Scenarios and frames in derivation: a case study of derivational families based on animal names

Daniele Sanacore
CLLE, CNRS
Université de Toulouse
daniele.sanacore@univ-tlse2.fr

Nabil Hathout
CLLE, CNRS
Université de Toulouse
nabil.hathout@univ-tlse2.fr

Fiammetta Namer
ATILF, CNRS
Université de Lorraine
fiammetta.namer@univ-lorraine.fr

Abstract

This paper presents an ongoing work on a descriptive device for the characterization of semantic relations in the French derivational lexicon. We call this device the "morphosemantic frame" (MF). In order to describe morphosemantic regularities in the lexicon, we take inspiration from Fillmore's Frame Semantics. We use a case study of derivational families based on animal names to introduce "morphosemantic frames" and to illustrate how derivational families could be described by means of these script-like scenarios.

1 Introduction

Morphological relations are relations of form and meaning. While the formal properties of these relations have been the object of numerous studies, the organization of sense relations and of the structures from which meaning is calculated have been less explored. Our work deals with this latter point. We present a method for the representation of morphosemantic relations in derivational families and provide some elements for its automation and application to large amount of data. More precisely, in this work we present a case study where we apply this methodology on derivational families initiated by animal names in French.

2 Theoretical background

In the framework of paradigmatic derivational morphology (Štekauer, 2014, for a panorama), two notions are central: derivational family and paradigm. Families are sets of derivationally related lexemes (Hathout, 2011). An example is the family of *laver* 'to wash' provided in (1). Derivational families form paradigms, which are sets of families containing the same morphosemantic relations. An example of derivational paradigm is provided in (2), where the families of *laver*, *former* 'to train' and *gonfler* 'to inflate' present the same content relations (Bonami and Strnadová, 2019).

(1) *laver.v* 'to wash', *laveur.n* 'washer', *laveuse.n* 'female washer', *lavage.n* 'washing', *laverie* 'laundromat', *lavable* 'washable'.

(2) *laver.v* 'to wash', *laveur.n* 'washer', *lavage.n* 'washing';
former.v 'to train', *formateur.n* 'trainer', *formation.n* 'training';
gonfler.v 'to inflate', *gonfleur.n* 'inflater', *gonflement.n* 'inflating';

The approach that we adopt for the representation of morphosemantic relations in derivational families is inspired by the principles of Frame Semantics (Fillmore, 1976) and *FrameNet* (Ruppenhofer et al., 2006), a lexical resource that implements it. In Frame Semantics, frames are defined as structures that represent cognitive situations or objects along with their participants or features (called "frame elements"). A semantic frame is described by a sort of "story" that makes the semantic relations between the frame elements explicit. For instance, the COMMERCE-PAY frame is described by the gloss in (3) and is characterized by the frame elements (i.e. the participants) BUYER, SELLER, MONEY and GOODS. Moreover, frames are evoked by some lexical units (LUs), for example by the verb *to pay* or the noun *payment* in the case of , and realized by corpus sentences like the ones presented in (4).

(3) COMMERCE-PAY: This frame involves BUYERS paying MONEY for GOODS to a SELLER. In this frame the MONEY is the direct object, and is mapped to the theme of the transfer.

(4) a. I PAID her 50 dollars for a video game.
BUYER SELLER MONEY GOODS

b. Eurotunnel has offered PAYMENT in shares but TML doesn't want shares.
BUYER MONEY SELLER MONEY

3 Frames for morphosemantic description

This work is based on an approach to semantics in derivational morphology that adapts the semantic frames presented in Section 2 to morphosemantic description. Since we are interested in morphosemantic relations existing in derivational families, the paradigms that we want to represent are morphosemantic paradigms (i.e. sets of families structured by the same semantic relations). In this approach, frame-like structures contribute to the semantic characterization of the relations in derivational families. Derivational families are seen as implementations of such frames, in the same way as corpus sentences are seen as the concrete realizations of frames in *FrameNet* . Let us consider the example of the animal morphosemantic frame in (5). Such a frame contains as frame elements several features, concepts and participants that are generally associated with animals. As it can be seen, the frame is rather general; some aspects related to animals are missing; it involves concepts that are related to animals in different ways. This frame can however be used as a starting point to illustrate the method we propose.

(5) ANIMAL: An animal is a living being with certain PHYSICAL FEATURES (SIZE, COLOR, FUR, FLESH, OTHER PHYSICAL PECULIARITIES) and with a BEHAVIOR. The animal has a certain relationship with humans and it can be involved in human activities such as HUNTING, FISHING, BREEDING and SCIENTIFIC RESEARCH conducted by HUNTERS, FISHERMEN, BREEDERS and SCIENTISTS. If the animal is hunted, fished or bred, it is usually eaten by humans and used in RECIPES to prepare FOOD. The animal can be associated with some STEREOTYPES. If a the animal is a pest, it can be the target of a REMOVAL PROCEDURE realized by SPECIALISTS which use INSTRUMENTS.

A morphosemantic frame is a structure that describes a set of concepts related to a sort of semantic pivot, in this case, animal. If we consider derivational families initiated by animal names, such a conceptual structure can help characterize the meaning of derived nouns, relational adjectives and derived verbs. For example, verbs such as *zébrer* 'to stripe' and *léopardiser* 'to stain' are semantically related to the FUR of the animal, while a verb like *renarder* 'to fox' is semantically related to a STEREOTYPE associated with the fox (to be a cunning animal). Such a frame could also help the description of polysemy. For example, a verb like *saumoner* may mean 'to add salmon to something (for example in recipes)' or 'to give something the color of salmon'.

Others concepts contained in the morphosemantic frame in (5) are directly realized in the derivational families. For example, *renardier* 'fox hunter' and *louvetier* 'wolf hunter' concretely realize the HUNTER frame element, while *chevrier* 'goat breeder' or *apiculteur* 'beekeeper' realize the BREEDER frame element. Finally, lexemes like *dératisation* 'rodent control' or *démoustication* 'mosquito control' realize the REMOVAL PROCEDURE frame element.

4 Methodology for frame creation

We collected derivational families initiated by animal names using the derivational resource *Glawinette* (Hathout et al., 2020) and the *GLÀFF* lexicon (Hathout et al., 2014; Sajous and Hathout, 2015). We also collected lexicographic definitions from two electronic dictionaries, *Wiktionnaire* and *TLFi* (Pierrel et al., 2004).

Families built around animal names often seem to evoke distinct scenarios. For example, let us consider the derivational family built around *sardine* 'sardine' in (6). The lexeme *sardine* has two meanings. The first is the fish itself, while the second defines an object (a metallic pin) that has been named after the fish because of its shape (7). The derived noun *sardinier* is associated with four possible meanings, as shown by the lexicographic definitions in (8). It can denote a fisherman specialized in sardines, a ship used to

fish sardines, the owner of a factory that stocks and sells sardines or a worker of such a factory. These four senses, in a frame-based perspective, seem to realize two distinct scenarios. The first concerns the fishing activity and involves participants such as the fisherman, the tools used for fishing, the boat used for fishing, the fishing activity, the fish itself etc. The other scenario concerns an industrial activity, in this case the production and distribution of sardine cans in factory, the owner of that factory, the product, the grocery shops where this product will be sold, etc. These two scenarios are also evoked by the derived relational adjective *sardinier*, which can refer either to the fishing activity or the industrial distribution of sardines.

(6) *sardine.n, sardinade.n, sardinerie.n, sardinier.n, sardinier.a, sardinière.n, sardinière.a, sardinal.n, sardiner.v, ensardiner.v*[1]

(7) a. SARDINE.N: *Poisson de mer au corps fuselé d'une vingtaine de centimètres de long.*
'sea fish with a streamlined body and around twenty centimeters long'.
b. SARDINE.N: *Broche métallique servant à fixer une tente de camping au sol.*
'metal pin used to fix a camping tent to the ground'.

(8) a. SARDINIER: *Pêcheur de sardines* 'sardine fisherman'.
b. SARDINIER: *Ouvrier, ouvrière qui prépare les sardines* 'worker that prepares sardines'.
c. SARDINIER: *Industriel de la sardine* 'sardine industrialist'.
d. SARDINIER: *Bateau qui se consacre à la pêche à la sardine* 'boat used for fishing sardines'.

(9) SARDINIER.A: *Relatif à la pêche ou aux industries de la sardine* 'related to fishing or to fish industry'.

The derived noun *sardinerie* (10) denotes the factory where sardines are canned and thus realizes a concept inscribed in the scenario of an industrial activity, while the noun *sardinal* (11) is used to refer to the nets used for fishing sardines (along with anchovies and or other species of similar size) and is inscribed in the scenario of the fishing activity.

The derived noun *sardinade* belongs to a third scenario. It denotes a recipe that makes use of sardines and the meal prepared using this recipe (12). In this case, the scenario concerns the preparation of meals with recipes that make use of the flesh, the grease, or other parts of a given animal.

(10) SARDINERIE.N: *Usine où l'on prépare les sardines pour les conserver* 'factory where sardines are prepared in order to be conserved'.

(11) SARDINAL.A: *filets dont les mailles sont calibrées pour prendre des sardines, des anchois, etc.*
'nets whose knits are calibrated to catch sardines, anchovies, etc'.

(12) SARDINADE.N: *Recette de cuisine méditerranéenne où des sardines sont cuites entières.*
'Mediterranean cuisine recipe where the whole body of sardines is cooked'

The denominal verbs *se sardiner* (13a) and *ensardiner* (13b) have a similar meaning related to a stereotype associated with the animal, in this case, the fact of being crammed in cans.

(13) a. SARDINER.V: *(Pronominal) S'entasser comme des sardines dans une boîte de conserve* 'to cram like sardines in a sardine can'.
b. ENSARDINER.V: *Entasser comme des sardines* 'to cram something like sardines'.

The derivational family of *sardine* seems to realize several different scenarios and, on this basis, its lexemes could be semantically described by means of five morphosemantic frames described in (14, 15, 16, 17, 18). These subframes can be merged into a general frame like the one presented in (5).

(14) ANIMAL_FISHING: An ANIMAL is fished for its FLESH, for its GREASE, its SKIN or OTHER BODY PARTS. This ANIMAL is fished by some FISHERMEN, who may use some special BOATS or some special FISHING TOOLS in their activity.

[1]We excluded from the analysis terms that where marked as aged or not representative of contemporary French

(15) ANIMAL_INDUSTRIAL_PREPARATION: The FLESH, the GREASE, SKIN of an animal is generally prepared by some WORKERS in some INDUSTRIES in order to be COMMERCIALIZED in some GROCERIES.

(16) ANIMAL_RECIPES: The FLESH, the GREASE, the ORGANS or other parts of an animal are generally used in some RECIPES to prepare some FOOD.

(17) ANIMAL_OBJECT_ASSOCIATION: An OBJECT resembles an ANIMAL or a PART OF THE ANIMAL in its appearance.

(18) ANIMAL_BEHAVIOR: An ANIMAL is associated with a given stereotyped BEHAVIOR. A PERSON that adopts this BEHAVIOR can be given the name of the ANIMAL in a metaphoric sense.

These subframes describe a set of concepts that are concretely realized by the lexemes in the family of *sardine*. Their relations with the other members of the family could be glossed as in (19a, 19b, 19c, 19d, 19e), where these lexemes are substituted for the frame elements they realize.

(19) a. Une SARDINE est pêchée par un SARDINIER qui se trouve à bord d'un SARDINIER et utilise un SARDINAL.
'a sardine is fished by a sardine fisherman who is sailing on a sardine boat and using a sardine net'.

b. Un SARDINIER est un industriel qui possède une usine où des SARDINIERS entassent des SARDINES dans des boîtes de conserve.
'a sardine industrialist owns a factory where workers cram sardines in cans'.

c. Une personne prépare un plat à base de SARDINES en suivant une recette, la SARDINADE.
'a person prepares a dish based on sardines following a recipe that uses sardines'.

d. Une SARDINE est un objet qui rassemble à une SARDINE.
'a metal pin (sardine) is an object that resembles to a sardine'

e. Une SARDINE est associée à l'état d'être ENSARDINÉ.
'a sardine is associated with the state of being crammed in cans'

We extended this operation to other derivational families in an iterative way in order to (*i*) validate the subframes already created, and (*ii*) create other subframes based on the concepts realized in families.

To illustrate the point (*i*), we can find other derivational families that fit the scenarios that we proposed. These families are initiated by animals that are involved in fishing, industrial distribution or cuisine like *saumonier* 'fisherman specialized in salmons ', *morutier* 'fisherman specialized in cods', and *carpiste* 'fisherman specialized in carps' which realize the FISHERMAN frame element, while *harenguier* 'ship used for fishing herrings', *homardier* 'ship used for fishing lobsters', and *thonier* 'ship used for fishing tuna' realize the BOAT frame element. In other words, several derivational families realize the same ANIMAL_FISHING morphosemantic frame and as a consequence, can be aligned with respect to the frame.

To illustrate the point (*ii*), let us consider the family of *renard* 'fox' in (20). The analysis of this family involves a hunting scenario (25) which includes the fur of the hunted animal (*renard*), the animal itself (*renard)*, the hunter (*renardier*), and the lair where the animal hides (*renardière*). In addition, the family of *renard* presents a group of lexemes related to the stereotyped behavior associated with the animal (*renard*, *renardie*, *renardise*).

(20) *renard.n, renarder.v, renardie.n, renardise.n, renardier.n, renardière.n*

(21) a. RENARD: *Mammifère carnivore, au museau pointu et aux oreilles droites.*
' carnivorous mammal, with a pointed snout and straight ears'

b. RENARD: *fourrure de renard* ' fox fur'

c. RENARD: *Personnage cauteleux, fin et rusé* ' cunning, shrewd person'

(22) RENARDIER: *(Chasse) Celui qui est chargé de prendre les renards*
'person in charge of catching foxes'

(23) RENARDIÈRE: *Tanière du renard.* ' fox lair'.

(24) RENARDIE[2]: *Ruse, déloyauté, action de renard* ' disloyalty, cunning action'

(25) ANIMAL_HUNTING: An ANIMAL is hunted by a HUNTER because of its FLESH, FUR, GREASE or BODY PARTS. The hunter make use of WEAPONS, TRAPS and HUNTING ANIMALS.

5 Lexicographic information for an automatic feeding of morphosemantic frames

Another question we are intersted in is the (partial) automation of our method. What are the elements present in lexicographic definitions that could be helpful to assign a derived lexeme to the frame element FISHERMAN in the ANIMAL_FISHING frame? How can we find the derivational families that realize a given morphosemantic frame?

We can use some recurring structures in definitions and various keywords and labels. Consider the definitions of *carpiste* ' carp fisherman', *saumonier* 'salmon fisherman' and *morutier* 'cod fisherman' in 26) which realize the FISHERMAN frame element in the ANIMAL_FISHING subframe. Their definitions have a similar structure and contain regular key-phrases like "*pêcheur*+ [gerundive]" or "*marin-pêcheur*". Similar regularities are identifiable for the lexemes that realize the BOAT, the FOOD and the FUR frame elements, as shown in (27), (28) and (29).

(26) a. CARPISTE: *Pêcheur se consacrant uniquement à la pêche de la carpe* 'fisherman who dedicates himself to carp fishing'
b. SAUMONIER: *Personne pratiquant la pêche au saumon* 'person that fishes salmon'
c. MORUTIER: *Marin-pêcheur pratiquant la pêche à la morue* 'fisherman who fishes cod'

(27) a. HARENGUIER: *Bateau spécialisé dans la pêche du hareng* 'boat specialized in herring fishing'
b. THONIER: *Bateau destiné à la pêche au thon* 'boat used for tuna fishing'.

(28) a. ANCHOÏADE: *Préparation culinaire à base d'anchois pilés et de câpres.* 'culinary preparation made using piled anchoivies and and capers'.
b. HOMARDINE: *(Cuisine) Sauce à base de homard* 'sauce made with lobsters'

(29) a. VISON: *(Par métonymie) Fourrure de cet animal* '(metonymy) mink fur'
b. LOUTRE:*(Par métonymie) Fourrure de cet animal* '(metonymy) otter fur'

Some examples of markers that can denote the realizations of the cited frame elements are provided in Table 1.

FISHERMAN	BOAT	FUR	FOOD
pêcheur de	*bateau*	*fourrure de*	*préparation culinaire*
pêcheur+gerund.	*bateau*+ [part.]	*par métonymie* + fourrure	*cuisine* + à base de
pêche	*pêche*	*par ellypse*+ fourrure	*sauce*

Table 1: Frame element markers in dictionnary definitions

6 Conclusion

In this work, we used a case study of derivational families based on animal nouns in order to introduce and illustrate morphosemantic frames. We showed that such families seem to realize distinct scenarios where animals are involved and that several families can be aligned with respect to those scenarios on the basis of their morphosemantic regularities.

[2]The same definition is provided for *renardise*.

Acknowledgements

This work benefited from the support of the project DEMONEXT ANR-17-CE23-0005 of the French National Research Agency (ANR).

References

Olivier Bonami and Jana Strnadová. 2019. Paradigm structure and predictability in derivational morphology. *Morphology* 29(2):167–197.

Charles J. Fillmore. 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*. pages 20–32.

Nabil Hathout. 2011. Morphonette: a paradigm-based morphological network. *Lingue e linguaggio* 10(2):245–264.

Nabil Hathout, Franck Sajous, and Basilio Calderone. 2014. Glàff, a large versatile french lexicon. In *Conference on Language Resources and Evaluation (LREC)*. pages 1007–1012.

Nabil Hathout, Franck Sajous, Basilio Calderone, and Fiammetta Namer. 2020. Glawinette: a linguistically motivated derivational description of French acquired from GLAWI. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*.

Jean-Marie Pierrel, Jacques Dendien, and Pascale Bernard. 2004. Le tlfi ou trésor de la langue française informatisé. *Actes de EURALEX* 4.

Josef Ruppenhofer, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California. Distributed with the FrameNet data.

Franck Sajous and Nabil Hathout. 2015. GLAWI, a free XML-encoded Machine-Readable Dictionary built from the French Wiktionary. In *Proceedings of the of the eLex 2015 conference*. Herstmonceux, England, pages 405–426.

Pavol Štekauer. 2014. Derivational paradigms. *The Oxford handbook of derivational morphology* 354:369.