecting two lexemes

- *Démonette* is a directed graph where a relation ($w_1 \leftarrow w_2$) describes the morphological motivation of w_1 with respect to w_2. Most of the lexemes are connected with each other in both directions (Hathout and Namer, 2016).

In its current state, Démonette does not provide what we are looking for.

Semantic description in *Démonette*

There are four fields used for semantic description of derivational relations in *Démonette*, two fields for the semantic type of w_1 and w_2, one for the concrete definition and one for the abstract definition, where w_2 is replaced by its semantic type.

Semantic types, concrete and abstract definitions in 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 @"

Different levels of information

The current semantic representations in *Démonette* is characterised by a description of the relation provided by the concrete and abstract definition and a description of the semantic role and the ontological type.

However, the ontological category and the semantic role are merged in an unique label (e.g. @AGF for *laveuse, nettoyeuse* and *imprimeuse*). Since the ontological category associated to a lexeme is independent from the semantic role, two separated labels should be provided.

e.g. With the current semantic typing system, an instrument noun like *aspirateur* 'vacuum cleaner' and an human agent noun like *observateur* 'observer' are typed with the same label @AGM.

Frame Semantics

Frame Semantics is based on the fundamental assumption that people understand language by means of situations evoked in their mind by words. These representations of real world situations evoked in our mind are called **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

The implementation of Frame Semantics is **FrameNet** (Ruppenhofer et al., 2006), a lexical resource for English. FrameNet relates words to their meanings via the frames they instantiate and records the way in which sentences and phrases are structured around them. The main objectives of FrameNet are:

- characterize frames and find the lexical units that evoke them
- develop a descriptive terminology for each frame
- extract sample sentences for each frame.

Frames in FrameNet

Frames represent story fragments characterised by a given number of participant involved in it. These elements are called **frame elements**. For instance, let us take a look at the *revenge* frame:

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.

Sentences instantiating the *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.

Frames in FrameNet

FrameNet defines two types of frame:

Core frame elements: elements that are essential for the understanding of the frame (e.g. AVENGER, PUNISHMENT, OFFENDER, INJURY, INJURED_PARTY for the REVENGE frame)

Non-core frame elements: elements that are more generally applicable across frames (e.g. DEGREE, DURATION, MANNER, PLACE, TIME for the REVENGE frame)

Moreover, each frame is associated with a given set of **lexical units** that evoke it. For the REVENGE frame, these lexical units are:

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 in FrameNet

The frame representation also provides partial sentences describing the role of each frame element, providing thus **redundant definitions** for each one of them:

Core frame elements of the REVENGE frame

AVENGER: The **Avenger** exacts revenge from the **Offender** for the **Injury**.

INJURED_PARTY: 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: 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**.

Semantic typing for frame elements

At least a part of the frame elements composing each frame is associated to an ontological category, for the REVENGE frame, the ontological labels associated to its core and non-core frame elements are:

Ontological labels in the REVENGE frame

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

Differences between *FrameNet* and *Démonette*

FrameNet vs *Démonette*

- *FrameNet* is a lexical resource for English, while *Démonette* describes the French lexicon
- *FrameNet* aims to characterise situations evoked in our mind by words in the lexicon; while *Démonette* seeks a paradigmatic representation of morphosemantic relations between lexemes
- *FrameNet* is also used to annotate corpora, while *Démonette* is not.

Frame-like structures for *Démonette*

Even though the differences between *Démonette* and FrameNet need to be considered, frames could be adapted to improve the morphosemantic description of derivational relations in resources like *Démonette*.

We can interpret the elements of a derivational family like frame elements in a frame and insert them in a frame-like structure. In a second moment, we can find other derivational families that fit the same structure in order to highlight paradigmatic regularities.

Derivational families structured like frames

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

Démonette:
Un **laveur lave** quelque chose dans un **lavoir**...
'A washer washes something in a wash house...'

Frame-like structures for *Démonette*

Global definitions of frames in *FrameNet* would be too complicate to be adapted for *Démonette*, however, partial and redundant frame definitions like those of *can* be used for our task...

Redundant definition of *lavage*

Quand on **lave** quelque chose on fait un **lavage**.
'When we wash something we do a washing.'

Un **laveur** fait le **lavage** de quelque chose.
'A washer does the washing of something.'

On fait le **lavage** de quelque chose dans une **laverie**.
'We do the washing of something in a laundromat.'

On fait le **lavage** de quelque chose avec une **lavette**.
'We do the washing of something with a dishcloth.'

Frame-like structures for *Démonette*

The first subfamily that we represent with a frame-like structure is the family of *laver*:

laver	to wash
laveur,laveuse	person who washes
lavoir,laverie	public place where people do the laundry
lavette	hard sponge use for washing
lavable	able to be washed
lavement	procedure or medicinal product for intestinal washing
lavage	action or result of the action of washing

The description must be structured on three levels of semantic representation: **ontological**, **relational** and **argumental**.

Ontological level

The reference ontology chosen for the semantic typing are the **unique beginners for nouns** used by Wordnet, an English lexical database that groups nouns, verbs and adjectives into sets of cognitive synonyms (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			

Each lexeme of the considered subfamily is associated with one of these unique beginners...

Semantic types for the *laver* subfamily

laver	activity
lavage	activity
laveur, **laveuse**	human
lavoir, **laverie**	artifact
lavable	attribute
lavette	artifact
lavement	activity

Relational level

The information on the relational level shows the type of relation that connects the elements of the family by means of sentences like those used by *FrameNet*. The important condition is that all the elements of the considered subfamily must be involved in at least one derivational relation, so in at least one sentence:

Un **laveur lave** quelque chose.
Une **laveuse lave** quelque chose.
' A washer washes something.'

Quelque chose est **lavable** si on peut la **laver**.
'Something is washable if it can be washed'.

On **lave** quelque chose dans une **laverie**.
'Something is washed in a laundromat.'
On **lave** quelque chose dans un **lavoir**.
'Something is washed in a wash house,'

Relational level

Un **lavement lave** l'intestin.
'An enema washes the intestine'.

On realise le **lavage** de quelque chose avec une **lavette**.
'We do the washing of something with a dishcloth'

Un **laveur** procède au **lavage** de quelque chose.
Une **laveuse** procède au **lavage** de quelque chose.
'A washer does the washing of something.'

On réalise un **lavage** quand on **lave** quelque chose.
'A washing is realised when we wash something.'

On pratique un **lavage** sur quelque chose qui est **lavable**.
'The washing is done on something that can be washed.'

Relational level

Once binary relations have been established for each element of the considered subset, we procede by creating sentences with three family elements, thus establishing relations between three lexemes:

Quelque chose est **lavable** si un **laveur** peut la **laver**.
Quelque chose est **lavable** si une **laveuse** peut la **laver**.
'Something is washable if a washer can wash it.'

Un **laveur** fait le **lavage** de quelque chose avec une **lavette**.
Une **laveuse** fait le **lavage** de quelque chose avec une **lavette**.
'A washer does the washing of something with a dishcloth.'

Un **laveur lave** quelque chose dans un **lavoir** avec une **lavette**.
Une **laveuse lave** quelque chose dans un **lavoir** avec une **lavette**.
'A washer washes something in a wash house with a dishcloth.'

Semantic roles

Finally, the representation provides also the semantic roles for each element of the subfamily. The role in the argumental structure is deduced from the category of relation in which they are inscribed:

laver	predicate	
lavage	predicate	preceded by light verb
laveur, **laveuse**	agent	
lavoir, **laverie**	place	
lavette	instrument	
lavable	modifier	

In this case, the relation between *laver* and *lavement* should be considered in a separate structure.

laver	predicate	
lavement	predicate	preceded by light verb

Catching paradigmatic generalizations

The frame-like representation is also fit to show the paradigmatic organization of the derivational lexicon. Let us test the structure we constructed for *laver* on three other subfamilies: *observer* 'to observe', *imprimer* 'to print' and *nettoyer* 'to clean':

observer 'to observe'	observateur , 'observer(m.)'	observatrice 'observer' (f.)	observation , 'observation'	observable 'observable'	observatoire 'observatory'
imprimer 'to print'	imprimeur 'printer'(m.)	imprimeuse 'printer' (f.)	impression 'printing'	imprimable 'printable'	imprimerie 'copy shop'
nettoyer 'to clean'	nettoyeur 'cleaner' (m.)	nettoyeuse 'cleaner' (f.)	nettoyage 'cleaning'	nettoyable 'cleanable'	–

Table: Partial derivational families of *observer*, *imprimer* and *nettoyer*

Catching paradigmatic generalizations

Un	**laveur**	**lave**	quelque chose
Un	**nettoyeur**	**nettoie**	quelque chose
Un	**observateur**	**observe**	quelque chose
Un	**imprimeur**	**imprime**	quelque chose
—	**[agent; human]**	**[predicate; activity]**	—

Table: Masculine human agent and activity

On	**lave**	quelque chose dans une	**laverie** **lavoir**
On	**imprime**	quelque chose dans une	**imprimerie**
On	**observe**	quelque chose dans un	**observatoire**
On	**nettoye**	quelque chose dans un	-
—	**[predicate;activity]**	—	**[place; artifact]**

Table: Activity and artifact

Catching paradigmatic generalizations

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

Table: Attribute and activity

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

Table: Feminine human agent, activity and artifact

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

Table: Potentiality, human agent and activity

Conclusions and next steps

- Semantic frames can be easily adapted to represent derivational relations in a paradigmatic morphology framework
- However, a larger and more heterogeneous number of derivational families needs to be analysed
- Next steps will involve researches on how to perform the automatic generation of derivational frames like the ones we showed
- Researches need to be done to find the best type of linguistic data for the generation of derivational frames (lexicographic data, corpora, etc.).

Thank you for your attention!!!

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.

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. and Namer, F. (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*.

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.

Ruppenhofer, J., Ellsworth, M., Schwarzer-Petruck, M., Johnson, C. R., and Scheffczyk, J. (2006). Framenet ii: Extended theory and practice.

Polysemy in Frame Semantics

Typically, each sense of a polysemous word belongs to a different frame (Ruppenhofer et al., 2006). For instance, the lemma *bake* evokes three different frames:

- 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.