



HAL
open science

What do derivational paradigms tell us about back-formation and what does back-formation tell us about derivational paradigms?

Nabil Hathout, Fiammetta Namer

► To cite this version:

Nabil Hathout, Fiammetta Namer. What do derivational paradigms tell us about back-formation and what does back-formation tell us about derivational paradigms?. *Word Structure*, 2025, 18 (3), pp.239-280. <10.3366/word.2025.0256>. <hal-05407592>

HAL Id: hal-05407592

<https://hal.science/hal-05407592v1>

Submitted on 9 Dec 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

What do derivational paradigms tell us about back-formation and what does back-formation tell us about derivational paradigms?

Nabil Hathout 

CNRS & University of Toulouse – Jean Jaurès

Fiammetta Namer 

CNRS & University of Lorraine

Abstract

Back-formation is the process that leads to the coinage of a lexeme by deleting a sequence present in the base, such as the verb *babysit* from the noun *babysitter*. Back-formed words exhibit a distortion between a meaning that is more complex than that of their base and a form that is simpler. The very nature of this process remains to be established. Our study suggests that back-formation (i) encompasses multiple realities, (ii) performs a formal regularization within paradigms, and (iii) most of the time involves meaning–form discrepancies. Based on five case studies of French verbs that are either back-formed from nouns or adjectives, or have the structure of neoclassical compounds, we propose that these verbs fall into three main groups, depending on the structures they involve. Our analysis highlights the importance of a paradigmatic description where formal and semantic dimensions are distinguished. We also show how a paradigmatic framework such as ParaDis sheds new light on the nature of back-formation. Conversely, back-formation helps us better characterize the structure of paradigmatic families and the relations between newly-created families and existing ones.

Keywords: back-formation, derivational paradigm, paradigmatic family, family center, stolon, quasi-stolon, neoclassical verb compounds

Word Structure 18.3 (2025): 239–280

DOI: 10.3366/word.2025.0256

© The Authors. The online version of this article is published as Open Access under the terms of the Creative Commons Attribution–NonCommercial–NoDerivs 4.0 Licence (<http://www.creativecommons.org/licenses/by-nc-nd/4.0/>)
www.eupublishing.com/word

I. Introduction

In morphology, back-formation is the process that leads to the coinage of a lexeme or word-form by deleting a sequence present in the base, such as the verbs *babysit*_V (1a) from *babysitter*_N and *helicopt*_V (1b) from *helicopter*_N. The formal mechanism at work in back-formation is similar to clipping (e.g., *bro*_N from *brother*_N), so that the result of back-formation is a word that is formally simpler than its base. However, unlike back-formation, clipping always produces a lexeme of the same part-of-speech as its base. Moreover, according to Bauer (2003: 81), “[c]lipping always provides synonymous words from the same base but here the different style level (not the different meaning) allows both to co-exist.” In other words, clipped words are evaluative variants of their base, whereas back-formed words are semantically distinct from their base.

- (1) a. How do you deal with a creepy babysitter if you’re the child being babysat?
(<https://www.quora.com/How-do-you-deal-with-a-creepy-babysitter-if-youre-the-child-being-babysat>)
- b. They were alive and very well and had been helicopted off the Zmutt ridge that morning.
(https://www.ukclimbing.com/forums/expedition+alpine/four_days_out_on_the_matterhorn-141550)

Back-formation is a word-formation device that has challenged many theoretical frameworks of morphology, from those based on morpheme concatenation to lexeme-based, word-based models, and analogy-based approaches. Back-formation raises many questions: Is it a derivational process in its own right, or is it the inversion of a rule? Is the deleted sequence always the exponent of an affix? Is back-formation universal? Does it cover derivation and inflection? But what really intrigues the researchers is the identification of the process itself: In what cases can we assert that two lexemes are in a back-formation relation and not in a standard derivational relation?

This paper does not aim to answer all these questions but to propose a classification of various types of back-formed derivatives in French, using a paradigmatic model of morphology. Our purpose is to show that (i) the paradigmatic organization of the morphologically complex lexicon is an appropriate level for identifying and describing the behavior of the different classes of back-formed lexemes documented in the literature, and (ii) the analysis of these lexemes depends crucially on the ability of the adopted model to handle separately the semantic, categorical (part-of-speech) and formal (phonological) levels of description.

The paper is organized as follows. §2 provides an overview of work on back-formation and derivational paradigms, as well as a brief presentation of the ParaDis model we use for our analyses. In §3, we focus on back-formed verbs, which form the most frequent category of back-formed words (Manova 2018). The process is considered first diachronically, then synchronically. §§4–7 present case studies on different types of back-formed verbs, illustrated by the analysis of the following examples: *lacter*_V ‘produce milk’ (← *lactation*_N ‘lactation’), *allocuter*_V ‘allocute’ (← *allocation*_N ‘speech’), *intuiter*_V

‘guess’ (← *intuition*_N ‘intuition’), *cryograver*_V ‘cryograph’ (← *cryogravure*_N ‘cryograph’), *inoxyder*_V ‘inoxidize’ (← *inoxidable*_A ‘stainless’). Through these analyses, we show that the concept of back-formation encompasses multiple realities whose specificities are made visible thanks to the representation provided by a paradigmatic framework such as ParaDis (§8).

2. Related work on back-formation and on derivational paradigms

2.1 Back-formation

According to the comprehensive survey of the issue by Štekauer (2015), back-formation is typically regarded by researchers as a word-formation process in which “something” is removed (in the words of Booij 2010). As Štekauer points out, this subtractive operation is characteristic of marked processes like clipping, which Dressler (2005), among others, describes as “unnatural”. In contrast, there is no consensus on the nature of what is deleted in a back-formed word: For some (Pennanen 1966; 1975) it is a morpheme, for others (Nagano 2007) a pseudo-affixal exponent. Rainer (2004: 497) lists several examples from Italian in which a suffix was deleted to back-form a pseudo-base. Thornton (2007: 264), quoting D’Achille (2005), distinguishes cases like *cappuccio*_N ‘large cappuccino’ ← *cappuccino*_N ‘latte’ where back-formation involves deletion of the diminutive suffix *-ino*, from cases like *batto*_N ‘large rowboat’ ← *battello*_N ‘boat’ where the removed *-ello* sequence is a pseudo-suffix, *battello*_N ‘boat’ being borrowed from Old French.

This uncertainty raises many questions: Is back-formation the result of a sequence of derivational rules? Is it a form of compounding? Does it follow from some analogy-based principles? Are back-formed words the outcome of the reversed application of a regular process? Are they formed by means of bidirectional schemas?

The difficulty of defining back-formation and describing how it works is reflected in debates about this process, which focus on: (i) the diachronic or synchronic dimensions of this word-formation process, (ii) the analysis of the verbal compounds, which make up a significant proportion of back-formed verbs, and (iii) the choice of the best suited theoretical approach – morpheme- or lexeme-based, rule- or analogy-based, using binary, oriented or paradigmatic relations, etc.

The description of back-formation given below reflects our position on these issues.

(a) Following Bauer (1983) and Becker (1993), among others, we consider that back-formation is to be analyzed synchronically, but of course the identification of the process requires taking into account diachronic clues to estimate and compare the dates of first occurrence of the back-formed word and its formally more complex base. In particular, we rely on the presence of the former and the absence of the latter in the major French general dictionaries and in large corpora like frCOW (Schäfer & Bildhauer 2012), as well as on attestation and frequency in the Google Books Ngram corpus.

(b) Štekauer (2015) proposes an analysis of compound verbs (which make up an important subset of back-formed lexemes, in English at least) in which they are formed

directly from their base verb, by nominal incorporation (e.g., *brainstorm*_V by incorporation of *brain*_N into the verb *storm*_V) or adjectival incorporation (e.g., *slowmove*_V by incorporation of *slow*_A into the verb *move*_V). This analysis is refuted by Plag (2003: 154–155), who rejects the use of incorporation in the formation of English compound verbs. He observes that the English lexicon does include morphologically complex verbs having a noun-verb (e.g., *proofread*_V, *talent-spot*_V) or an adjective-verb (e.g., *shortcut*_V) structure. However, he notes that this analysis is contradicted by:

(...) the general impossibility of verbs to incorporate adjectival/adverbial non-heads. For instance, neither *read a book*, *steal a car* nor *drive fast*, *move slowly* can be readily turned into compounds (**bookread*, **carsteal*, **fastdrive*, **slow(ly)-move*), whereas nominalized verbs and their arguments (as in *the reading of books*, *a driver of trains*) and deverbal adjectives and their adverbial/adjectival modifiers are happily condensed to compounds (*book-reading*, *train-driver*, *a fast-driving chauffeur*, *a slowmoving animal*). (Plag 2003: 155).

This is why Plag also claims that:

The majority of compounds involving a verbal head is best analyzed as the result of a back-formation or conversion process. Thus, [*proof-read* and *talent-spot*] are (...) back-formations from noun-noun compounds with either a verbal noun in *-ing* or a person noun in *-er* in head position (e.g. *proof-reading*, *talent-spotter*). With regard to [*short-cut*], conversion is involved (...). (Plag 2003: 154).

On the other hand, incorporation is defended by Kastovsky (2009), among others, and, according to Štekauer (2015), this word-formation process solves all the problems posed by back-formation, which other approaches are unable to deal with in a fully satisfactory way. It follows from the incorporation analysis that the formation of compound verbs is distinct from that of their related compound nouns, which could be formed by regular derivation from the verb. Yet, Štekauer's (2015) analysis fails to explain why certain verbs like *sky react*_V or *eye rotate*_V seem not to be attested, even though the corresponding nouns *sky reaction*_N or *eye rotation*_N are. Similar observations on neoclassical compound verbs in French make us refute their analysis as resulting from nominal or adjectival incorporation: As we will see in §6.1, our position is in line with Plag's (2003) arguments.

(c) Our analysis is carried out within a paradigmatic framework of derivation (§2.2) and differs in this respect from most of the work summarized in Štekauer (2015). It is based on the assumption that (i) in the lexicon of a language, all paradigms are determined by semantics, (ii) any new lexical unit – and in particular back-formed lexemes – fits into a paradigm, and (iii) the form of a neologism results from a balance between the size of the stem and the faithfulness to the base form. These hypotheses rely in part on two types of word-based, non-oriented back-formation mechanisms. The first corresponds to Booij & Masini's (2015) second-order schemas that regularize the complex relations between a

back-formed verb and its base; the second draws on formal analogy and is proposed by Arndt-Lappe (2015) to explain the forms of back-formed verbs.

2.2 Paradigms in derivation

The analysis of back-formation proposed in this article is based on a paradigmatic approach to derivation. This approach has gained an increasingly solid legitimacy among morphologists (Štekauer 2014; Hathout & Namer 2018; 2019; Fernández-Domínguez et al. 2020; among others), particularly following work showing that (i) the two morphology modules, namely inflection and derivation, are not separate, and (ii) a continuum exists from inflection to derivation. It also stems from the fact that the paradigmatic nature of inflection has been widely accepted since Carstairs-McCarthy (1994). The exact nature of the boundary between inflection and derivation is, of course, still a matter of debate (Štekauer 2014; Booij 2000; Spencer 2013; Haspelmath 2024), and there is no consensus on the similarities and differences between derivational and inflectional paradigms (Spencer 2021). Recently, Bonami & Strnadová (2019), Boyé & Schalchli (2019) and Hathout & Namer (2022) have proposed derivational transpositions of inflectional paradigms, their objects and mechanisms. The derivational paradigms and the other objects proposed in these works are defined as follows:

(i) A **derivational family** (or word family) is a set of lexemes connected by derivational relations (i.e., relations of formal, categorical and semantic contrast), such as those in (2), taken from Sanacore et al. (2021).

(2) *pot*_N ‘pot’, *poterie*_N ‘pottery’, *potier*_N ‘potter.M’, *repoter*_V ‘repot’, *repoteur*_N ‘repotter.M’, *repotage*_N ‘repotting’, *repotable*_A ‘repottable’, ...

(ii) A **derivational paradigm** (3) is a superposition of **paradigmatic families**, i.e., of subsets of derivational families (Hathout & Namer 2022). For example, *pot*_N ‘pot’, *potier*_N ‘potter.M’, and *poterie*_N ‘pottery’ form a paradigmatic family, subset of the word family (2) and part of the derivational paradigm (3).

(3) PRODUCT_N	AGENT_N.M	ACTIVITY_N
<i>pot</i> _N ‘pot’	<i>potier</i> _N ‘potter.M’	<i>poterie</i> _N ‘pottery’
<i>verre</i> _N ‘glas’	<i>verrier</i> _N ‘glass maker.M’	<i>verrierie</i> _N ‘glass making’
<i>carrelage</i> _N ‘tile’	<i>carreleur</i> _N ‘tiler.M’	<i>carrelage</i> _N ‘tiling’
<i>lithographie</i> _N	<i>lithographiste</i> _N	<i>lithographie</i> _N
‘lithography’	‘lithographer’	‘lithography’
<i>tissu</i> _N ‘fabric’	<i>tisserand</i> _N ‘weaver.M’	<i>tissage</i> _N ‘weaving’

(iii) Two paradigmatic families are superposed within the same derivational paradigm if all the relations that exist between their elements are **aligned**, i.e., if they show the same content contrasts (Bonami & Strnadová 2019). For example, in (3), the nouns *potier*_N ‘potter.M’, *verrier*_N ‘glass maker.M’, *carreleur*_N ‘tiler.M’, *lithographiste*_N

‘lithographer’, and *tisserand*_N ‘weaver.M’ are **semantically** aligned, because they are in the same semantic relation with the other members of their respective family: *Carreleur*_N ‘tiler.M’ is the AGENT_N of the ACTIVITY_N *carrelage*_N ‘tiling’, whose purpose is to produce the PRODUCT_N *carrelage*_N ‘tile’. In addition, the nouns *potier*_N ‘potter.M’ and *verrier*_N ‘glass maker.M’ are **fully** aligned, because they show the same semantic contrast with the other members of their respective family and the same formal contrast, which could be represented by means of formal alternation patterns like ([X]~[X]e), [X]~[X]əri), [X]e~[X]əri) where X stands for a stem shared by all members of the paradigmatic family.

- (iv) Defining a derivational paradigm then boils down to **identifying** the paradigmatic families it contains.
- (v) In a derivational paradigm, each cell contains a lexeme which belongs to a derivational family and to a **derivational series** made up of all the lexemes it aligns with, based on its part-of-speech, meaning, and morphological pattern. For example, in (3), *potier*_N ‘potter.M’ aligns with *verrier*_N ‘glass maker.M’. Note that the formal criterion is not considered necessary in some theories. For example, according to Bonami & Strnadová (2019), rival patterns align in the same derivational series. In their model, *potier*_N ‘potter.M’ aligns also with *tisserand*_N ‘weaver.M’, *carreleur*_N ‘tiler.M’, and *lithographiste*_N ‘lithographer’.
- (vi) Sanacore et al. (2021) observe that derivational paradigms are lexical realizations of **semantically motivated networks** (Fradin 2020; Roché 2023). For example, the triple of concepts {AGENT_N, ACTIVITY_N, PRODUCT_N}, which semantically specifies the derivational paradigm (3), is an instance of an *activity network*, defined by Fradin (2018: 166) as the interaction of at least three concepts: a nominal activity (‘pottery’), the human characterized by the ability or habit to perform this activity (‘potter’), and the object or substance that motivates the activity (‘pot’).

The previous definitions and propositions have the following corollaries:

- (i) The lexicon of a language contains all the derivational paradigms in which morphologically complex units are formed.
- (ii) A paradigmatic family necessarily fits into a paradigm.
- (iii) A derivational family usually contains several paradigmatic families.
- (iv) The coinage of a new word (i.e., lexeme) creates a relation with a base word within the same family, within a different family of the same paradigm, or within a different paradigm. This observation echoes a quote from Horecký et al. (1989: 42) reported and translated in Štekauer (2014: 360): “(...) new complex words do not come into existence as isolated units but rather as complete paradigms. So, for example, if a new field of science is established, the coining of its name motivates the coming-into-existence of other vitally needed derivational “cases” (...)”.

Considering back-formation from a paradigmatic point of view means assuming that back-formed words realize new relations, and that these relations belong to a family

which is itself part of a paradigm. The differences between the various types of back-formation are reflected in differences in the family that hosts the back-formation relation and in the paradigm in which that family is included.

2.3 The ParaDis framework

Several features of the ParaDis model are relevant to the analysis of verb back-formations:

- (i) The model includes four levels of representation: morphological, semantic, categorical and formal (i.e., phonological).
- (ii) These levels are organized paradigmatically into families and paradigms.
- (iii) Formal, categorical and semantic paradigms are independent of one another.
- (iv) A paradigmatic family (respectively, a morphological paradigm) is the mapping of a semantic family (respectively, a semantic paradigm), a categorical family (respectively, a categorical paradigm) and a formal family (respectively, a formal paradigm). A semantic family can be defined as a structured set of interpredictable semantic values (concepts), and a formal family as a structured set of interpredictable formal (phonological) representations. As a consequence, a paradigmatic family is a subset of a morphological family (i.e., a derivational family) whose members are formally and semantically interpredictable.
- (v) Word-formation, including back-formations (as we will see), operates within the morphological paradigms. Families and paradigms at the other levels of representation (formal, categorical and semantic) motivate these formations and help explain their analysis.

Hathout & Namer (2019; 2022) and Namer & Hathout (2020) have shown that (i) performing the analysis at the family level, and (ii) dissociating meaning, form and category allow for a unified analysis of regular processes, meaning-form discrepancies, defective families, suppletion, cross-formation, affix competition, duplication, etc. In the next sections of this article, we show how ParaDis can also be used to analyze different cases of verb back-formation, and to discriminate between different types of back-formation.

3. Back-formed verbs

3.1 Grammatical category of back-formed lexemes

To the best of our knowledge, the grammatical category of back-formed words has never been investigated specifically. Manova (2018) suggests that back-formation produces mainly verbs; and in fact the vast majority of examples illustrating the phenomenon in the literature (including the present paper) are verbs formed from nouns denoting an action, agent, or instrument. If this turns out to be a universal trend, and if back-formation does actually produce primarily verbs in all languages where this formation process is present, then the most likely explanation for this observation fundamentally lies in the semantic

and pragmatic properties of verbs: In its derivational family, (i) a verb is conceptually central, and (ii) it is usually structurally simpler than the action noun and the nouns referring to the participants in the process it denotes.

Point (i) means that the absence of a verb, more than that of any other lexical category, creates a “gap” and consequently the need to fill it, even though no affixal process is available to achieve this, given (ii). Back-formation can also yield lexemes of other lexical categories, as in the case of the adjective *nocent*_A ‘guilty’ in (4), back-formed from the adjective *innocent*_A ‘innocent’.

- (4) *innocent*_A ‘innocent’ → *nocent*_A ‘guilty’

Le Wokenisme n’est pas innocent. Il est nocent.

‘Wokenism is not innocent. It is nocent’

(<https://brunobertez.com/2022/05/17/le-woke-le-capital-la-reproduction-changer-lhomme-pour-que-rien-ne-change/>)

For the reasons we have just discussed, our study focuses on back-formation in French, and on verb back-formation in particular. Many of the cases we examine involve an action noun suffixed in *-ion*, such as *lactation*_N ‘lactation’, and the corresponding back-formed verb, in this case *lacter*_V ‘produce milk’. However, our analyses can easily be adapted to verb back-formations from other lexeme classes such as action nouns suffixed with another exponent (§6), nouns ending in *-ure* and denoting the result of a process (5) or its concrete product (6), abstract nouns ending in *-ie* (7), adjectives ending in *-able* (§7), *-eux* (8), *-ant/-ent* (9), or *-oire* (10), agent nouns ending in *-eur* (11), or instrument nouns ending in *-erie* (12).

- (5) *sépulture*_N ‘burial’ → *sépulter*_V ‘bury’

Enfin, il écrit à Humboldt « Je passe une vie tranquille et je vais mourir là où sépulter (sic) mes tristes restes à l’ombre des arbres nombreux. »

‘Finally, he wrote to Humboldt: “I’m spending a quiet life and I’m going to die where I can bury (sic) my sad remains in the shade of many trees”’

(<http://www.humboldt-bonpland.eu/Aime-Bonpland>)

- (6) *progéniture*_N ‘offspring’ → *progéniter*_V ‘offspring’

celles qui n’auront d’autres tâches futures que de « progéniter » des marmots...

‘those whose only future tasks will be to “offspring” brats’

(<https://www.lepetitsolognot.fr/billet-dhumeur-de-fabrice-simoes-22/>)

- (7) *stratégie*_N ‘strategy’ → *stratéger*_V ‘plan’

[Elle] est ce qui peut être de meilleur pour l’humain, le contraignant sans cesse à s’interroger, se positionner, “stratéger”, rester en éveil sans aucune zone de confort. [It] is what can be best for human being, constantly forcing one to question oneself, to position oneself, to “plan”, to stay awake without any comfort zone’

(https://fr.linkedin.com/posts/bruno-tertrais-0793b464_bruno-tertrais-la-guerre-des-mondes-sera-activity-7126108047336202240-sLZM)

- (8) *délictueux*_A ‘criminal’ → *délictuer*_V ‘commit a crime’
 Qu’il fait bon “délictuer” en ce beau pays.
 ‘How nice it is to “commit a crime” in this beautiful country’
 (https://www.lepoint.fr/societe/l-ex-convoyeur-musulin-condamne-a-verser-270-000-a-son-ancien-employeur-04-02-2011-135765_23.php)
- (9) *turbulent*_A ‘turbulent’ → *turbuler*_V ‘disturb’
 Et si le « système », stable en apparence, finissait par « turbuler », comme disait Jean-Pierre Chevènement en 2002 ?
 ‘What if the seemingly stable “system” ended up “disturbing”, as Jean-Pierre Chevènement put it in 2002?’
 (<https://www.lesechos.fr/elections/presidentielle/presidentielle-2022-le-systeme-peut-il-turbuler-1348264>)
- (10) *réhibitoire*_A ‘redhibitory’ → *réhibiter*_V ‘revoke’
 La procédure pénale, en tant que servante du droit substantiel, doit être exécutée avant que s’écoule un temps susceptible de frustrer, de « réhibiter » ou de dirimer l’application des règles de fond.
 ‘Criminal procedure, as the servant of substantive law, must be carried out before a period of time elapses that could frustrate, “revoke” or nullify the application of substantive rules’
 (<https://corpus.ulaval.ca/bitstreams/ba5dfb3e-a4ad-4a13-a3c4-b20826f49440/download>)
- (11) *précurseur*_N ‘precursor.M’ → *précurser*_V ‘precurese’
 Il faut laisser les précurseurs précurser, on en a besoin.
 ‘We must let the precursors precurese, we need them’
 (<https://www.espritroue.fr/topic/3667-essai-du-tesla-gotway-en-boutique-et-retours-b%C3%AAAs-testeurs/?page=18>)
- (12) *quincaillerie*_N ‘firearm’ → *quincailler*_V ‘shoot’
 La nuit du 5 au 6, ça a quincailé de tous les bords.
 ‘On the night of the 5th to the 6th, there was shooting on all sides’
 (<https://www.ouest-france.fr/normandie/saint-lo-50000/74e-d-day-la-nuit-du-5-juin-ca-quincaille-de-tous-les-bords-5797671>)

3.2 Identifying back-formed words

3.2.1 Historical account

Marchand (1969) and Kastovsky (1968), among others, defend the idea that back-formation is a phenomenon of purely diachronic interest (see the discussion in Štekauer 2015). In this view, the date of first occurrence is the only clue we have to establish the anteriority of the noun relative to the verb from which it derives. For example, in (13), *oppression*_N ‘oppression’ and *division*_N ‘division’ predate *oppresser*_V ‘oppress’ and *diviser*_V

‘divide’, according to the *Trésor de la Langue Française* (TLF) dictionary. These verbs could therefore be regarded as back-formations. Conversely, *annexion*_N ‘annexation’ and *agression*_N ‘aggression’ came after *annexer*_V ‘annex’ and *agresser*_V ‘aggress’. Note that, according to Corbin (2001), *agresser*_V ‘aggress’ is posterior to *agression*_N ‘aggression’ and should therefore be regarded as back-formed.

- | | | |
|------|--|---|
| (13) | <i>oppresser</i> _V ‘oppress’ (1350) | <i>oppression</i> _N ‘oppression’ (1175) |
| | <i>diviser</i> _V ‘divide’ (1377) | <i>division</i> _N ‘division’ (1119) |
| | <i>annexer</i> _V ‘annex’ (1274) | <i>annexion</i> _N ‘annexation’ (end of the 14th century) |
| | <i>agresser</i> _V ‘aggress’ (1375) | <i>agression</i> _N ‘aggression’ (1395) |

In all the examples in (13), the noun denotes an action and can be paraphrased as “action of @”, where @ represents the semantic content of the corresponding verb. These verbs and nouns have been part of the general French lexicon since around the 12th century, and no speaker is able to tell which of the noun or verb predates the other. From a morphological point of view, speakers perceive pairs like *oppresser*_V ‘oppress’/*oppression*_N ‘oppression’ and *diviser*_V ‘divide’/*division*_N ‘division’ as perfectly regular formations similar to the other two.

3.2.2 Back-formation and neologisms

Sense of novelty: Speakers may feel that a verb is new compared to the corresponding formally more complex noun or adjective because the relation between them is not perfectly regular, and that the verb is morphologically complex, despite its simpler form. Among web users, this uncertainty about the direction of the relation sometimes shows up in the use of discourse markers such as “(sic)” and typographic markers like quotation marks (on nonce formations marking, see Dal & Namer 2018). Examples (5) to (10) above illustrate this, as do examples (15) to (18) below. Contexts indicate that the speaker may express surprise (15) or indignation (16) at the use of the verb, apologize for its use (17), or claim its creation (18) (on speakers’ ambivalent feelings about their own lexical creations, see Dal & Namer 2012), but not always, as can be seen in examples (14) and (19) to (21).

- (14) Votre vétérinaire pourra vous fournir des instructions et des médicaments pour aider votre lapin à lacter et à nourrir ses bébés.
 ‘Your vet will be able to provide you with instructions and medication to help your rabbit lactate and feed its babies’
 (<https://planete-lapin.com/gestation-lapin-comportement-et-soins/>)
- (15) Ellipse d’un verbe ? Et pourtant, on dit : *une allocution en direct*, et non *une allocution directe*, bien que le verbe *allocuter* n’existe pas.
 ‘Ellipsis of a verb? And yet we say: *une allocution en direct*, not *une allocution directe*, even though the verb *allocuter* doesn’t exist’
 (<https://www.question-orthographe.fr/question/en-presentiel-ou-presentielle/>)

- (16) C'est à propos du terme « prédateur ». (...) Je dépose plainte contre X pour l'invention de ce *néologisme* trop mal défini pour qu'il ne soit pas louche
 'It's about the term "predate". (...) I'm reporting the crime of inventing this *neologism*, which is too ill-defined not to be suspicious'
 (<https://papiermachine.be/tribunal-des-mots/>)
- (17) (Oui je sais qu'*intuiter* est un beau néologisme mais je ne trouvais pas d'équivalent dans notre belle langue).
 '(Yes I know *intuiter* is a nice neologism but I couldn't find an equivalent in our beautiful language)'
 (<https://les-mathematiques.net/vanilla/discussion/2330273/paradoxe-dachille-et-de-la-tortue>)
- (18) Info du jour: le verbe de « *perception* » pour moi c'est « *percepter* ».
 'Tip of the day: for me, the verb for "perception" is "percept"'
 (<https://twitter.com/kimmyphoria/status/1433781205342773252>)
- (19) Et voilà, vous pouvez commencer à pyrograver !
 'Now you can start pyrographing!'
 (<https://www.cultura.com/comment-debuter-une-activite/comment-debuter-loisirs-creatifs/je-me-lance-dans-la-pyrogravure.html>)
- (20) (...) le chrome ajouté pour inoxyder certains aciers, se retrouve dans la scorie inox ou ALZ
 '(...) the chromium added to inoxidize certain steels is found in stainless steel or ALZ slag'
 (<https://docplayer.fr/212831627-Granulats-secondaires-sous-produits-d-acierie.html>)
- (21) Vous pouvez l'utiliser pour infalsifier des documents et protéger vos documents critiques.
 'You can use it to make documents tamper-proof and protect your critical documents'
 (<https://www.guru99.com/fr/google-drive-alternatives.html>)

Frequency and date of first occurrence: In addition to the above examples, Table 1 lists verbs (first column) that probably entered the French lexicon after the corresponding noun or adjective. The clues are the fact that the verb is not attested in the TLF, its frequency in the frCOW corpus is much lower than that of the other lexeme, and its date of first occurrence in the Google Books Ngram corpus, over the period 1800–2019, is later than that of the noun or adjective. These convergent clues support the conclusion that the verbs in the table are all back-formed from the corresponding nouns and adjectives.

Some of the verbs in Table 1, such as *superstiter*_v 'have a superstition', are listed as "potential" constructions in Matthias Kramer's 1712 dictionary (Bray 2000). Kramer's

Table 1. French verb back-formation in synchrony. The TLF column indicates the absence (0) or presence (1) of the verb and its nominal or adjectival base in the dictionary wordlist. The frCOW column gives the number of occurrences of the verb and its base in the corpus. The GBooks column shows the date of first occurrence of the verb and its base in the Google Books Ngram corpus.

Verb	Base word	Verb definition	TLF	frCOW	GBooks
<i>délater</i> _V 'denounce'	<i>délation</i> _N 'informing'	faire une délation 'inform'	0 - 1	7 - 6,586	NA - 1800
<i>allocter</i> _V 'allocute'	<i>allocution</i> _N 'speech'	faire une allocution 'make a speech'	0 - 1	0 - 18,501	NA - 1800
<i>ablater</i> _V 'remove'	<i>ablation</i> _N 'ablation'	faire une ablation 'make an ablation'	0 - 1	61 - 11,502	1898 - 1800
<i>prédater</i> _V 'pray on'	<i>prédation</i> _N 'predation'	faire une prédation 'make a predation'	0 - 1	118 - 7,004	1897 - 1800
<i>additer</i> _V 'add'	<i>addition</i> _N 'addition'	faire une addition 'make an addition'	0 - 1	2 - 40,602	1804 - 1800
<i>combuster</i> _V 'combust'	<i>combustion</i> _N 'combustion'	subir une combustion 'undergo a combustion'	0 - 1	0 - 34,596	NA - 1800
<i>bijecter</i> _V 'relate one-to-one'	<i>bijection</i> _N 'bijection'	être en bijection 'be in bijection with'	0 - 1	1 - 1,313	NA - 1830
<i>superstitier</i> _V 'have a superstition'	<i>superstition</i> _N 'superstition'	avoir de la superstition 'have superstition'	0 - 1	0 - 14,870	NA - 1800
<i>inflexer</i> _V 'bow'	<i>inflexion</i> _N 'bowing'	avoir une inflexion 'have an inflection'	0 - 1	1 - 8,821	NA - 1800
<i>percepter</i> _V 'perceive'	<i>perception</i> _N 'perception'	utiliser sa perception 'make use of perception'	0 - 1	2 - 160,362	NA - 1800
<i>intuiter</i> _V 'guess'	<i>intuition</i> _N 'intuition'	utiliser son intuition 'make use of intuition'	0 - 1	201 - 52,941	1860 - 1800
<i>inoxyder</i> _V 'inoxidize'	<i>inoxydable</i> _A 'stainless'	rendre inoxydable 'make stainless'	0 - 1	8 - 13,708	NA - 1830
<i>infalsifier</i> _V 'unfalsify'	<i>infalsifiable</i> _A 'unfalsifiable'	rendre infalsifiable 'make tamper-proof'	0 - 0	2 - 776	NA - 1800

potential lexicon aims to fill the gaps in the French derivational system. Among the listed forms, some were attested in Old or Middle French, such as *ablater*_V 'remove' or *bijecter*_V 'relate one-to-one', before disappearing from the lexicon. This dictionary is echoed by Corbin's (1987: 733ff.) lists of possible words, which includes verbs such as *délater*_V 'denounce' or *additer*_V 'add'.

By (re)creating the verbs in Table 1, contemporary French speakers are driven by the same motivations as Matthias Kramer, namely the reconstitution of the verb that is the base of a noun in *-ion* or an adjective in *-able* so that the relation between the two is as transparent as possible. In some cases, such as *inflexer*_V 'bow', the back-formation creates

a doublet of an attested verb (i.e., *infléchir*_V ‘bow’) which is formally more distant from the lexeme that motivates the reformation (i.e., *inflexion*_N ‘bowing’). On this point, see Lignon et al. (2014).

3.2.3 Inventory of back-formed verbs

To our knowledge, there is no systematic inventory of French back-formed lexemes. As a result, while our conviction is that the different types of back-formation discussed in the rest of the article cover the main configurations of back-formation in the French lexicon, we have no way of testing this hypothesis against an already established dataset.

Some of the criteria mentioned above, such as date of first occurrence, could be used to extract candidate pairs of back-formed lexemes and their base from machine-readable dictionaries such as GLAWI (Sajous & Hathout 2015; Hathout & Sajous 2016) or Kaikki (Ylonen 2022). In practice, this criterion is not effective, due to the lack of systematic recording of first attestation dates in the etymological sections and the difficulty of parsing them.

A more effective solution is to use meaning, in this case the morphological definitions found in dictionaries (Martin 1983). For example, in the French Wiktionary, one of the meanings of *intuiter*_V ‘guess’ is defined in relation to *intuition*_N ‘intuition’ (22), but no meaning of *intuition*_N ‘intuition’ is defined in relation to *intuiter*_V ‘guess’. The opposite happens with regular verb-noun derivation: For example, one meaning of *adoption*_N ‘adoption’ is defined in relation to *adopter*_V ‘adopt’ (23), but no meaning of *adopter*_V ‘adopt’ is defined in relation to *adoption*_N ‘adoption’. This difference in definition orientation can be exploited by using a lexicon like Glawinette (Hathout et al. 2020) or GLeMM (Hathout et al. 2025).

(22) Intuiter: Faire appel à l’intuition (French Wiktionary)
 ‘Guess: To call upon intuition’

(23) Adoption: Action d’adopter (French Wiktionary)
 ‘Adoption: The act of adopting’

Using one of these resources, back-formed verb acquisition consists in selecting the pairs in which the first lexeme (*intuiter*_V ‘guess’) is defined relative to the second (*intuition*_N ‘intuition’) and whose alternation pattern is present in both orientations. For example, the alternation pattern of *intuiter*_V ‘guess’/*intuition*_N ‘intuition’ is “+er/+ion” while that of *adoption*_N ‘adoption’/*adopter*_V ‘adopt’ is “+ion/+er”, where “+” represents the stem of the lexemes. The number of pairs equipped with each of these alternation patterns enables us to identify the “default” and the “atypical” orientations. For example, in GLeMM-2.0, 54 pairs have the “+ion/+er” pattern and 47 the “+er/+ion” pattern. The pairs in the second group can be considered as candidates whose first lexeme is potentially back-formed. They include the examples in (24), where verbs are clearly back-formed.

- (24) *adducter*_V ‘adduce’ / *adduction*_N ‘adduction’
*antéverser*_V ‘antevert’ / *antéversion*_N ‘anteversion’
*asserter*_V ‘assert’ / *assertion*_N ‘assertion’
*circonlocuter*_V ‘circumlucate’ / *circonlocution*_N ‘circumlocution’
*compétiter*_V ‘compete’ / *compétition*_N ‘competition’
*constricter*_V ‘constrict’ / *constriction*_N ‘constriction’
*contracepter*_V ‘conceit’ / *contraception*_N ‘contraception’
*convulser*_V ‘seizure’ / *convulsion*_N ‘seizure’
*diffractioner*_V ‘diffract’ / *diffraction*_N ‘diffraction’
*digresser*_V ‘digress’ / *digression*_N ‘digression’
*décocter*_V ‘decoct’ / *décoction*_N ‘decoction’
*défecter*_V ‘defect’ / *défection*_N ‘defection’
*dépléter*_V ‘deplete’ / *déplétion*_N ‘depletion’
*effracter*_V ‘break and enter’ / *effraction*_N ‘breaking and entering’
*exciser*_V ‘excise’ / *excision*_N ‘excision’
*excréter*_V ‘excrete’ / *excrétion*_N ‘excretion’
*expanser*_V ‘expand’ / *expansion*_N ‘expansion’
*généuflexer*_V ‘genuflect’ / *généuflexion*_N ‘genuflection’
*hydrocuter*_V ‘be a victim of hydrocution’ / *hydrocution*_N ‘hydrocution’
*impulser*_V ‘impulse’ / *impulsion*_N ‘impulsion’
*intersecter*_V ‘intersect’ / *intersection*_N ‘intersection’
*intuiter*_V ‘guess’ / *intuition*_N ‘intuition’
*liquater*_V ‘liqueate’ / *liquation*_N ‘liquation’
*léser*_V ‘lesion’ / *lésion*_N ‘lesion’
*mondoviser*_V ‘worldwide telecast’ / *mondovision*_N ‘worldwide telecast’
*perfuser*_V ‘perfuse’ / *perfusion*_N ‘perfusion’
*prédater*_V ‘prey on’ / *prédation*_N ‘predation’
*préempter*_V ‘preempt’ / *préemption*_N ‘preemption’
*péréquater*_V ‘balance out’ / *péréquation*_N ‘balancing out’
*rotater*_V ‘rotate’ / *rotation*_N ‘rotation’
*rédempter*_V ‘redeem’ / *rédemption*_N ‘redemption’
*rétroverser*_V ‘turn’ / *rétroversion*_N ‘retroversion’
*secréter*_V ‘secrete’ / *secrétion*_N ‘secretion’
*translater*_V ‘translate’ / *translation*_N ‘translation’
*téléprospector*_V ‘telemarket’ / *téléprospection*_N ‘telemarketing’
*téléviser*_V ‘televise’ / *télévision*_N ‘television’
*émulser*_V ‘emulate’ / *émulsion*_N ‘emulation’
*érupter*_V ‘erupt’ / *éruption*_N ‘eruption’
*évictér*_V ‘evict’ / *éviction*_N ‘eviction’

4. Back-formed verbs in existing paradigms

As already mentioned, the term *back-formation* covers a variety of word-formation configurations. The simplest are likely to be verbs like *lacter*_V ‘produce milk’, *allocuter*_V

‘allocute’, and *délater*_V ‘denounce’. These verbs are all back-formed, according to the criteria presented in §2. All three are missing in dictionaries such as the TLF; they are not attested in the Google Books Ngram corpus; and only *délater*_V ‘denounce’ appears in the frCOW corpus. In contrast, the corresponding predicate nouns *lactation*_N ‘lactation’, *allocution*_N ‘speech’, and *délation*_N ‘informing’ are included in the TLF word list, and are attested in both the Google Books Ngram corpus and the frCOW corpus.

4.1 *Lacter*

In a paradigm-based framework, the formation of *lacter*_V ‘produce milk’ takes place within the paradigmatic family of the noun *lactation*_N ‘lactation’. This paradigmatic family contains the lexemes in (25).

(25) *lait*_N ‘milk’, *lactation*_N ‘lactation’, *lactaire*_A ‘which produces milk’

As indicated above, a paradigmatic family is a semantically coherent subset of a derivational family, and can therefore be relatively small. For example, the paradigmatic family of a predicate like *lactation*_N ‘lactation’ is made up of lexemes describing the typical participants in the event denoted by the noun, and its relational adjective. This is why the adjectives *lactique*_A ‘lactic’ and *lactescent*_A ‘milk-white’, meaning ‘white as milk’, for example, do not belong to the paradigmatic family of *lactation*_N ‘lactation’. The same goes for derived words like *lactose*_N ‘lactose’, *lactate*_N ‘lactate’, *lactémie*_N ‘lactaemia’, and compounds like *lactifère*_A ‘lactiferous’ and *lactosérum*_N ‘whey’.

The paradigmatic family of *lactation*_N ‘lactation’ is included in a paradigm P1, an excerpt of which is given in Table 2. Three of the families in this paradigm contain a verb (Pred_V) whose meaning is the same as that of the predicative noun in *-ation* (Pred_N). The presence of the Pred_V cell is expected as verbs are the most appropriate word class for expressing predicates (Croft 1991). The inclusion of the family of *lactation*_N ‘lactation’ into paradigm P1 brings into play the predictive power of P1 resulting in a kind of regularization of the family. The form of the verb *lacter*_V ‘produce milk’, predicted by the paradigm, yields the same semantic and formal contrast between the contents of the Pred_V and Pred_N cells as in the other families.

The analysis of the back-formation of *lacter*_V ‘produce milk’ within a paradigmatic framework therefore consists in (i) identifying the morphological paradigm that contains the paradigmatic family of the noun *lactation*_N ‘lactation’, and (ii) inserting the form predicted by the paradigm into the empty verb cell.

4.2 *Allocuter*

The analysis of the back-formed verb *allocuter*_V ‘allocute’ is slightly different from that of *lacter*_V ‘produce milk’. It is carried out within the paradigmatic family of *allocution*_N ‘speech’ (26). This family is part of a paradigm P2 presented in Table 3, which predicts the attested form of the verb.

Table 2. Excerpt from paradigm P1 of the paradigmatic family of *lactation*_N ‘lactation’. In each family, lexemes are connected to each other by the same form and meaning relations.

Obj_N	Pred_V	Pred_N	Relation_A
<i>lait</i> _N ‘milk’		<i>lactation</i> _N ‘lactation’	<i>lactaire</i> _A ‘milk-producing’
<i>onde</i> _N ‘wave’	<i>onduler</i> _V ‘wave’	<i>ondulation</i> _N ‘undulation’	<i>ondulaire</i> _A ‘which produces waves’
<i>argument</i> _N ‘argument’	<i>argumenter</i> _V ‘argue’	<i>argumentation</i> _N ‘arguments’	<i>argumentaire</i> _A ‘of arguments’
<i>glace</i> _N ‘ice’	<i>glacer</i> _V ‘ice’	<i>glaciation</i> _N ‘glaciation’	<i>glaciaire</i> _A ‘of glaciation’

Table 3. Excerpt from paradigm P2.

Pred_V	Pred_N	Agent_N.M	Agent_N.F	Agentive_A
<i>exécuter</i> _V ‘execute’	<i>exécution</i> _N ‘execution’	<i>exécuteur</i> _N ‘executioner.M’	<i>exécutrice</i> _N ‘executioner.F’	<i>exécutif</i> _A ‘executive’
<i>persécuter</i> _V ‘persecute’	<i>persécution</i> _N ‘persecution’	<i>persécuter</i> _N ‘persecutor.M’	<i>persécutrice</i> _N ‘persecutor.F’	<i>persécutif</i> _A ‘which causes a feeling of persecution’

- (26) *allocation*_N ‘speech’, *allocuteur*_N ‘he who gives a speech’, *allocutrice*_N ‘she who gives a speech’, *allocutif*_A ‘related to a speech’, *allocutaire*_N ‘person for whom a speech is intended’

Alternatively, family (26) could also partly fit into the paradigm P3 of Table 4, whose size (i.e., the number of its paradigmatic families) is significantly larger than that of P2. However, the verb form predicted by P3, *allocuer*_V, is not attested in the Google-indexed web pages (queried on 2024-04-12).

Two possible reasons may explain why speakers prefer to insert (26) into P2 rather than P3: First, the form *allocuter*_V ‘allocute’ predicted by P2 is more similar to the forms of the lexemes in (26); and second, the lexemes in (26) end in [kysjō], [kytær], [kytris], [kytif], and these endings are shared with the corresponding members of the other P2 families. As a consequence, (26) is formally closer to the families of P2 than to those of P3.

The analysis of *allocuter*_V ‘allocute’ is therefore the same as that of *lacter*_V ‘produce milk’. The only difference is that two paradigms are available here to accommodate the *allocation*_N ‘speech’ paradigmatic family. The form of the verb shows that speakers prefer the paradigm able to accommodate a larger number of lexemes in the family and to host the paradigmatic families that are most formally similar to the family of *allocation*_N ‘speech’. On the other hand, paradigm size in terms of number of families, and lexeme frequency within these families do not appear to be determining factors in the selection of the paradigm.

Table 4. Excerpt from paradigm P3.

Pred_V	Pred_N	Agentive_A
<i>évoluer</i> _V ‘progress’	<i>évolution</i> _N ‘progression’	<i>évolutif</i> _A ‘progressing’
<i>constituer</i> _V ‘constitute’	<i>constitution</i> _N ‘constitution’	<i>constitutif</i> _A ‘constitutive’
<i>rétribuer</i> _V ‘remunerate’	<i>rétribution</i> _N ‘remuneration’	<i>rétributif</i> _A ‘remunerative’
<i>distribuer</i> _V ‘distribute’	<i>distribution</i> _N ‘distribution’	<i>distributif</i> _A ‘distributive’

Table 5. Excerpt from paradigm P4.

Pred_V	Pred_N	Agent_N.M	Agent_N.F
<i>ventiler</i> _V ‘ventilate’	<i>ventilation</i> _N ‘ventilation’	<i>ventilateur</i> _N ‘he who ventilates’	<i>ventilatrice</i> _N ‘she who ventilates’
<i>gesticuler</i> _V ‘gesticulate’	<i>gesticulation</i> _N ‘gesticulation’	<i>gesticulateur</i> _N ‘gesticulator.M’	<i>gesticulatrice</i> _N ‘gesticulator.F’
<i>réguler</i> _V ‘regulate’	<i>régulation</i> _N ‘regulation’	<i>régulateur</i> _N ‘regulator.M’	<i>régulatrice</i> _N ‘regulator.F’
<i>isoler</i> _V ‘isolate’	<i>isolation</i> _N ‘isolation’	<i>isolateur</i> _N ‘isolator.M’	<i>isolatrice</i> _N ‘isolator.F’

4.3 Délater

The back-formation of the verb *délater*_V ‘denounce’ presents a similar situation. The paradigmatic family (27) of the noun *délation*_N ‘informing’ fits into the paradigm P2 (Table 2), enabling the prediction of the attested form *délater*_V ‘denounce’. But (27) could also fit into paradigm P4 (Table 5), which not only has a much larger size, but also contains numerous families in [lasjō], [latœr], [latris] whose forms are close to those of (27). However, the inclusion of (27) into P4 results in the prediction of a verb whose (infinitive) form would be *déler*. This form is suboptimal compared to *délater*_V ‘denounce’ for two reasons: (i) it is too short, and (ii) it is not sufficiently similar to *délation*_N ‘informing’ to enable its identification as the missing verb in the family (27).

(27) *délation*_N ‘informing’, *délateur*_N ‘informer.M’, *délatrice*_N ‘informer.F’

To sum up, we have analyzed three back-formed verbs in a paradigmatic framework. In all three cases (*lacter*_V ‘produce milk’, *délater*_V ‘denounce’, *allocuter*_V ‘allocute’), back-formation simply consists in including the family of the existing predicate noun (*lactation*_N ‘lactation’, *délation*_N ‘informing’, *allocution*_N ‘speech’) in a paradigm and then inserting the verb predicted by the paradigm into this family. When several paradigms are available, speakers seem to give priority to formal similarity (*délater*_V ‘denounce’, *allocuter*_V ‘allocute’), then to similarity with the lexemes of the families already present in the paradigm (*allocuter*_V ‘allocute’).

4.4 ParaDis

Morphological paradigms P1 to P4 all map onto (variants of) the same semantic paradigm, which connects a dynamic predicate, its action noun, agent nouns and agentive adjectives. This is what Fradin (2020) and Roché (2023) have called *action networks*. On the other hand, P1 to P4 correspond to different formal paradigms, all of which have in common a cell containing a form ending in [sjô].

The dissociation of formal and semantic paradigms in ParaDis enables the selection of the formal paradigm that best satisfies the well-formedness constraints for the three back-formed verbs. The selection is the result of a trade-off between verb stem minimum size and faithfulness (i.e., optimal similarity) to the content of other cells.

In the examples we analyzed, the separation of levels enables ParaDis to discard the most productive formal paradigms (those with the most families) and the most extensive ones (those with the largest families), and to favor instead (i) the formal paradigm corresponding to P1 to host the formal family of *lactation*_N ‘lactation’ because it predicts an optimal verb form in terms of faithfulness (the sequence [lakt] is immediately recognized since it is present in *lactose*_N ‘lactose’, *lactaire*_A ‘which produces milk’, *lactique*_A ‘lactic’), and (ii) the formal paradigm corresponding to P2, for the formal families of *délation*_N ‘informing’ and *allocution*_N ‘speech’ because the corresponding verb stem represents the best compromise between size ([delat] and not [del]), hiatus avoidance ([alokyt] and not [alokɥ]), and faithfulness to the forms of their family members.

5. Back-formation in a family at the crossroads of two paradigms

Prima facie, the relation between *intuition*_N ‘intuition’ and *intuiter*_V ‘guess’ in Table 6 (like that between *perception*_N ‘perception’ and *percepter*_V ‘perceive’) is analyzed in the same way as the relation between *délation*_N ‘informing’ and *délater*_V ‘denounce’: The verb fills a cell in a family that includes a noun in *-ion* and, for *intuition*_N ‘intuition’ and *perception*_N ‘perception’, an adjective in *-if*; the forms of the noun and adjective determine that of the verb; the first attestation of the verb is more recent than that of the other members of the family. As with *délater*_V ‘denounce’ and *allocuter*_V ‘allocute’, the formation of these verbs takes place within the families of the nouns *intuition*_N ‘intuition’ and *perception*_N ‘perception’ and within the paradigms that contain these families. Furthermore, the stem of the newly-formed verb (represented by X in the table header) is a sequence that maximizes formal similarity to the family members and, consequently, ensures the best formal transparency of the relations with them. Moreover, in the case of *percepter*_V ‘perceive’, the verb is a doublet of *percevoir*_V ‘perceive’ which has been attested for a long time – in all its occurrences, *percepter*_V ‘perceive’ is synonymous with *percevoir*_V ‘perceive’. The effect of its back-formation is therefore to reduce the formal distance between the verb and the rest of the family.

The formal relations found in the “regular” families in Table 7 are exactly the same. These families are all rooted in a verb (first column), and include a noun in *-ion* (second column) and an adjective in *-if* (third column). They are formally aligned with the

Table 6. Paradigmatic families of *intuition*_N ‘intuition’ and *perception*_N ‘perception’.

X_V	Xion_N	Xif_A
<i>intuiter</i> _V ‘make use of one’s faculty of intuition’	<i>intuition</i> _N ‘intuition’	<i>intuitif</i> _A ‘intuitive’
<i>perceper</i> _V ‘make use of one’s faculty of perception’	<i>perception</i> _N ‘perception’	<i>perceptif</i> _A ‘perceptive’

Table 7. Paradigm made up of families formally aligned with the families of *intuiter*_V ‘guess’ and *perceper*_V ‘perceive’. These families form action networks.

X_V “@”	Xion_N “act of @”	Xif_A “which @”
<i>prospecter</i> _V ‘prospect’	<i>prospection</i> _N ‘prospection’	<i>prospectif</i> _A ‘prospective’
<i>exécuter</i> _V ‘execute’	<i>exécution</i> _N ‘execution’	<i>exécutif</i> _A ‘executive’
<i>inventer</i> _V ‘invent’	<i>invention</i> _N ‘invention’	<i>inventif</i> _A ‘inventive’
<i>adopter</i> _V ‘adopt’	<i>adoption</i> _N ‘adoption’	<i>adoptif</i> _A ‘adoptive’
<i>asserter</i> _V ‘assert’	<i>assertion</i> _N ‘assertion’	<i>assertif</i> _A ‘assertive’

families of *intuiter*_V ‘guess’ and *perceper*_V ‘perceive’. Yet, are they all part of the same paradigm?

A first observation is that the alignment of the families of *intuiter*_V ‘guess’ and *perceper*_V ‘perceive’ with those of the paradigm in Table 7 only holds at the formal level. The families in Table 7 belong to an action network: The noun in *-ion* is an action noun, and its semantic content is the same as that of the verb (e.g., *invention*_N ‘invention’ is the action of *inventer*_V ‘invent’). This is not the case with *intuition*_N ‘intuition’ and *perception*_N ‘perception’ which are cognitive faculties that characterize entities that are intuitive or perceptive. *Intuiter*_V ‘guess’, like *perceper*_V ‘perceive’, describes an event in which someone uses this faculty, either consciously (28) or not (29).

(28) Mais cela fait plusieurs décennies qu’Edgar Morin a *intuité* que l’hypermodernité est indissociable de la complexité.

‘But it’s been several decades since Edgar Morin intuited that hypermodernity is inseparable from complexity’

(<https://www.lesechos.fr/thema/dossiers/special-gestion-des-risques-lart-de-surmonter-des-crisis-simultanees-1902510>)

(29) Oui d’accord, mais si personne n’est là pour *perceper* ce son, qui nous dit qu’il existe vraiment ?

‘OK, but if no one is there to perceive the sound, who’s to say it really exists?’

(<https://www.jeuxvideo.com/forums/1-7926-14075908-2-0-1-0-0.htm>)

In other words, the function of verb back-formation is not to fill the cell corresponding to the **center** of an action network as in the case of *délater*_V ‘denounce’ and *allocuter*_V

Table 8. Paradigm containing families in which semantic relations form a cognition network.

Cog_N “@”	Pred_inst_V “use @”	Agentive_A “who has @”
<i>mémoire</i> _N ‘memory’	<i>mémoriser</i> _V ‘memorize’	
<i>pensée</i> _N ‘thought’	<i>penser</i> _V ‘think’	<i>pensif</i> _A ‘thoughtful’
<i>raison</i> _N ‘reason’	<i>raisonner</i> _V ‘reason’	<i>rationnel</i> _A ‘rational’

‘allocute’.¹ Moreover, the verb is not the central element of the families in Table 6: The center of the network is the noun of cognition, and the verb meaning (‘make use of intuition’, ‘make use of perception’) is built on the semantic content of this noun, interpreted as the faculty used in the mental process expressed by the verb.

This leads us to a second observation. The formal paradigm where the verbs in Table 6 are formed is also in correspondence with a larger morphological paradigm whose cells are connected by different semantic relations. Notably, the *-ion* suffix present in *intuition*_N ‘intuition’ and *perception*_N ‘perception’ is not normally used in French to form nouns of instruments used in the process described by the base verb – unlike *-eur* (*aspirateur*_N ‘vacuum cleaner’ ← *aspirer*_V ‘vacuum’) or *-oir* (*arrosoir*_N ‘watering can’ ← *arroser*_V ‘water’), for example.

Table 8 shows derivational families consisting of a noun denoting a cognitive faculty (Cog_N), an agentive adjective (Agentive_A), and an instrumental verb (Pred_inst_V) based on this noun through a regular derivational process. These families form cognitive networks consisting of interpredictable relations. The adjective denotes the possession of the faculty, e.g., they who are *pensif*_A ‘thoughtful’ have *pensées*_N ‘thoughts’, they who are *rationnel*_A ‘rational’ appeal to their *raison*_N ‘reason’. The verb denotes the use of the faculty: *Mémoriser*_V ‘memorize’ is to make use of one’s *mémoire*_N ‘memory’, and *raisonner*_V ‘reason’ is to make use of one’s *raison*_N ‘reason’. The use of the faculty implies its possession. Conversely, an entity that possesses a faculty is expected to use it.

The families hosting *intuiter*_V ‘guess’ and *perceper*_V ‘perceive’ are specific because, semantically, they belong to the same cognition network as the families in Table 8, but formally, they are aligned with the families in Table 7, which form action networks. The back-formation of these two verbs is therefore more complex than that of the verbs *délater*_V ‘denounce’ and *allocuter*_V ‘allocute’: It is the result of an operation that crosses two paradigms. On the one hand, we have a semantic paradigm that is part of a cognition network containing an explicit relation between a cognitive faculty and a verb predicate (30). On the other hand, we have a formal paradigm (31) where all the relations are interpredictable (the forms [Xtif] and [Xsjö] are correlated with each other, and they

¹ As discussed in more detail in §6.3.1, we call “center” a distinguished member in a paradigmatic family in relation to which all other members are defined. The center is often, but not necessarily, the derivational ancestor in its paradigmatic family.

both predict [Xt]), and that minimizes the formal distance between the contents of its cells.

In a model such as ParaDis where the semantic and formal descriptions of the families are dissociated, the back-formation of *intuiter*_V ‘guess’ and *perceper*_V ‘perceive’ can be analyzed as a paradigmatic hybridization: The families of these verbs correspond to a semantic paradigm and to a formal paradigm not normally associated with each other. Our hypothesis is that it is precisely this atypical combination that makes speakers feel that these verbs are marked and perceive them as “new” in their respective families.

(30) {“@cognitive_faculty”, “use @”, “who has @”}

(31) {[Xsjō], [Xt], [Xtif]}

6. Networks, centers and quasi-stolons

6.1 (Neoclassical) Compounding does not form verbs in French

A third type of back-formed verbs consists of verbs that have a neoclassical component in their form, such as *cryograver*_V ‘cryograph’, *hydromasser*_V ‘hydromassage’, and *cyberharceler*_V ‘cyberharass’. Namer (2012) analyses these verbs as back-formations of the corresponding suffixed predicative noun: *cryogravure*_N ‘cryograph’, *hydromassage*_N ‘hydromassage’, and *cyberharcèlement*_N ‘cyberharassment’. According to her analysis, *cryograver*_V ‘cryograph’ means ‘to perform a *cryogravure*_N ‘cryograph’’. The paradigmatic analysis of these verbs proposed in this section basically follows that of Namer (2012).

Namer (2012) motivates her analysis with a number of arguments, including the three following ones.

- (i) French verbs whose form contains a nominal or adjectival neoclassical component cannot be the result of a noun or adjective incorporation, because incorporation is not available in French. Note that “incorporation” refers to two phenomena. One is morphological, namely the formation of verbs composed of a verbal head and a nominal (or adjectival) component. The other, characterized as “semantic incorporation” or “pseudo-incorporation”, is semantic. The notion we are interested in here is morphological incorporation. According to Gerdtz (1998) and Massam (2009), among others, the French morphological system does not possess this construction mechanism. On the other hand, lexical semantics studies (Mathieu 2009; Hirschbühler & Labelle 2009; Knittel 2010) have shown that the semantic behavior of certain verb phrases in languages like French exhibits similarities to nominal incorporation.
- (ii) For any verb that appears to have a structure YX, where Y is a neoclassical component and X is a verb stem, the corresponding action or activity noun is always attested. The noun structure usually matches a pattern YXsfx, where sfx is a suffix used in the formation of verb-based nouns (e.g., *-ure* in *cryogravure*_N

- ‘cryograph’). Some verbs like *cyberattaquer*_V ‘cyberattack’ are related by conversion to a suffixless noun (*cyberattaque*_N ‘cyberattack’), therefore matching a YX pattern.
- (iii) In contrast, there are nouns with a YX or YXsfx structure for which no morphologically related verbs are attested nor are semantically possible. This is notably the case when the verb component X denotes an inaccusative predicate (e.g., an intransitive verb like *résister*_V ‘resist’ or *réagir*_V ‘react’ used with an inanimate subject) and the neoclassical component Y refers to the patient of the predicate. The semantic relation between X and Y corresponds to a predicate-argument relation. It is compatible with a nominal realization of the process, as in *cardiobattement*_N ‘cardiobeating’, *gastrorésistance*_N ‘gastroresistance’, *ophtalmoréaction*_N ‘ophtalmoreaction’, *neurocafoillage*_N ‘neuromess’, *thermoaccumulation*_N ‘thermoaccumulation’, and *hydroévaporation*_N ‘hydroevaporation’. On the other hand, this relation is syntactically incompatible with a YX verbal form of the process, because the subject position would already be saturated by the Y component. For example, the sentence (32a) cannot be used to express (32b). The same applies to *gastrorésister*_V ‘gastroresist’, *ophtalmoréagir*_V ‘ophtalmoreact’, etc.
- (32) a. *{Ça/il} cardiobat.
 ‘It is cardio-beating’
 b. Le cœur bat.
 ‘The heart is beating’

Namer (2012) provides a detailed analysis of the conditions governing the different argument or adjunct functions played by the neoclassical component Y with respect to X. One of her findings is that Y never fulfils the agent role of the verb predicate X but can realize other functions, such as means in *hydromasser*_V ‘hydromassage’ or place in *cyberharceler*_V ‘cyberharass’.

The impossibility to create verb compounds is not specific to French. As already mentioned in §2.1, Plag (2003: 155) points out that this is the case in English too, where it is not possible to express phrases like *read a book* or *steal a car* using compounds like *bookread*_V or *carsteal*_V. In English, compounding only forms nouns and adjectives. As a consequence, verbs in compounds like *proof-read*_V or *ghost-write*_V are analyzed as back-formed from the nouns *proof-reading*_N and *ghost-writer*_N (Plag 2003: 154).

6.2 Compounding produces subtypes

There is also a semantic explanation for the fact that verbs like *cryograver*_V ‘cryograph’ cannot be neoclassical compounds formed from a verb component (*graver*_V ‘engrave’). This explanation is correlated with the endocentric nature of these compounds, where endocentricity is understood in the sense of Bauer (2008) – see also Haspelmath (2002) and Booij (2005), presented in Scalise & Bisetto’s (2009) classification of compounds. According to Bauer (2008: 35), in English endocentric compounds, “one of the elements (typically the right-hand element) is the head of the construction. Headedness is

shown most clearly by hyponymy: the compound as a whole is a hyponym of its head.”² In other words, both standard and neoclassical endocentric compounds refer to subtypes (or hyponyms) of existing types (for a psycholinguistic account of hyponymy in compound words, see, e.g., Gagné et al. 2020). This is particularly true for French.

For example, from a nominal component like *journal*_N ‘newspaper’, neoclassical compounding forms nouns like *webjournal*_N ‘web newspaper’, *radiojournal*_N ‘radio newspaper’, *vidéojournal*_N ‘video newspaper’, and *écojournal*_N ‘environmental newspaper’. Each of these compounds denotes a type of newspaper, which in turn is seen as a (superordinate) type. The same applies to *cryogravure*_N ‘cryograph’, which is a type of *gravure*_N ‘engraving’, where *gravure*_N ‘engraving’ has a generic meaning, as in (33) and (34).

(33) La gravure est un art graphique.
‘Engraving is a graphic art form’

(34) La gravure est un art centenaire.
‘Engraving is a century-old art form’

In these examples, *gravure*_N ‘engraving’ refers to a generic activity, as does *cryogravure*_N ‘cryograph’. More generally, from a nominal component that denotes a process, neoclassical compounding produces nouns that refer to subtypes of that process and have the meaning of a generic activity.

If *cryograver*_V ‘cryograph’ were a compound verb formed from the verb *graver*_V ‘engrave’, a similar interpretation would be expected, with *cryograver*_V ‘cryograph’ referring to a way of engraving. However, *cryograver*_V ‘cryograph’ means engraving with a particular instrument, namely cold. Moreover, *cryograver*_V ‘cryograph’ does not refer to an activity, but to an action, i.e., to an event that is a specific occurrence of the generic activity denoted by *cryogravure*_N ‘cryograph’ (Haas et al. 2008; Haas & Huyghe 2010): *Cryograver X* means ‘to perform a *cryogravure*_N ‘cryograph’ on X’, as in examples (35b) and (36b). The verb *cryograver*_V ‘cryograph’ refers to an event that is grounded in space and time by the inflectional properties of its forms (tense, aspect, mode, person, gender, number, etc.) and by the argumental and discursive relations established in the syntactic contexts in which it occurs. The same is true for the verb *graver*_V ‘engrave’ in examples (35a) and (36a). Consequently, the verb *graver*_V ‘engrave’ does not refer to a type of which *cryograver*_V ‘cryograph’ could be a subtype, thereby providing a further argument against the analysis of *cryograver*_V ‘cryograph’ as a neoclassical compound formed from *graver*_V ‘engrave’.

² “L’un des éléments (typiquement l’élément le plus à droite) est la tête de la construction. La présence de la tête se manifeste clairement par l’hyponymie : le mot composé est un hyponyme de la tête”; see <https://www.wgtn.ac.nz/lals/about/staff/Publications/Bauer-EnglishExocentricCompounds.pdf>, for Bauer’s own translation.

- (35) a. Pierre grave un vase en cristal de Baccarat.
 ‘Pierre engraves a Baccarat crystal vase’
 b. Pierre cryograde un vase en cristal de Baccarat.
 ‘Pierre cryographs a Baccarat crystal vase’
- (36) a. Pierre grave le prénom de sa fille sur un médaillon.
 ‘Pierre engraves his daughter’s name on a locket’
 b. Pierre cryograde le prénom de sa fille sur un médaillon.
 ‘Pierre cryographs his daughter’s name on a locket’

6.3 Centers and stolons in the ParaDis model

6.3.1 Centers

We now turn to the paradigmatic account of the back-formation of French verbs with a neoclassical compound structure. We draw on the proposal of Roché (2023) and Fradin (2020) that paradigmatic families are organized into networks, and that some of these networks are verb-centered (i.e., verb-based), like the action networks (37), while others are noun-centered, like the fruit or activity networks (38)–(40).

- (37) *laver*_V ‘wash’, *lavoir*_N ‘wash house’, *lavage*_N ‘washing’, *laveur*_N ‘washer.M’, *laveuse*_N ‘washer.F’
- (38) *pomme*_N ‘apple’, *pommier*_N ‘apple tree’, *pommerie*_N ‘apple orchard’
- (39) *journal*_N ‘newspaper’, *journalisme*_N ‘journalism’, *journaliste*_N ‘journalist’
- (40) *natation*_N ‘swimming’, *nageur*_N ‘swimmer.M’, *nageuse*_N ‘swimmer.F’, *nager*_V ‘swim’, *nage*_N ‘swimming stroke’

In (37), the family of *laver*_V ‘wash’ is organized as an action network whose other members include the nominal expression of the verb meaning (i.e., the action noun *lavage*_N ‘washing’) and verb arguments (agent, place). By contrast, the family of *pomme*_N ‘apple’ (38) is part of a verbless network organized around the fruit noun: The other members in the family are nouns semantically defined in relation to *pomme*_N ‘apple’, either directly as *pommier*_N ‘tree that produces apples’ or indirectly as *pommerie*_N ‘orchard planted with apple trees’. Similarly, the family of *journal*_N ‘newspaper’ (39) is centered around the noun *journal*_N ‘newspaper’, which, in its “information content” facet, denotes the purpose of *journalisme*_N ‘journalism’ (on the notion of facets, see, e.g., Cruse 2004: 114).

Family (40) is centered around the activity noun *natation*_N ‘swimming’. In this family, a *nageuse*_N ‘swimmer.F’ is “a woman who practices swimming”, not “a woman who is swimming”; in this same family, the meaning of *nager*_V ‘swim’ is “to practice swimming”

(42a). Besides, family (40) contains the noun *nage*_N ‘swimming stroke’, which refers to a way of swimming (42b). The verb *nager*_V ‘swim’ may also refer to an event in the action network (41). In this case, the verb is the center of the network. The action network (41) also includes another reading of *nage*_N ‘swimming’, which is an action noun used in contexts such as (42c) and (42d).

(41) *nager*_V ‘swim’, *nage*_N ‘swimming’, *nageur*_N ‘swimmer.M’, *nageuse*_N ‘swimmer.F’

- (42) a. Elles nagent quatre heures par jour en moyenne.
 ‘They swim an average of four hours a day’
 (<https://www.lacote.ch/valaud/la-cote/nyon-district/une-traversee-de-la-manche-a-vocation-humanitaire-555824>)
- b. La brasse est la nage la plus pratiquée.
 ‘Breaststroke is the most practised way of swimming’
 (https://www.guide-piscine.fr/apprendre-a-nager/differentes-nages-brasse-crawl-dos-papillon-variantes-natation-piscine-apprendre-410_A)
- c. Il est important de faire attention à votre posture pendant la nage.
 ‘It is important to pay attention to your posture while swimming’
 (<https://oppq.qc.ca/blogue/natation-6-conseils-un-physiotherapeute/>)
- d. La nage a duré 13 heures, 50 minutes et 47 secondes pour le pré-adolescent.
 ‘The swim lasted 13 hours, 50 minutes and 47 seconds for the pre-teen’
 (<https://petitesannoncesafrique.com/nage-sainte-lucie-martinique/>)

6.3.2 *Stolons*

In §6.2, we mentioned the example of *webjournal*_N ‘web newspaper’, which we consider to be a compound in the same way as *écojournal*_N ‘environmental newspaper’, *radiojournal*_N ‘radio newspaper’, and so on. Like its base component *journal*_N ‘newspaper’ (39), *webjournal*_N ‘web newspaper’ belongs to a family centered around the noun denoting the object the activity is about (43). Examples (39) and (43) clearly show that the families of *journal*_N ‘newspaper’ and *webjournal*_N ‘web newspaper’ are structured in the same way and belong to the same paradigm, as illustrated in Figure 1.

(43) *webjournal*_N ‘web newspaper’, *webjournalisme*_N ‘webjournalism’, *webjournaliste*_N ‘webjournalist’

This means that the nominal compounding illustrated in Figure 1 produces a kind of copy of the source family. We call this type of construction a **stolon**, by analogy with the way certain plants reproduce. (In botany, a stolon is a plant stem enabling a propagation strategy that consists of producing a set of individuals by cloning from a mother plant.) A stolon therefore creates a relation between two paradigmatic families. Such an analysis is possible once we consider that derivational morphology forms families from families, rather than (complex) lexemes from (simpler) lexemes. This hypothesis assumes a paradigmatic approach to derivation.

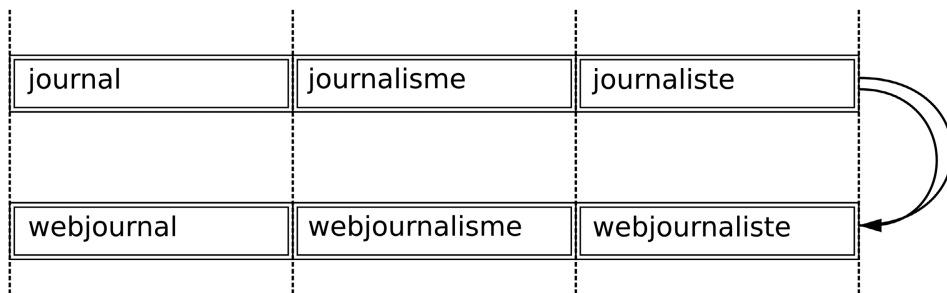


Figure 1: Compounding in *web-X* creates a paradigmatic family within the paradigm of the component X. The relation between the families of X and *web-X*, represented by a double arrow, is called a stolon.

The situation is essentially the same for some prefixed verbs. For example, consider the family of the verb *programmer*_V ‘schedule’ as in (44) and that of *déprogrammer*_V ‘cancel’ (that is, “unschedule” some show, for instance) as in (45), where *déprogrammeur*_N ‘canceler.M’ and *déprogrammatrice*_N ‘canceler.F’ are interpreted as a person who carried out a (single) occurrence of *déprogrammation*_N ‘cancelling’ and not as a person whose activity is to cancel shows. An example of a sentence containing the noun *déprogrammeur*_N ‘canceler.M’ could be as in (46).

(44) *programmer*_V ‘schedule’, *programmation*_N ‘scheduling’, *programmeur*_N ‘scheduler.M’, *programmatrice*_N ‘scheduler.F’

(45) *déprogrammer*_V ‘cancel’, *déprogrammation*_N ‘cancelling’, *déprogrammeur*_N ‘canceler.M’, *déprogrammatrice*_N ‘canceler.F’

(46) Le déprogrammeur, c’est lui. (Celui qui a déprogrammé l’émission, c’est lui.)
‘He is the canceler (He is the one who cancelled the show)’

The paradigmatic family of *programmer*_V ‘schedule’ makes up an action network. The relations within this network are the same as those in the family of *déprogrammer*_V ‘cancel’. As a consequence, both families fit into the same paradigm. We can then consider them to be linked by a stolon (Figure 2).

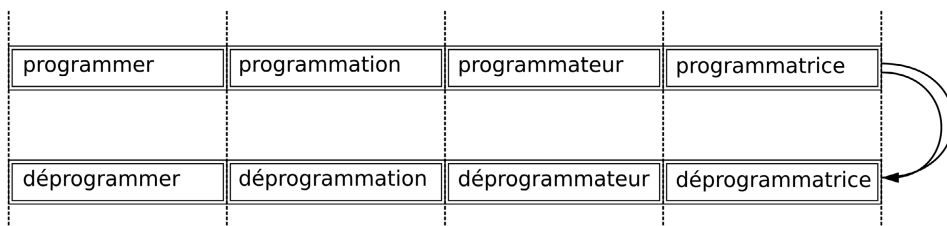


Figure 2: Prefixation in *dé-* creates a family in the same paradigm as the source family. The two families are connected by a stolon (double arrow).

6.3.3 A paradigmatic analysis of verbs structured as neoclassical compounds: stolons or quasi-stolons?

We have seen in §6.2 that the family of the neoclassical compound *cryogravure*_N ‘cryograph’, repeated in (47), is noun-centered (i.e., around *cryogravure*_N ‘cryograph’), whereas the source family repeated in (48) is verb-centered around *graver*_V ‘engrave’, which denotes an event. The two families are therefore not organized in the same way: Their cells are not connected by the same semantic relations, consequently the two families do not belong to the same paradigm. The same remarks can be made when comparing any other neoclassical compound family (e.g., the families of *hydromassage*_N ‘hydromassage’ or *cyberharcèlement*_N ‘cyberharassment’) with the corresponding source family (e.g., those of *masser*_V ‘massage’ and *harceler*_V ‘harass’, respectively). In all these cases, neoclassical compounding creates families within paradigms that are part of semantic networks distinct from those of the source families.

(47) *cryogravure*_N ‘cryograph’, *cryographeur*_N ‘cryographer.M’, *cryographeuse*_N ‘cryographer.F’

(48) *graver*_V ‘engrave’, *gravure*_N ‘engraving’, *graveur*_N ‘engraver.M’, *graveuse*_N ‘engraver.F’

In contrast to the families shown in Figures 1 and 2, the family of *cryograver*_V ‘cryograph’ is not connected to the family of *graver*_V ‘engrave’ by a (true) stolon. The two families are nonetheless formally aligned. They only differ at the semantic level. Similarly, the families of *hydromassage*_N ‘hydromassage’ and *masser*_V ‘massage’, and *cyberharcèlement*_N ‘cyberharassment’ and *harceler*_V ‘harass’ are aligned formally, but not semantically. In all these cases, neoclassical compounding duplicates families on the formal level, while shifting semantic paradigms.

In a model like ParaDis, where the formal and semantic levels are separated, the formal families of *cryogravure*_N ‘cryograph’ and *graver*_V ‘engrave’ belong to the same formal paradigm; this is also true for the families of *hydromassage*_N ‘hydromassage’ and *masser*_V ‘massage’, and those of *cyberharcèlement*_N ‘cyberharassment’ and *harceler*_V ‘harass’. They can therefore be considered to be connected by a formal stolon. On the other hand, the semantic families of *cryogravure*_N ‘cryograph’ (resp. *hydromassage*_N ‘hydromassage’, *cyberharcèlement*_N ‘cyberharassment’) and *graver*_V ‘engrave’ (resp. *masser*_V ‘massage’, *harceler*_V ‘harass’) belong to two separate paradigms, one centered on the noun and the other on the verb. We call the relation between these morphological families **quasi-stolons**, since the families are aligned only formally (Figure 3).

Once the creation of the family of *cryogravure*_N ‘cryograph’ has been initiated, it is filled in the same way as any other family in an activity network centered on an activity noun. We have seen that this family includes a cell for a verb having the meaning ‘perform the activity’. In the family of *cryogravure*_N ‘cryograph’, the realization of this verb, namely *cryograver*_V ‘cryograph’, involves the same formal mechanisms as those used to form the verb *sépulter*_V ‘bury’ in the family of *sépulture*_N ‘burial’ (5). The overall

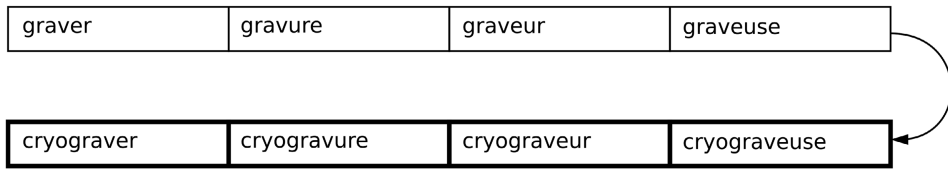


Figure 3: The creation of the family of *cryograver_V* ‘cryograph’ involves a semantic paradigm shift. The families of *graver_V* ‘engrave’ and *cryograver_V* ‘cryograph’ are aligned formally, but not semantically since the family of *graver_V* ‘engrave’ forms an action network, while the family containing *cryograver_V* ‘cryograph’ forms an activity network. The relation between the two families is a quasi-stolon represented by a single arrow.

mechanism is the same as the one described in §4.1 for the formation of *lacter_V* ‘produce milk’ in the family started by *lactation_N* ‘lactation’. It is based on the following principles:

- (i) a family is initiated by the realization of one of its members;
- (ii) a family necessarily belongs to a paradigm;
- (iii) form and meaning relations between the members of a family are known and predicted by its paradigm.

7. A case of deadjectival back-formation

The cases of verb back-formation discussed in the previous three sections involve families containing nouns but lacking a verb: *Lacter_V* ‘produce milk’ is inserted in a family belonging to a well-established paradigm that is part of an action network; *intuiter_V* ‘guess’ completes a “blended” family mixing of a semantic paradigm expressing a cognitive network and a formal paradigm usually associated with an action network; *cryograver_V* ‘cryograph’ is added into a family formed as a quasi-stolon, which combines an activity network with an action network. The fourth type of back-formation we consider in this article is a case of adjective-verb relation which is both less common and slightly more complex. It involves the formation of verbs such as *inoxyder_V* ‘inoxidize’, which means ‘make non-oxidizable, or stainless’ (49)–(51), and *infalsifier_V* ‘unfalsify’, which means ‘make tamper-proof’ (52)–(53).

- (49) Où faire inoxyder des lames de couteaux ?
 ‘Where to get knife blades inoxidized?’
 (<https://www.abcelectronique.com/recherche/?db=forum&np=10&q=sous+traitant>)
- (50) Inoxyder les guidons.
 ‘Inoxidize the handlebars’
 (Céline, *Mort à crédit*, 1936, p. 406)

- (51) Ainsi, le chrome ajouté pour inoxyder certains aciers, se retrouve dans la scorie inox ou ALZ.
 ‘For example, chromium added to oxidize certain steels is found in stainless slags or ALZ’
 (<https://docplayer.fr/212831627-Granulats-secondaires-sous-produits-d-acierie.html>)
- (52) Vous pouvez l’utiliser pour infalsifier des documents et protéger vos documents critiques.
 ‘You can use it to make documents tamper-proof and protect your critical documents’
 (<https://www.guru99.com/fr/google-drive-alternatives.html>)
- (53) Connecting Software s’associe à Cryptowerk pour lancer une solution complémentaire permettant de sceller et d’infalsifier des documents dans des logiciels commerciaux standard grâce à la technologie Blockchain.
 ‘Connecting Software partners with Cryptowerk to launch a complementary solution for sealing and make documents tamper-proof in standard business software using Blockchain technology’
 (<https://www.connecting-software.com/fr/press-release/>)

First, we can rule out the possibility that *inoxyder*_V ‘inoxidize’ is formed by prefixing the verb *oxyder*_V ‘oxidize’ with *in-*. In French, prefixation in *in-* almost exclusively forms adjectives from adjectives, and only marginally forms nouns from nouns, such as *irrespect*_N ‘disrespect’, *inaction*_N, ‘inaction’, or *inexpérience*_N ‘inexperience’. By contrast, no French verbs are prefixed with *in-* (Apothéloz 2003). Verb negative prefixation is realized using other prefixes, such as *dé-* (e.g., *déservir*_V ‘disserve’).

7.1 Paradigmatic context

The formation of the verb *inoxyder*_V ‘inoxidize’ involves three paradigmatic families. The first is the initial family shown in (54).³ It is part of a well-established paradigm where we find families like (55).

- (54) *oxyder*_V ‘oxidize’, *oxydation*_N ‘oxidization’, *oxydateur*_N ‘oxidizer.M’, *oxydable*_A ‘oxidizable’
- (55) *dégrader*_V ‘degrade’, *dégradation*_N ‘degradation’, *dégradateur*_N ‘degradator.M’, *dégradable*_A ‘degradable’

³ Family (54) also contains the noun *oxyde*_N ‘oxide’, which plays the role of means in the process denoted by *oxyder*_V ‘oxidize’ (Fradin 2020).

The adjective *oxydable*_A ‘oxidizable’ is also part of a second paradigmatic family, shown in (56), which belongs to an adjectival paradigm formed by adjectives and their quality nouns as in (57). The fact that *oxydable*_A ‘oxidizable’ belongs to both families (54) and (56) is not a problem, since nothing prevents two derivational paradigms from overlapping (Beecher 2004).

(56) *oxydable*_A ‘oxidizable’, *oxydabilité*_N ‘oxydizability’

(57) a. *fidèle*_A ‘faithful’, *fidélité*_N ‘faithfulness’
 b. *facile*_A ‘easy’, *facilité*_N ‘easiness’

Paradigm (57) corresponds to what Fradin (2020) calls “metapatterns” and what Sanacore (2023) and Sanacore et al. (2024) call “fringes”. It is characterized by the fact that in these families, the adjective and the noun share the same semantic content in the sense of Croft (1991) (see also Spencer 2013 on adjective-noun transposition).

Moreover, families of paradigm (57) are hardly constrained: With the notable exception of noun-based relational adjectives (e.g., *barbu*_A ‘bearded’), a very large number of adjectives are paired to a quality noun (Roché 2006; Koehl 2012; Strnadová 2014). In addition, the noun mainly has a discourse function: The relation in (57) enables a meaning normally realized by an adjective to be realized using a noun (Croft 1991).

The third family is (58), resulting from prefixing the members of family (56) with *in-*. Families (56) and (58) belong to the same paradigm – *inoxydabilité*_N ‘non-oxidizability’ is the quality of being *inoxydable*_A ‘stainless’, as *oxydabilité*_N ‘oxidizability’ is the quality of being *oxydable*_A ‘oxidizable’. They are therefore connected by a stolon.

(58) *inoxydable*_A ‘stainless’, *inoxydabilité*_N ‘non-oxidizability’

7.2 Analysis

The starting point for the word-formation process leading to the verb *inoxyder*_V ‘inoxidize’ is the need to express the meaning ‘make stainless’. This meaning must be included in a family that also includes the adjective *inoxydable*_A ‘stainless’ to ensure the correct interpretation of the verbal lexeme. This family must therefore be part of a paradigm containing a series of adjectives denoting a property “PPT” and a series of change-of-state verbal predicates that can be glossed as “make PPT”. The corresponding semantic paradigm looks like (59), where “@” stands for the change-of-state predicate, “@ACT” for the corresponding action noun, and “@AGT” for the agent noun. The paradigm includes families realized as in (60) or (61). In these families, the predicate @ can be realized by a verb in *-iser* like *banaliser*_V ‘banalize’ in (60) or by a convert like *lisserv*_V ‘smooth’ in (61). Among the semantic families included in (59), we have (62) where the value of PPT is ‘stainless’. The question then is to associate a formal family to the

semantic family (62) knowing that the formal value in correspondence with PPT is [inɔksidabl].

- (59) {“PPT”, “@ = make PPT”, “@ACT”, “@AGT”}
- (60) *banal*_A ‘common’, *banaliser*_V ‘banalize’, *banalisation*_N ‘banalization’, *banalisateur*_N ‘banalizer.M’
- (61) *lisse*_A ‘smooth’, *lisser*_V ‘smooth’, *lissage*_N ‘smoothing’, *lisseur*_N ‘smoother.M’
- (62) {‘stainless’, ‘make stainless’, ‘act of making stainless’, ‘he who makes stainless’}

Speakers have several options for formally realizing the meanings of the verb of the semantic family (62). We can first rule out conversion, since this process does not seem to be available in French for the formation of verbs from adjectives in *-able*. One option would be to create a family within the most regular formal paradigm, namely that of the verbs suffixed with *-iser*. The presence of the suffix *-iser* results in a more canonical derivation in the sense of Corbett (2010). The paradigmatic family corresponding to (62) would then be (63).

- (63) *inoxydable*_A ‘stainless’, *inoxydabiliser*_V ‘inoxidize’, *inoxydabilisation*_N ‘inoxidization’, *inoxydabilisateur*_N ‘inoxidizer.M’

However, family (63) is not optimal because, as Dal & Namer (2010; 2012) have shown, speakers are reluctant to form verbs suffixed in *-iser* from adjectives in *-able*. This is confirmed by the very few attested verbs ending in *-abiliser* compared to the number of adjectives in *-able* in the French lexicon. An additional reason that penalizes here the suffixation in *-iser* is that the verb stem produced is too long, since it contains 6 syllables ([i.nɔ.ksi.da.bi.liz]) whereas the stem optimal size for French derivatives is 3 syllables (Roché 2003; Plénat 2009).

Another option is to create a formal stolon from the family that contains the adjective *oxydable*_A ‘oxidizable’ and the verb *oxyder*_V ‘oxidize’, prioritizing formal paradigmatic cohesion over semantic paradigmatic cohesion. We then obtain a formal family like (64) by prefixing the forms of the lexemes in family (54) with *in-*. The resulting paradigmatic family is connected to family (54) by a quasi-stolon similar to the one we proposed for the back-formation analysis of the verb *cryograver*_V ‘cryograph’.

- (64) *inoxydable*_A ‘stainless’, *inoxyder*_V ‘inoxidize’, *inoxydation*_N ‘inoxidization’, *inoxydateur*_N ‘inoxidizer.M’

This option creates the verb *inoxyder*_V ‘inoxidize’ through back-formation from the adjective *inoxydable*_A ‘stainless’. The global picture is illustrated in Figure 4.

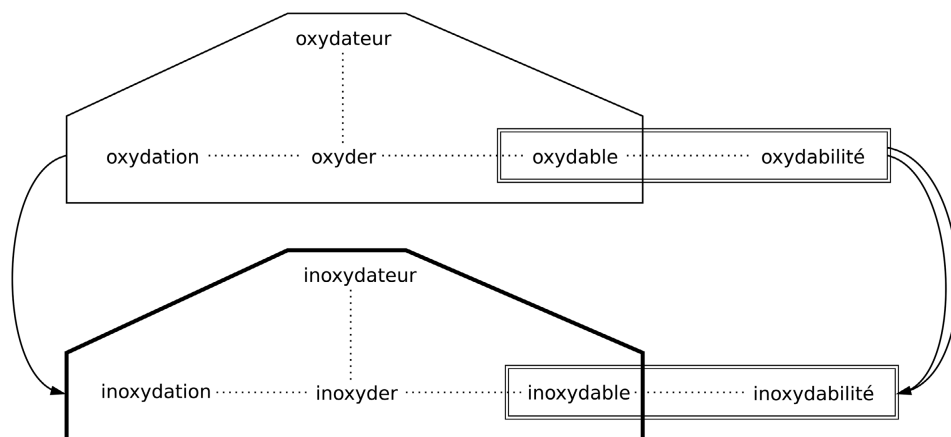


Figure 4: Schematic representation of the families involved in the analysis of back-formation for *inoxyder_V* ‘inoxidize’. Families framed in double lines are part of the same paradigm. They are linked by a stolon represented by a double arrow. Families framed in thick and thin lines belong to different paradigms. They are linked by a quasi-stolon represented by a single arrow.

7.3 Discussion

The analysis of *inoxyder_V* ‘inoxidize’ calls for three comments.

The first is that a Google search⁴ reveals that the infinitive form *inoxydabiliser_V* ‘inoxidize’ does not occur in its index. By contrast, Google indicates that 625 web pages contain the form *inoxyder_V* ‘inoxidize’. This result is consistent with the observation that speakers prioritize the formal cohesion of the morphological family, and that quasi-stolons are acceptable combinations as long as the formal and semantic families they associate are part of well-established formal and semantic paradigms. The use of quasi-stolons therefore does not seem to be limited to cases where a semantic paradigm is not regularly associated with any formal paradigm, but to be an additional mechanism for associating formal and semantic regularities.

The second point is that *inoxyder_V* ‘inoxidize’ does not mean ‘practice inoxidization’, which makes this back-formation different from that of *cryograver_V* ‘cryograph’. Family (64) corresponds to an action network, as does family (54). But the two families cannot align. The relation that exists between *oxyder_V* ‘oxidize’ and *oxydable_A* ‘oxidizable’ (65a) does not hold between *inoxyder_V* ‘inoxidize’ and *inoxydable_A* ‘stainless’ (66a). Conversely, the relation that exists between *inoxyder_V* ‘inoxidize’ and *inoxydable_A* ‘stainless’ (66b) does not hold between *oxyder_V* ‘oxidize’ and *oxydable_A* ‘oxidizable’ (65b).

- (65) a. *oxydable* = que l’on peut oxyder ‘which can be oxidized’
 b. *oxyder* ≠ rendre *oxydable* ‘make oxidizable’

⁴ The search was conducted on February 10, 2024.

- (66) a. *inoxydable* ≠ *que l'on peut inoxyder* 'which can be oxidized'
 b. *inoxyder* = *rendre inoxydable* 'make stainless'

The third comment is that the analysis of *inoxyder*_V 'oxidize' highlights the existence of a rivalry between a number of formal paradigms capable of realizing a semantic paradigm that contains the meaning speakers wish to express. In other words, rivalry classically conceived as a competition between several affixes that can be used to realize a given meaning is in fact a competition between paradigms capable of accommodating a family containing the words that are meant to realize that meaning.

8. General discussion

We have described various cases of verb back-formation in French and shown that they fall into different categories, depending on the paradigmatic configurations of the verbs analyzed. Before looking at what back-formation and paradigms tell us about each other (§8.2 and §8.3), we briefly address the comprehensiveness of the types of back-formed verbs we have presented.

8.1 Did we capture all the cases?

In order to answer this question, let us first review the back-formed verbs examined in this article. We described five types of back-formed verbs in French and proposed fully paradigmatic descriptions for each of them. These types constitute groups that are distinguished by (i) the complexity of their paradigmatic analyses, and (ii) their frequency.

- (i) The interaction between the formal and semantic paradigmatic dimensions manifests itself to varying degrees reflected in the increasingly complex paradigmatic structures involved in the formation of these verbs: regular meaning-form matching in the case of *lacter*_V 'produce milk'; formal stem reanalysis in favor of paradigmatic regularization in the case of *délater*_V 'denounce' and *allocter*_V 'allocate'; a more marked meaning-form mismatch in the case of *intuiter*_V 'guess', at the intersection of formal and semantic paradigms that are usually disjoint; a shift of center for verbs that are back-formed from neoclassical compounds, such as *cryograver*_V 'cryograph', where two morphological paradigms are in correspondence with the same formal paradigm; and a greater meaning-form discrepancy with *inoxyder*_V 'oxidize', which involves several morphological paradigms within the same word family.
- (ii) The back-formation processes that produce these types of verbs are, in our view, different in terms of productivity. Our intuition (which an exhaustive inventory of French back-formed verbs, as proposed in §3.2.3, could corroborate or refine) is that verbs with compound structure are the most frequent, due to the high productivity of neoclassical compounding processes forming activity nouns.

Moreover, we also believe that verbs such as *lacter*_V ‘produce milk’, *allocuter*_V ‘allocute’ and *délater*_V ‘denounce’ are frequent because they are reconstructed from nouns of Latin origin through a paradigmatically regular formation. Conversely, cases such as *intuiter*_V ‘guess’ and *inoxyder*_V ‘inoxidize’, neither of which are part of profitable series (Corbin 1987; Plag 1999; Bauer 2001), are scattered throughout the lexicon and therefore harder to catch.

We believe our study has identified most of the back-formation cases at the intersection of (i) and (ii), which probably form the majority of back-formed verbs in French. The remaining fringe consists of more atypical cases, such as *inoxyder*_V ‘inoxidize’ and *intuiter*_V ‘guess’, as well as perhaps other types of verbs resulting from processes involving paradigmatic shifts and novel paradigmatic recompositions. These types have yet to be discovered through an extensive morphological approach (Hathout et al. 2008; 2009; Hathout 2016), and will undoubtedly be the most interesting objects of analysis because of their rarity and ability to shed new light on observed phenomena.

8.2 What does back-formation tell us about paradigms?

The fundamental principle underlying this study is that all paradigms are already included in the lexicon, that all the paradigmatic families are created within a paradigm, and that all words belong to a paradigmatic family. As stated before, a paradigmatic family is defined as a subset of a word family where the relations are maximally interpredictable.

The main lesson to be learned from back-formation is that understanding how paradigms work requires an **unbundled analysis** of their formal and semantic dimensions. The paradigmatic model adopted for the analysis must allow for such a separation of levels of description. This first observation confirms our choice of the ParaDis framework, which offers a multi-dimensional organization, composed of formal, categorical and semantic paradigms that are independent of each other and in correspondence with a morphological level.

The second lesson is that back-formation is a phenomenon that must be described at the level of the paradigmatic family. Back-formation also shows that paradigmatic families can be **blended**, that is, they can result from an association between a formal family and a semantic family that belong to paradigms that are rarely associated.

Thirdly, back-formation informs us about how paradigmatic families are related to each other within a word family. The relation between two paradigmatic families is established by **stolons** when the families belong to the same paradigm, and by **quasi-stolons** when the two families are in correspondence with two formal families that belong to the same formal paradigm.

Back-formation also sheds light on the way paradigms are organized into networks, and in particular on the importance of their **center**, which represents the concept from which the other members of the network are defined.

Our analysis of back-formations helps us characterize three types of derivational processes within a paradigmatic framework. All three types are family-based. They are illustrated by the following examples, already presented in §7.

- (67) *oxyder*_V ‘oxidize’, *oxydation*_N ‘oxidization’, *oxydateur*_N ‘oxidizer.M’, *oxydable*_A ‘oxidizable’
- (68) *oxydable*_A ‘oxidizable’, *oxydabilité*_N ‘oxidizability’
- (69) *inoxydable*_A ‘stainless’, *inoxydabilité*_N ‘inoxidizability’
- (70) *inoxydable*_A ‘stainless’, *inoxyder*_V ‘inoxidize’, *inoxydation*_N ‘inoxidization’, *inoxydateur*_N ‘inoxidizer.M’

The first process is the creation of a family like (68) in a paradigm that corresponds to a metapattern (Fradin 2020) from the adjective *oxydable*_A ‘oxidizable’ in family (67). The relation between the members of this paradigm is essentially categorical and grammatical. The members of family (68) have the same meaning overall. The center of the family is the element of the category that realizes this meaning by default (Croft 1991). Specifically, the center is the adjective when the meaning is a property. The semantic properties of the paradigm in (68) place it somewhere in between derivation and inflection. Following Spencer (2013), who describes relations such as that between the adjective and the noun in (68) as transpositional, we will call the paradigm in (68) a **transpositional paradigm**.

The second process is the creation of a new family by stolon within the paradigm of a source family such as (69) obtained from (68). In this case, the two families have exactly the same structure. In particular, their centers are aligned.

The third process is the creation of a target family like (72) in a paradigm different from that of a source family like (71). In this case, as in the first process, the source and target families share a common lexeme. The difference with the second process is that here, the lexeme at the origin of the derivation is not at the center of the target family. For example, in a pollination situation, pollen plays the role of theme. As such, *pollen*_N ‘pollen’ is part of the family (72). The center of the family (72) is determined by the semantic properties of the target paradigm. In this case, it is the verb *polliniser*_V ‘pollenize’. The target family can be regular, like (72), or it can be a quasi-stolon, like (70) from (67).

- (71) *pollen*_N ‘pollen’, *pollinique*_A ‘of pollen’
- (72) *pollen*_N ‘pollen’, *polliniser*_V ‘pollenize’, *pollinisation*_N ‘pollenation’, *pollinisateur*_N ‘pollenizer.M’

8.3 What do paradigms tell us about back-formation?

The way in which French back-formed verbs fit into their paradigmatic families helps us better characterize this phenomenon. In particular, the analysis of verb back-formations in a paradigmatic framework reveals three types of back-formed verbs.

Type 1: The verb fills a cell in an existing paradigm (*lacter*_V ‘produce milk’, *allocuter*_V ‘allocute’). In this case, back-formation does not involve any meaning-form discrepancy.

It is simply a **formal regularization of an existing family**. A principle of completeness is probably at work, favoring those families with the least possible number of cells without formal realization.

Type 2: The verb (*intuiter*_V ‘guess’) extends a morphological family by combining structures of two paradigms that are not usually associated with each other. In this case, as in the previous one, back-formation is a **formal regularization**, resulting in a **blended family**.

Type 3: The verb belongs to a new family which is a formal copy of the source family but corresponds to a different semantic paradigm from that of the source family (*cryograver*_V ‘cryograph’, *inoxyder*_V ‘inoxidize’). The target family is then a **quasi-stolon of the source family**. The back-formation here involves a new family and new semantic relations: The event denoted by the verb *cryograver*_V ‘cryograph’ is an instance of the activity *cryogravure*_N ‘cryograph’ (whereas *graver*_V ‘engrave’ and *gravure*_N ‘engraving’ share the same semantic content); *inoxyder*_V ‘inoxidize’ means ‘to make *inoxydable*_A ‘stainless’’, whereas *oxyder*_V ‘oxidize’ does not mean ‘to make *oxydable*_A ‘oxidizable’’. Furthermore, we observed that speakers prioritize formal family resemblance.

9. Conclusion

Using examples of French back-formed verbs, we have shown in this article that it is possible to shed new light on the question of the nature of back-formation. The actualization of the relation between an affixed noun denoting a process and a related verb which is formally simpler and whose first attestation is clearly more recent than that of the noun may have two outcomes.

- (i) It may lead to the completion of an existing family. The repairs are transparent when the formal and semantic dimensions of the family paradigm are tightly associated (*prédater*_V ‘pray on’, *lacter*_V ‘produce milk’, *délater*_V ‘denounce’). They may be more marked when the paradigm combines a semantic dimension and a formal dimension that are not commonly associated (*intuiter*_V ‘guess’).
- (ii) It may lead to the creation of a new family that formally mirrors a source family (*cryograver*_V ‘cryograph’). The new family is produced by a mechanism we call “quasi-stolon”: The two families are in correspondence with the same formal paradigm, but with two different semantic paradigms.

The back-formation of *inoxyder*_V ‘inoxidize’ from *inoxydable*_A ‘stainless’ is much less common: The verb is the outcome of a process which forms a causative predicate of change of state. The quasi-stolon mechanism solves the problem posed by the sub-optimality of the forms produced by the more conventional means of derivation, such as conversion or the addition of the suffix *-iser*.

All the back-formation cases we have examined are motivated by a paradigmatic regularization that prioritizes the formal cohesion of the newly-formed verb family,

possibly at the cost of an atypical association between the formal and semantic paradigms. More generally, the conclusion of our study is that in a paradigm-based approach to morphology, back-formation is rarely a fully autonomous derivational process: It is mostly an operation governed by the morphological paradigms, whose aim is to fill a gap in a family with a lexeme that represents the best trade-off between faithfulness to the family's stems, insertion in a well-identified formal paradigm, and inclusion in the semantic network that best expresses the intended semantic content.

Our study focused on French. Derivational paradigms in other languages present different shapes, sizes and contents, notably due to variations in their inflectional morphology, in the prominence of compounding, in the agglutinative, fusional or other nature of their derivational morphology, and so on. We therefore need to check that the analyses we propose remain valid in these languages and to identify the conditions for their validity.

We have also suggested in this article that speakers favor formal regularities over morphosemantic regularities. Experimentally testing the reality of this strategy will be another direction in which we plan to pursue this study. The experiments we plan will be carried out by interviewing subjects, but also using computational models. We also plan to carry out a specific study of the notions of stolon and quasi-stolon introduced in this article. Which French affixes generate stolons? Are quasi-stolons used for formations other than back-formations? Are there quasi-stolons in which the shift concerns the formal families?

Last but not least, the biggest challenge is how derivational paradigms fit together. Figure 4 gives an idea. The study of many derivational phenomena within a paradigmatic framework will be greatly facilitated once the paradigm network is specified with sufficient precision.


Acknowledgements


We thank Livio Gaeta and Fabio Montermini for organizing the workshop "Back-formation in a new theoretical universe" at the IMM-20 (2022) conference and for editing this special issue. We also thank the workshop participants for their questions and comments. We thank the two anonymous reviewers and Richard Huyghe for their comments and suggestions, which helped us to improve the article. All remaining errors are ours.

Author Contribution Statement

The two authors contributed equally to collecting the data, developing the theoretical analysis, and writing the manuscript.

ORCID

Nabil Hathout  <https://orcid.org/0000-0003-4492-171X>

Fiammetta Namer  <https://orcid.org/0000-0002-6144-3011>

References

- Apothéloz, Denis. 2003. Le rôle de l'iconicité constructionnelle dans le fonctionnement du préfixe négatif *in-*. *Cahiers de Linguistiques Analogique* 1(1). 35–63.
- Arndt-Lappe, Sabine. 2015. Word-formation and analogy. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. 2, 822–841. Berlin: De Gruyter.
- Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2003. *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.
- Bauer, Laurie. 2008. Les composés exocentriques de l'anglais. In Amiot, Dany (ed.), *La composition dans une perspective typologique*, 35–47. Arras: Artois Presses Université.
- Becker, Thomas. 1993. Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1993*, 1–25. Dordrecht: Kluwer.
- Beecher, Henry. 2004. Derivational paradigm in word formation. Research Paper, UCSD.
- Bonami, Olivier & Strnadová, Jana. 2019. Paradigm structure and predictability in derivational morphology. *Morphology* 29. 167–197.
- Booij, Geert. 2000. Inflection and derivation. In Booij, Geert & Lehmann, Christian & Mugdan, Joachim & Skopeteas, Stavros (eds.), *Morphology: An international handbook on inflection and word-formation*, 360–369. Berlin: De Gruyter.
- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Booij, Geert & Masini, Francesca. 2015. The role of second order schemas in the construction of complex words. In Bauer, Laurie & Körtvélyessy, Lívía & Štekauer, Pavol (eds.), *Semantics of complex words*, 47–66. Cham: Springer.
- Booij, Geert E. 2005. Compounding and derivation: Evidence for construction morphology. In Dressler, Wolfgang U. & Kastovsky, Dieter & Rainer, Franz (eds.), *Morphology and its demarcations*, 109–132. Amsterdam: John Benjamins.
- Boyé, Gilles & Schalchli, Gauvain. 2019. Realistic data and paradigms: The Paradigm Cell Finding Problem. *Morphology* 29. 199–248.
- Bray, Laurent. 2000. *Matthias Kramer et la lexicographie du français en Allemagne au XVIIIe siècle : Avec une édition des textes métalexigraphiques de Kramer*. Tübingen: Max Niemeyer. <https://doi.org/10.1515/9783110960853>.
- Carstairs-McCarthy, Andrew. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70(4). 737–788.
- Corbett, Greville. 2010. Canonical derivational morphology. *Word Structure* 3(2). 141–155.
- Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Max Niemeyer.
- Corbin, Danielle. 2001. Du nouveau sur *beurre laitier*. Note sur une fausse conversion. In Buridant, Claude & Kleiber, Goerges & Pellat, Jean-Christophe (eds.), *Par monts et par vaux. Itinéraires linguistiques et grammaticaux. Mélanges de linguistique générale et française offerts au professeur Martin Riegel pour son soixantième anniversaire par ses collègues et amis*, 127–143. Leuven: Peeters.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organisation of information*. Chicago: University of Chicago Press.
- Cruse, D. Alan. 2004. *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford University Press.

- D'Achille, Paolo. 2005. Le retroformazioni in italiano. In Giovanardi, Claudio (ed.), *Lessico e formazioni delle parole. Studi offerti a Maurizio Dardano per il suo 70° compleanno*, 75–102. Firenze: Cesati.
- Dal, Georgette & Namer, Fiammetta. 2010. French property nouns based on toponyms or ethnic adjectives: A case of base variation. In Rainer, Franz & Dressler, Wolfgang U. & Kastovsky, Dieter & Luschützky, Hans Christian (eds.), *Variation and change in morphology: Selected papers from the 13th International Morphology Meeting, Vienna, February 2008*, 53–74. Amsterdam: John Benjamins.
- Dal, Georgette & Namer, Fiammetta. 2012. Faut-il brûler les dictionnaires ? Ou comment les ressources numériques ont révolutionné les recherches en morphologie. In Neveu, Franck & Muni Toke, Valelia & Blumenthal, Peter & Klingler, Thomas & Ligas, Pierluigi & Prévost, Sophie & Teston-Bonnard, Sandra (eds.), *Congrès mondial de linguistique française 2012*. 1261–1276. Paris: ILF.
- Dal, Georgette & Namer, Fiammetta. 2018. Playful nonce-formations in French, creativity and productivity. In Arndt-Lappe, Sabine & Braun, Angelika & Moulin, Claudine & Winter-Froemel, Esme (eds.), *Expanding the lexicon: Linguistic innovation, morphological productivity, and ludicity*, 209–228. Berlin: De Gruyter.
- Dressler, Wolfgang U. 2005. Word-formation in Natural Morphology. In Štekauer, Pavol & Lieber, Rochelle (eds.), *Handbook of word-formation*, 267–284. Dordrecht: Springer.
- Fernández-Domínguez, Jesús & Bagasheva, Alexandra & Lara-Clares, Cristina. 2020. What paradigms and what for? In Fernández-Domínguez, Jesús & Bagasheva, Alexandra & Lara-Clares, Cristina (eds.), *Paradigmatic relations in word formation*, 1–20. Leiden: Brill.
- Fradin, Bernard. 2018. Paradigms and the role of series in derivational morphology. *Lingue e Linguaggio* 17(2). 155–172.
- Fradin, Bernard. 2020. Characterizing derivational paradigms. In Fernández-Domínguez, Jesús & Bagasheva, Alexandra & Lara-Clares, Cristina (eds.), *Paradigmatic relations in word formation*, 49–84. Leiden: Brill.
- Gagné, Christina L. & Spalding, Patricia, Thomas L. Spicer & Wong, Dixie & Rubio, Beatriz & Perez Cruzns, Karen. 2020. Is buttercup a kind of cup? Hyponymy and semantic transparency in compound words. *Journal of Memory and Language* 113. 104–110.
- Gerdts, Donna. 1998. Incorporation. In Spencer, Andrew & Zwicky, Arnold M. (eds.), *The handbook of morphology*, 84–100. London: Blackwell Publishing.
- Haas, Pauline & Huyghe, Richard. 2010. Les propriétés aspectuelles des noms d'activités. *Cahiers Chronos* 21. 103–118.
- Haas, Pauline & Huyghe, Richard & Marín, Rafael. 2008. Du verbe au nom : Calques et décalages aspectuels. In Durand, Jacques & Habert, Benoît & Laks, Bernard (eds.), *Congrès mondial de linguistique française 2008*. 2051–2065. Paris: ILF.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Oxford University Press.
- Haspelmath, Martin. 2024. Inflection and derivation as traditional comparative concepts. *Linguistics* 62(1). 43–77.
- Hathout, Nabil. 2016. La question des données en morphologie. *Cahiers de l'ILSL* 45. 123–160.
- Hathout, Nabil & Montermini, Fabio & Tanguy, Ludovic. 2008. Extensive data for morphology: Using the World Wide Web. *Journal of French Language Studies* 18(1). 67–85.
- Hathout, Nabil & Namer, Fiammetta. 2018. La parasynthèse à travers les modèles : des RCL au ParaDis. In Bonami, Olivier & Boyé, Gilles & Dal, Georgette & Giraudo, Hélène & Namer, Fiammetta (eds.), *The lexeme in descriptive and theoretical morphology*, 365–399. Berlin: Language Science Press.

- Hathout, Nabil & Namer, Fiammetta. 2019. Paradigms in word formation: What are we up to? *Morphology* 29. 153–165.
- Hathout, Nabil & Namer, Fiammetta. 2022. ParaDis: A family and paradigm model. *Morphology* 32. 153–195.
- Hathout, Nabil & Namer, Fiammetta & Plénat, Marc & Tanguy, Ludovic. 2009. La collecte et l'utilisation des données en morphologie. In Fradin, Bernard & Kerleroux, Françoise & Plénat, Marc (eds.), *Aperçus de morphologie du français*, 267–287. Saint-Denis: Presses universitaires de Vincennes.
- Hathout, Nabil & Sajous, Franck. 2016. Wiktionnaire's wikicode GLAWified: A workable French machine-readable dictionary. In Calzolari, Nicoletta & Choukri, Khalid & Declerck, Thierry & Goggi, Sara & Grobelnik, Marko & Maegaard, Bente & Mariani, Joseph & Mazo, Helene & Moreno, Asuncion & Odijk, Jan & Piperidis, Stelios (eds.), *Proceedings of the tenth international conference on language resources and evaluation (LREC'16)*. 1369–1376. Portorož, Slovenia: European Language Resources Association (ELRA).
- Hathout, Nabil & Sajous, Franck & Calderone, Basilio & Namer, Fiammetta. 2020. Glawinette: A linguistically motivated derivational description of French acquired from GLAWI. In Calzolari, Nicoletta & Béchet, Frédéric & Blache, Philippe & Choukri, Khalid & Cieri, Christopher & Declerck, Thierry & Goggi, Sara & Isahara, Hitoshi & Maegaard, Bente & Mariani, Joseph & Mazo, Hélène & Moreno, Asuncion & Odijk, Jan & Piperidis, Stelios (eds.), *Proceedings of the twelfth language resources and evaluation conference (LREC'20)*. 3877–3885. Marseille, France: European Language Resources Association (ELRA).
- Hathout, Nabil & Sajous, Franck & Calderone, Basilio & Namer, Fiammetta. 2025. GLeMM, a derivational resource that gives you more. Paper presented at the ISMo 2025 conference, University of Lille.
- Hirschbühler, Paul & Labelle, Marie. 2009. French locatum verbs and incorporation. *Lingua* 119(2). 263–279.
- Horecký, Ján & Bosák, Ján & Buzássyová, Klára. 1989. *Dynamika slovnej zásoby súčasnej slovenčiny*. Bratislava: Veda.
- Kastovsky, Dieter. 1968. *Old English deverbal substantives derived by means of a zero morpheme*. Esslingen: Langer.
- Kastovsky, Dieter. 2009. Diachronic perspectives. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of compounding*, 323–342. Oxford: Oxford University Press.
- Knittel, Marie Laurence. 2010. Possession vs. pseudo-incorporation in the nominal domain: Evidence from French event nominals dependencies. *The Linguistic Review* 27(2). 177–230.
- Koehl, Aurore. 2012. *La construction morphologique des noms désadjectivaux suffixés en français*. Nancy: Université de Lorraine. PhD dissertation.
- Lignon, Stéphanie & Namer, Fiammetta & Villoing, Florence. 2014. De l'agglutination à la triangulation ou comment expliquer certaines séries morphologiques. In Neveu, Franck & Blumenthal, Peter & Hriba, Linda & Gerstenberg, Annette & Meinschaefer, Judith & Prévost, Sophie (eds.), *Congrès mondial de linguistique française 2014*. 1813–1836. Paris: ILF.
- Manova, Stela. 2018. Subtractive morphology. In Aronoff, Mark (ed.), *Oxford bibliographies in linguistics*, Oxford: Oxford University Press. <https://doi.org/10.1093/obo/9780199772810-0116>.
- Marchand, Hans. 1969. *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. Munich: Beck.
- Martin, Robert. 1983. *Pour une logique du sens*. Paris: Presses Universitaires de France.

- Massam, Diane. 2009. Noun incorporation: Essentials and extensions. *Language and Linguistics Compass* 3(4). 1076–1096.
- Mathieu, Eric. 2009. Introduction to “noun incorporation and its kind”. *Lingua* 119(2). 141–147.
- Nagano, Akiko. 2007. Marchand’s analysis of back-formation revisited: Back-formation as a type of conversion. *Acta Linguistica Hungarica* 54(1). 33–72.
- Namer, Fiammetta. 2012. Nominalisation et composition en français : D’où viennent les verbes composés ? *Lexique* 20. 173–205.
- Namer, Fiammetta & Hathout, Nabil. 2020. ParaDis and Démonette – From theory to resources for derivational paradigms. *The Prague Bulletin of Mathematical Linguistics* 114. 5–33.
- Pennanen, Esko V. 1966. *Contributions to the study of back-formation in English*. Tampere: Yhteiskunnallinen Korkeakoulu.
- Pennanen, Esko V. 1975. What happens in back-formation? In Hovdhaugen, Elisabeth (ed.), *Papers from the second scandinavian conference of linguistics*. 216–229. Department of Linguistics Oslo: University of Oslo.
- Plag, Ingo. 1999. *Morphological Productivity. Structural Constraints in English Derivation*. Berlin: De Gruyter.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- Plénat, Marc. 2009. Les contraintes de taille. In Fradin, Bernard & Kerleroux, Françoise & Plénat, Marc (eds.), *Aperçus de morphologie du français*, 47–63. Saint-Denis: Presses Universitaires de Vincennes.
- Rainer, Franz. 2004. Retroformazione. In Grossmann, Maria & Rainer, Franz (eds.), *La formazione delle parole in italiano*, 493–497. Tübingen: Max Niemeyer.
- Roché, Michel. 2003. De la « bonne formation » des dérivés en *-on*. *Cahiers de Grammaire* 28. 91–112.
- Roché, Michel. 2006. Comment les adjectifs sont sémantiquement construits. *Cahiers de Grammaire* 30. 373–387.
- Roché, Michel. 2023. Les familles dérivationnelles : comment ça marche ? *Lexique* 33. 166–209.
- Sajow, Frank & Hathout, Nabil. 2015. Glawi, a free XML-encoded machine-readable dictionary built from the French Wiktionary. In Kosem, Iztok & Jakubiček, Miloš & Kallas, Jelena & Krek, Simon (eds.), *Electronic lexicography in the 21st century: Linking lexical data in the digital age. proceedings of the eLex 2015 conference*. 405–426. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd.
- Sanacore, Daniele. 2023. *Une histoire de famille : Description morphosémantique des lexèmes construits et des relations dérivationnelles*. Toulouse : Université de Toulouse Jean Jaurès. PhD dissertation.
- Sanacore, Daniele & Hathout, Nabil & Namer, Fiammetta. 2021. Frame-like structures for morphosemantic description. *Verbum* 43(1). 179–194.
- Sanacore, Daniele & Hathout, Nabil & Namer, Fiammetta. 2024. A scenario-based approach to predictability in derivation. *Lingue e Linguaggio* 2024(2). 337–359.
- Scalise, Sergio & Bisetto, Antonietta. 2009. The classification of compounds. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of compounding*, 34–53. Oxford: Oxford University Press.
- Schäfer, Roland & Bildhauer, Felix. 2012. Building large corpora from the web using a new efficient tool chain. In Calzolari, Nicoletta & Choukri, Khalid & Declerck, Thierry & Doğan, Mehmet Uğur & Maegaard, Bente & Mariani, Joseph & Moreno, Asuncion & Odijk, Jan &

- Piperidis, Stelios (eds.), *Proceedings of the eighth international conference on language resources and evaluation (LREC'12)*. 486–493. Istanbul, Turkey: European Language Resources Association (ELRA).
- Spencer, Andrew. 2013. *Lexical relatedness*. Oxford: Oxford University Press.
- Spencer, Andrew. 2021. Paradigms, features, and the lexeme. In Boyé, Gilles & Schalchli, Gauvain (eds.), *ParadigMo2*. Bordeaux: Université Montaigne.
- Štekauer, Pavol. 2014. Derivational paradigms. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of derivational morphology*, 354–369. Oxford: Oxford University Press.
- Štekauer, Pavol. 2015. Backformation. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Štekauer, Pavol (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. 1, 340–352. Berlin: De Gruyter.
- Strnadová, Jana. 2014. *Les réseaux adjectivaux : Sur la grammaire des adjectifs dénominaux en français*. Université Paris Diderot & Univerzita Karlova V Praze. PhD dissertation.
- Thornton, Anna Maria. 2007. Phénomènes de réduction en italien. In Delais-Roussarie, Elisabeth & Labrune, Laurence (eds.), *Des sons et des sens : Données et modèles en phonologie et en morphologie*, 241–268. Paris: Hermès Science/Lavoisier.
- Ylonen, Tatu. 2022. Wiktextextract: Wiktionary as machine-readable structured data. In Calzolari, Nicoletta & Béchet, Frédéric & Blache, Philippe & Choukri, Khalid & Cieri, Christopher & Declerck, Thierry & Goggi, Sara & Isahara, Hitoshi & Maegaard, Bente & Mariani, Joseph & Mazo, Hélène & Odijk, Jan & Piperidis, Stelios (eds.), *Proceedings of the thirteenth language resources and evaluation conference*. 1317–1325. Marseille, France: European Language Resources Association (ELRA).