

A story-based approach to derivational paradigms

Daniele Sanacore ¹ Nabil Hathout ¹ Fiammetta Namer ²

¹Université de Toulouse, CLLE

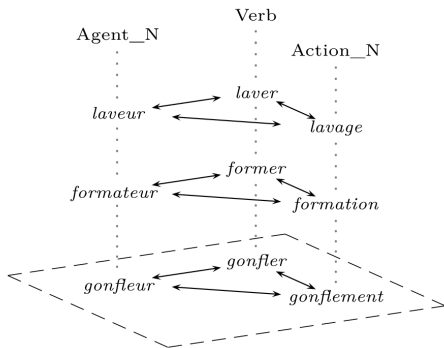
²Université de Lorraine, ATILF

September 13, ISMO 2023
ATILF, Nancy

Data and objective

For Bonami and Strnadová (2019), derivational paradigms are **alignments** of “slices” of derivational families that contain the same semantic relations. Hathout and Namer (2022) call these slices **paradigmatic families**.

Our objective is to **extract paradigmatic families** from derivational families. In **paradigmatic families, lexemes are all interpredictable**, as if they were part of a sort of “bundle”. **Relations in paradigmatic families are recurrent**.



- **Problem₁** : we don't know the boundaries of paradigmatic families
- **Problem₂** : we don't know how to align paradigmatic families

Multiple relations need to be described at the same time

F1= {POT 'pot', POTIER 'potter', POTERIE 'pottery', REMPOTER 'repot', REMPOTAGE 'repotting'}

Paradigmatic families :

- {POT, POTIER, POTERIE}
- {POT, REMPOTER, REMPOTAGE}

Our proposal : use stories to characterize lexemes contained in derivational families. Stories describe scenarios where participants are linked by relations. Their function is to make the relations between lexemes in paradigms explicit.

Story example :

Un jour, un potier nommé Marc créa un magnifique pot en argile dans son atelier.
'One day, a potter named Marc created a magnificent clay pot in his workshop.'

Stories to describe morphosemantic relations

All the possible subsets are considered :

$$\text{cov}(F1) = \{ \{ \text{POT}, \text{POTERIE} \}, \{ \text{POT}, \text{POTIER} \}, \dots, \{ \text{REMPOTER}, \text{REMPOTAGE} \}, \\ \{ \text{POT}, \text{POTIER}, \text{POTERIE} \}, \dots, \{ \text{POT}, \text{POTERIE}, \text{POTIER}, \text{REMPOTER}, \text{REMPOTAGE} \} \}$$

Hypothesis₁

Derivational paradigms are structured by recurrent semantic relations (Antoniová and Štekauer, 2016; Bauer, 2019; Fernández-Domínguez et al., 2020).

Concepts denoted by lexemes of the same paradigmatic family will **frequently** co-occur in stories.

Concepts denoted by lexemes that do not belong to the same paradigmatic family will **rarely** co-occur in stories.

- Since paradigmatic relations are recurrent relations, we need plenty of stories ;
- Stories are used as heuristic tools.

Obtain stories for morphosemantic description

Ideally, stories could be obtained by interviewing native speakers and by asking them to tell us a story that instantiates a set of lexemes that belong to the same family. However, we do not have the means to carry out such a large scale data collection. For this reason, we created some stories to illustrate the methodology that we propose.

This implies two approximations :

- We consider that speakers characterize lexical relations by means of stories ;
- Since we cannot access stories produced by real speakers, we produce them ourselves.

Obtain stories for morphosemantic description

Generative models such as *ChatGPT* are capable of generating stories that meet our needs.

Prompt : « Je veux une histoire de 50 mots qui contient les mots “pot” et “potier” »

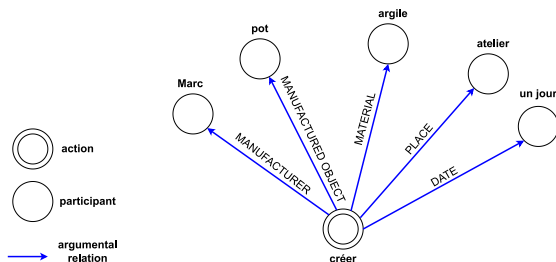
Il était une fois un potier nommé Marc, qui aimait créer des pots en argile dans son atelier. Un jour, il créa un magnifique pot, parfaitement équilibré et décoré avec des motifs complexes. Ce pot devint très populaire parmi les clients de Marc, qui ne tarissaient pas d'éloges sur son travail. Et c'est ainsi que Marc devint célèbre pour la qualité de son travail, en particulier pour ses pots exquis.

'Once upon a time, there was a potter named Marc, who loved to create clay pots in his workshop. One day, he created a magnificent pot, perfectly balanced and decorated with intricate designs. The pot became very popular with Marc's customers, who raved about his work. And so it was that Marc became famous for the quality of his work, especially his exquisite pots.'

Representing the semantic content of stories

Events, participants and relations in stories are represented by means of semantic networks (SN).

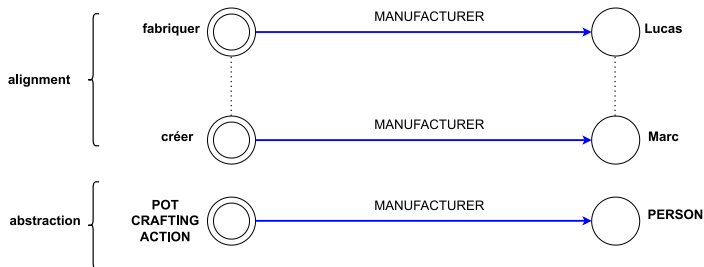
ex. *Un jour, un potier nommé Marc créa un magnifique pot en argile*
'One day, a potter named Mark created a magnificent clay pot'.



- Representation of events participants and relations;
- Participants are connected to events via argumental relations (Parsons, 1995; Higginbotham, 2000);
- SN contain “fine-grained” argumental roles.

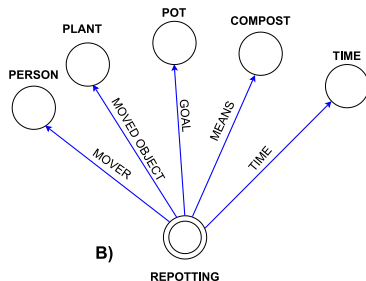
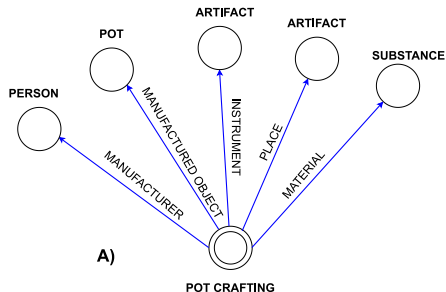
Aligning relations in stories

- Relations contained in different stories must be aligned on the basis of their type and the ontological class of their participants ;
- relations are abstracted using ontological categories.



Aligned relations are abstracted into semantic networks called **meaning bundles (MB)**. A meaning bundle is a semantic object that contains concepts and relations that regularly co-occur in stories. We represent MB as semantic networks.

Meaning bundles



(1) {POT, POTIER, POTERIE, REMPOTER, REMPOTAGE}

We use MB to slice derivational families and obtain paradigmatic families.

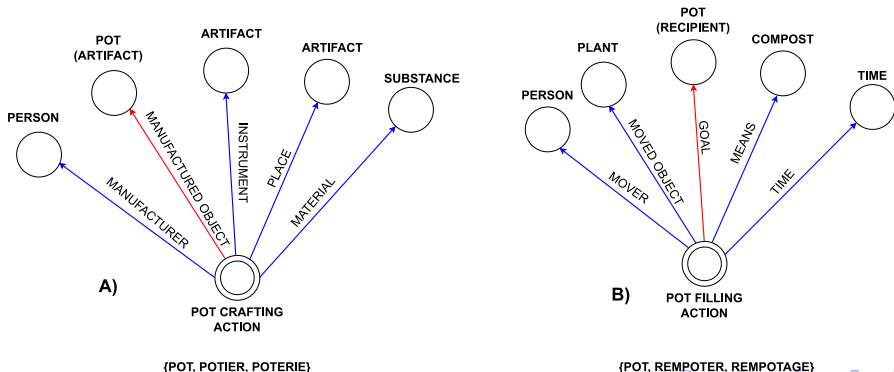
Two MB are obtained from the family in (1) : **a)** one describing pottery **b)** one describing repotting. These two MB are projected onto the derivational lexicon and produce the paradigmatic families in (2)

- (2) a. {POT, POTIER, POTERIE}
b. {POT, REMPOTER, REMPOTAGE}

Overlapping derivational paradigms

When they overlap, derivational paradigms select distinct semantic properties of the same lexeme.

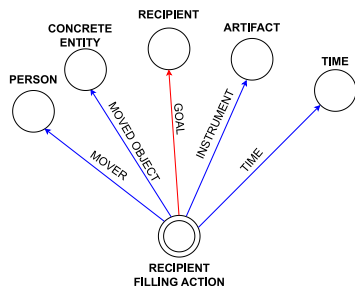
- For the paradigmatic family {POT, POTIER, POTERIE} the ontological category ARTIFACT is relevant ;
- For the paradigmatic family {POT, REMPOTER, REMPOTAGE} the RECIPIENT category is relevant.



Variety of derivational paradigms

Problem₃ : identify the relevant semantic properties that determine derivational paradigms

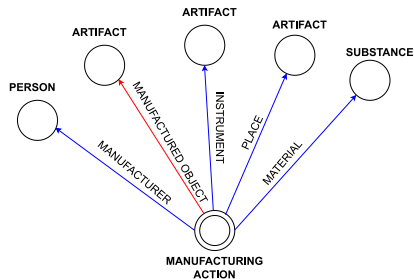
Relevant semantic properties are “revealed” by the aligned families :
RECIPIENT



GOAL RECIPIENT	ACTION.V	ACTION.N	INSTRUMENT ARTIFACT	MOVER PERSON
POT 'pot'	REMPOTER 'to repot '	REMPOTAGE 'repotting'	-	REMPOTEUR 'repotter'
CARAFE 'carafe'	CARAFER 'to decant'	CARAFAGE 'decanting'	-	-
TONNEAU 'barrel'	ENTONNER 'to barrel'	ENTONNAGE 'barreling'	ENTONNOIR 'funnel'	-
CAISSE 'crate'	ENCAISSER 'to crate'	ENCAISSEMENT 'crating'	-	ENCAISSEUR 'packer'
BOUTEILLE 'bottle'	EMBOUTEILLER 'to bottle'	EMBOUTEILLAGE 'bottling'	EMBOTEILLEUSE 'bottling machine'	EMBOUTEILLEUR 'bottler'
...

Variety of derivational paradigms

Relevant semantic properties are “revealed” by the aligned families :
ARTIFACT

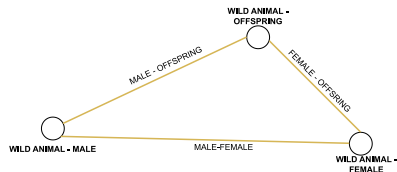


MANUF.OBJECT ARTIFACT	MANUFACTURER PERSON	MANUF.PLACE ARTIFACT	ACTIVITY.N
POT 'pot'	POTIER 'potter'	-	POTERIE 'pottery'
TUILE 'tile'	TUILIER 'tile maker'	TUILERIE 'tile factory'	-
BRIQUE 'brick'	BRIQUETIER 'brickmaker'	BRIQUETERIE 'brickyard'	BRIQUETERIE 'brickery'
BOTTE 'boot'	BOTTIER 'bootmaker'	BOTTERIE 'boot factory'	BOTTERIE 'boot making'
...

Variety of derivational paradigms

Some paradigms are not structured by argumental relations nor by events : families based on nouns denoting wild animals contain sort of inner “social” relations for a given animal species.

F2= { LOUP 'wolf', LOUVE 'she-wolf', LOUVETEAU 'wolf cub', LOUVETIER 'wolf hunter', LOUVETERIE 'wolf hunting' }

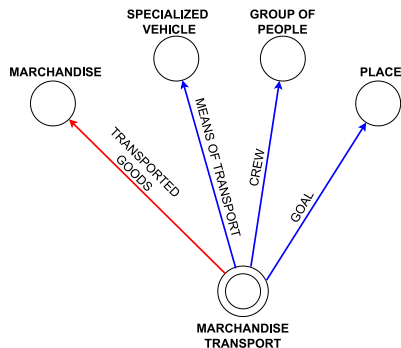


WILD ANIMAL-MALE	WILD ANIMAL-FEMALE	WILD ANIMAL-OFFSPRING
LOUP 'wolf'	LOUVE 'she-wolf'	LOUVETEAU 'wolf cub'
LION 'lion'	LIONNE 'lioness'	LIONCEAU 'lion cub'
ÉLÉPHANT 'male elephant'	ÉLÉPHANTE 'female elephant'	ÉLÉPHANTEAU 'elephant calf'
OURS 'male bear'	OURSE 'female bear'	OURSON 'bear cub'
HYÈNE 'male hyena'	HYÈNE 'female hyena'	HYÈNE 'hyena cub'
...

Variety of derivational paradigms

The same formal operation can select distinct semantic properties :
MARCHANDISE

F3= {BANANE 'banana', BANANIER₁ 'banana tree', BANANIER₂ 'banana cargo',
BANANIER₃ 'banana cake', BANANERAIE 'banana plantation'}

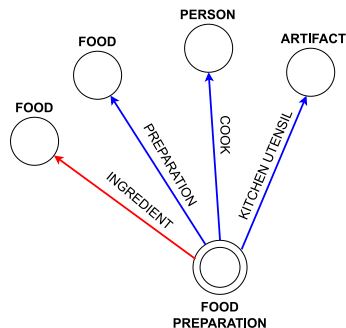


TRANSPORTED GOODS MARCHANDISE	MEANS OF TRANSPORT SPECIALIZED VEHICLE
BANANE 'banana'	BANANIER 'banana cargo'
PÉTROLE 'petroleum'	PÉTROLIER 'oil tanker'
GAZ 'gas'	GAZIER 'gas carrier'
ASPHALTE 'asphalt'	ASPHALTIER 'asphalt carrier'
...	...

Variety of derivational paradigms

The same formal operation can select distinct semantic properties :
INGREDIENT

{**BANANE 'banana'**, BANANIER₁ 'banana tree', BANANIER₂ 'BANANA CARGO',
BANANIER₃ 'banana cake', BANANERAIE 'banana plantation'}

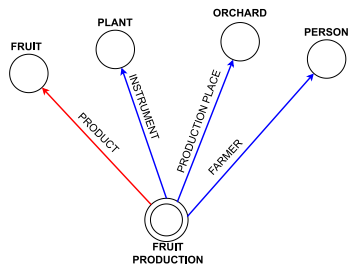


INGREDIENT FOOD	PREPARATION FOOD
BANANE 'banana'	BANANIER 'banana cake'
FRAMBOISE 'raspberry'	FRAMBOISIER 'raspberry cake'
FRAISE 'strawberry'	FRASIER 'strawberry cake'
CITRON 'lemon'	CITRONNIER 'lemon cake'
...	...

Variety of derivational paradigms

The same formal operation can select distinct semantic properties : FRUIT

{BANANE 'banana', BANANIER₁ 'banana tree', BANANIER₂ 'banana cargo', BANANIER₃ 'banana cake', BANANERAIE 'banana plantation'}



PRODUCT FRUIT	INSTRUMENT PLANT	ORCHARD PRODUCTION PLACE	FARMER PERSON
BANANE 'banana'	BANANIER 'banana tree'	BANANERAIE 'banana plantation'	-
PRUNE 'prune'	PRUNIER 'prune tree'	PRUNERAIE 'prune orchard'	-
POMME 'apple'	POMMIER 'apple tree'	POMMERAIE 'apple orchard'	-
MELON 'melon'	MELON 'melon plant'	MELONNIÈRE 'melon field'	MELONNIER 'melon farmer'
...

Conclusions

- Stories are used to identify the semantic relations that structure derivational paradigms ;
- Derivational paradigms need to be identified from data ;
- Derivational paradigms overlap ;
- Derivational paradigms are structured by relations of various kinds (ex. argumental relations vs lifecycle) ;
- Derivational relations paradigms may select distinct semantic properties of the same lexeme (ex. 'manufactured object' vs 'recipient', 'fruit' vs 'merchandise' vs 'ingredient').

Thank you for your attention !!!

- Antoniova, V. and Štekauer, P. (2016). Derivational paradigms within selected conceptual fields—contrastive research. *Facta Universitatis, Series : Linguistics and Literature*, pages 61–75.
- Bauer, L. (2019). Notions of paradigm and their value in word-formation. *Word Structure*, 12(2) :153–175.
- Bonami, O. and Strnadová, J. (2019). Paradigm structure and predictability in derivational morphology. *Morphology*, 29(2) :167–197.
- Fernández-Domínguez, J., Bagasheva, A., and Clares, C. L. (2020). *Paradigmatic Relations in Word Formation*. Brill.
- Hathout, N. and Namer, F. (2022). Paradis : a family and paradigm model. *Morphology*, pages 1–43.
- Higginbotham, J. (2000). On events in linguistic semantics. *Speaking of events*, 49 :80.
- Parsons, T. (1995). Thematic relations and arguments. *Linguistic Inquiry*, pages 635–662.

Fruit production paradigm if we include Xculteur and Xculture compounds

PRODUCT FRUIT	INSTRUMENT PLANT	ORCHARD PRODUCTION PLACE	FARMER PERSON	FRUIT PRODUCTION
BANANE	BANANIER	BANANERAIE	-	BANANICULTURE
PRUNE	PRUNIER	PRUNERAIE	PRUNICULTEUR	PRUNICULTURE
POMME	POMMIER	POMMERAIE	POMMICULTEUR	POMMICULTURE
MELON	MELON	MELONNIÈRE	MELONNIER	...
...	

Superposition of two paradigms that may be inscribed in an artifact creation scenario :

MANUF.OBJECT ARTIFACT	MANUFACTURER PERSON	MANUF.PLACE ARTIFACT	ACTIVITY.N	MATERIAL SUBSTANCE
POT 'pot'	POTIER 'potter'	- -	POTERIE 'pottery'	-
TUILE 'tile'	TUILIER 'tile maker'	TUILERIE 'tile factory'	- -	-
BRIQUE 'brick'	BRIQUETIER 'brickmaker'	BRIQUETERIE 'brickyard'	BRIQUETERIE 'brickery'	-
BOTTE 'boot'	BOTTIER 'bootmaker'	BOTTERIE 'boot factory'	BOTTERIE 'boot making'	-
BRONZE 'bronze (artifact)'	BRONZIER 'bronze smith'	-	BRONZERIE 'bronze smithing'	BRONZE 'bronze'
CÉRAMIQUE 'ceramics (artifact)'	CÉRAMISTE 'ceramist'	-	CÉRAMIQUE 'ceramics (activity)'	CÉRAMIQUE 'ceramics (material)'
...	

Inspiration for stories : semantic frames in FrameNet

Agriculture

Definition:

In this frame an **Agriculturist** cultivates **Food**. The location or **Ground** where the **Food** is cultivated by the **Agriculturist** is often expressed.

Peasants were encouraged to **CULTIVATE** basic food crops such as beans and corn.

Shrimp FARMING seems to be a recipe for disaster.

Some fields were probably **CULTIVATED** only in the dry season. **CNI**

FEs:

Core:

Agriculturist [See] The **Agriculturist** is the agent cultivating **Food**.

Semantic Type: Sentient

Food [Foo] **Food** is the item cultivated by the **Agriculturist**.

Semantic Type: Physical_entity They **CULTIVATE** **fungi**.

Ground [Gro]

This is the the background or context for the **Food** to which the **Agriculturist** pays attention in his agricultural efforts.
Romans were **CULTIVATING** the land in Northern Gaul.