ParaDis : a Family and Paradigm model

Nabil Hathout
CLLE, CNRS &
Université Toulouse Jean Jaurès

Fiammetta Namer
ATILF, Université de Lorraine &
CNRS

Abstract

The unification of inflectional and derivational morphology is an issue that is often debated but on which there is no consensus. On the other hand, it is well known that inflectional morphology is organized into paradigms. This paper contributes to the convergence between inflection and derivation by offering a new paradigmatic model of derivational morphology. This model, called ParaDis, is based on two types of objects, families and paradigms, and on an architecture made up of three levels of representation (formal, categorical and semantic) and a morphological level, which connects the other three. Morphological descriptions are distributed at three level, where they organize into formal, categorical and semantic families and paradigms. In this way, the analysis of a phenomenon adjusts closely to its level-specific regularities. The paper describes ParaDis and shows how it can be used to analyze a wide range of canonical and non-canonical derivational constructions in a simple and intuitive way. It compares ParaDis to five frameworks (Bochner, 1993; Koenig, 1999; Booij, 2010; Bonami & Strnadová, 2019; Jackendoff & Audring, 2020b) and illustrates its contributions through the analysis of parasynthetic derivatives.

1 Introduction

In recent years, there has been a growing interest in the paradigmatic approaches to derivational morphology with an increase in the number of articles, books, workshops and thematic issues dealing with this question. This activity has resulted in a number of specific proposals that essentially address three questions: (1) Is derivational morphology paradigmatic? (2) What do derivational paradigms look like? (3) What do they bring to derivational morphology?

Unlike inflectional morphology, whose paradigmatic nature is well known (Blevins, 2006, 2016), derivational morphology has, for a long time, been considered to be of a different nature because of its irregularities, induced for example by the rivalry between suffixes like *-ity* and *-ness* in English or *-age*, *-ion* and *-ment* in French. However, this view evolved in the last decade under the influence of works that argue that regularities and irregularities in inflectional and derivational morphology are basically of the same nature and that they only differ in degrees (Bauer, 1997; Štekauer, 2014; Spencer, 2013; Boyé & Schalchli, 2019; Bonami & Strnadová, 2019). This evolution toward the unification of morphology extends the notions, principles and theoretical frameworks of inflectional morphology to derivational morphology, and especially its paradigmatic organization. This organization is the subject of many works that discuss the structure and content of derivational paradigms. Yet, we do not know of any large-scale systematic descriptions of the derivational paradigms of any language similar to the verb conjugation tables of Romance languages. Most proposals extend the

inflectional paradigms to derivation in a more or less direct way (Stump & Finkel, 2013). Others, like Štekauer (2014) and Körtvélyessy et al. (2020), consider that derivational paradigms are onomasiological in nature and that they are structured by relations that hold between concepts. The derivational paradigms proposed in these different works are primarily intended for the analysis of non-canonical phenomena (Corbett, 2010) that are difficult to account for in "traditional" rule-based theoretical frameworks (Blevins, 2006). They are also adopted because they describe derivational morphology by means of networks of implications between words that reflect recent psycholinguistic advances in language acquisition and usage (Blevins et al., 2016).

In this paper, we present ParaDis (Paradigms vs Discrepancies), a paradigmatic model of derivational morphology. ParaDis is in line with paradigm-based approaches to morphology like the ones proposed by Bochner (1993) and Bonami & Strnadová (2019). It also has similarities with Jackendoff's (2002; 2007; 2009) Parallel Architecture, as it is based on interconnected autonomous levels of description.

In a nutshell, (*i*) ParaDis takes up the classical description of lexemes along three independent dimensions (i.e. levels of representation); (*ii*) it transposes the paradigmatic description of lexemes to morphological families; (*iii*) it generalizes the paradigmatic organization to all the levels of representation that are relevant to morphology: formal (phonological), categorical, semantic and morphological (lexical); (*iv*) at all levels of representation, concrete paradigms are alignments of families whose members are in the same relations of contrast; the aligned members form series; abstract paradigms generalize these alignments into networks of relations between series characterized by patterns; (*v*) abstract paradigms in the different levels of representation are not necessarily isomorphic; (*vi*) morphological paradigms can be superposed in order to account for linguistically relevant generalizations; (*vii*) ParaDis includes a set of constraints that control the production of derived lexemes and guide their analysis; however, for the sake of clarity, these constraints are not presented in detail in this article. These features give ParaDis a descriptive power that allows for a simple and intuitive analysis of a range of canonical and non-canonical phenomena. They also determine a method of analysis of morphological phenomena where data is first described by means of homogeneous paradigms which are then superposed to reconstruct heterogeneous structures that account for the phenomena in all their complexity. Moreover, the mechanisms required to deal with canonical and non-canonical derivation also operate in inflection but we will not develop their implementation for inflection here.

The remainder of the article is organized as follows. In Section 2, we present the key features of inflectional paradigms and discuss their transposition to derivational morphology. Section 3 reviews five existing models that can be used for the description of derivational paradigms. We then present the principles and mechanisms of ParaDis in Section 4. Section 5 is devoted to form-meaning discrepancies and to the description of some non-canonical phenomena. Section 6 compares the analysis of parasynthetic constructions in ParaDis and in four of the models presented in Section 3. Finally, Section 7 summarizes the main contributions of ParaDis.

2   Paradigmatic morphology

The purpose of this paper is to provide an answer to the recurring question: "Can we adapt the paradigmatic organization of inflectional morphology to derivational morphology?" This question has been addressed in many studies, from Van Marle (1985) to Bonami & Strnadová (2019) including Bauer (1997); Stump (1991, 2005); Štekauer (2014); Antoniova & Štekauer (2015); Boyé & Schalchli (2016, 2019) among many others. All these authors agree on the fact that: (*i*) inflection and

derivation are two parts of a single morphology, (*ii*) inflectional morphology is paradigmatic; (*iii*) inflectional paradigmatic organization can be transposed to derivational morphology. We adopt the same assumptions here.

The unity of morphology has been, and remains, subject to debate. Some theories, like Split Morphology (Matthews, 1972; Anderson, 1982), consider inflection and derivation as independent domains but this separation is rejected in more recent studies which defend the idea that they are part of a continuum (see (Walther, 2013, chapter 2) for a detailed presentation). This view owes much to work in typology like (Haspelmath, 1996); Štekauer (2015) presents the main arguments in support of this position; Spencer (2013) describes in detail a number of phenomena that illustrate the porosity of the boundary between the two sub-domains; Bonami & Strnadová (2019) and Boyé & Schalchli (2019) show how inflection and derivation can be described with the same model in a perfectly natural way. We come back to this question in Section 3.

2.1 Inflectional paradigms

There is an increasing acceptance of the fact that inflection is paradigmatic. This view is becoming a standard (Carstairs-McCarthy, 1994; Wunderlich & Fabri, 1995; Stump, 2001; Ackerman et al., 2009; Baerman et al., 2010). Its adoption was made possible by the growing importance of the abstractive word-based models in morphology (Blevins, 2016) and by the shift from realizational systems, where sets of rules yield inflected word forms from base units, to a system where inflected forms are conceived as realizations of words. The paradigmatic organization of inflection arises from the possibility to group inflected word forms into abstract representations, namely LEXEMES (Anderson, 1992; Aronoff, 1994). The set of inflected word forms gathered in a lexeme is its concrete paradigm, also known as its Paradigm$_2$ (Carstairs-McCarthy, 1994). For instance, the concrete paradigm of the French lexeme *laver* 'to wash' includes the inflected word forms in (1). The morphosyntactic properties they realize is given in the bottom line.

(1)   *lave*   *lavons*   *laverai*   *laverons*   *lavèrent*   *lavant* ...
      PRES.1SG   PRES.1PL   FUT.1SG   FUT.1PL   PST.3PL   PRESP

Grouping forms into lexemes reveals an important regularity: in a given language, the inflectional paradigms of the lexemes of a given grammatical category all have the same size. This is for example the case of French verbs, which have 51 inflected forms. The inflectional paradigmatic organization is also based on the fact that the inflected word forms of a lexeme are identified by their MORPHOSYNTACTIC FEATURES. These features and their combinations are determined by the language and the grammatical category of the lexeme. For instance, French morphosyntactic features include gender, number, person, tense, mode, etc. The lexemes whose word forms are in the same formal and morphosyntactic relations belong to the same INFLECTIONAL CLASS. They instantiate the same abstract paradigm or Paradigm$_1$ (Carstairs-McCarthy, 1994).

Lexemes highlight a third regularity: the interpredictability of their inflected forms (Wurzel, 1989; Boyé, 2011), in particular of their morphosyntactic features and their formal (i.e. phonological) representations. Interpredictability can be described in terms of implications as in (Wurzel, 1989) or implicative entropy as in (Bonami, 2014; Bonami & Beniamine, 2016; Beniamine, 2018). For instance, if we know that a French verb has an inflected form *laverai* whose morphosyntactic features are FUT.1SG, we can predict with a good level of confidence that it also has an inflected form *laverons* whose morphosyntactic features are FUT.1PL, and vice versa. The strength of the prediction depends on the grammatical category of the lexeme, its inflectional class and the paradigm cells involved.

The paradigmatic nature of inflectional morphology is well-known in Romance languages. In French dictionaries and textbooks, it often takes the form of tables like Table 1 which presents an excerpt from the paradigms of four verbs that have the same conjugation: *laver*, *casser* 'to break', *éclairer* 'to light' and *saluer* 'to greet'. This paradigm is indexed by the combination of morphosyntactic features in the table header. In any two columns of the paradigm the word forms display the same formal contrasts. For example, the formal difference between *laverons* and *laveraient* is the same as the one between *casserons* and *casseraient*; it could be described as a substitution of the *-eraient* ending for the *-erons* ending.

	PRES.1SG	PRES.2SG	FUT.1PL	FUT.2PL	COND.3SG	COND.3PL
laver	*lave*	*laves*	*laverons*	*laverez*	*laverait*	*laveraient*
casser	*casse*	*casses*	*casserons*	*casserez*	*casserait*	*casseraient*
éclairer	*éclaire*	*éclaires*	*éclairerons*	*éclairerez*	*éclairerait*	*éclaireraient*
saluer	*salue*	*salues*	*saluerons*	*saluerez*	*saluerait*	*salueraient*

Table 1: Excerpt from the inflectional paradigm of French verbs which conjugate like *laver*. Each line contains the word forms of one verb. The columns contain forms that realize the same morphosyntactic features. The features are given in the table header.

2.2 What do derivational paradigms bring to morphology?

The benefits brought to morphology by derivational paradigms are numerous and well known. We will essentially recall the most important of them, identified in particular by Bauer (1997). First of all, paradigms allow for the unification of inflectional and derivational morphology and thereby account for the fact that they use the same processes (prefixation, suffixation, reduplication, etc.) and the same stems (or themes). A second important contribution is that derivational paradigms account for the fact that lexemes form derivational families, and may be morphologically related without having a common ancestor, like in the English family (*aggressor*, *aggressive*, *aggression*) which does not normally contain a verb *to aggress* or in the French family (*prédateur* 'male predator', *prédatrice* 'female predator', *prédation* 'predation') which lacks a verb *préder* 'to predate'.

Paradigms also allow for the description of cases of systematic syncretism that word formation rule (WFR) frameworks cannot easily capture, such as the fact that, in French, the relational adjective of a demonym (e.g., *italien* 'inhabitant of Italy') has the same form as the corresponding language name (e.g., *italien* 'the language of the Italians'). They also account for the fact that derivationally related lexemes may share the same stem, e.g., French /fɔʁmat/, found in *formation* 'training' and *formateur* 'trainer', and /fɔʁm/, found in *déformer* 'deform' and *formable* 'formable'.

Another phenomenon that derivational paradigms are better able to describe than WFRs is form-meaning discrepancies that can be observed, for example, in French relational adjectives in *-istique* (e.g., *journalistique* 'journalistic'), which are regularly derived from a base noun (*journaliste* 'journalist') but may have multiple interpretations (Strnadová, 2014; Roché, 2011; Namer, 2021) because their meaning can also be formed on that of other nouns in their family (*journalisme* 'journalism' and *journal* 'newspaper'). More generally, complex lexemes may display a wide variety of form-meaning discrepancies which present regularities that derivational paradigms can grasp because they can describe relations between any pair of lexemes in a derivational family. By contrast, WFRs can only handle relations between derivatives and their bases.

2.3 Paradigmatic families

One question we address in this article is whether it is possible to design derivational paradigms that have the same structure and properties as inflectional paradigms. The possibility to group derivatives formed by regular WFRs in a table like Table 2 clearly suggests that the answer is affirmative. This is confirmed by the numerous examples of derivational paradigms presented in studies like (Körtvélyessy et al., 2020). Table 2 actually contains "slices" of four derivational families made up of lexemes which display the same formal, categorical and semantic contrasts, and therefore form a (partial) paradigm. Notice the lack of header in Table 2, because we do not know what properties may serve as indexes of the columns in derivational paradigms.

laver	*lavage*	*laveur*	*laveuse*	*lavable*
'to wash'	'washing'	'male washer'	'female washer'	'washable'
casser	*cassage*	*casseur*	*casseuse*	*cassable*
'to break'	'breaking'	'male breaker'	'female breaker'	'breakable'
éclairer	*éclairage*	*éclaireur*	*éclaireuse*	*éclairable*
'to light'	'lighting'	'male lighter'	'female lighter'	'lightable'
souder	*soudage*	*soudeur*	*soudeuse*	*soudable*
'to solder'	'soldering'	'male solderer'	'female solderer'	'solderable'

Table 2: Example of a French derivational paradigm represented in a tabular format.

Beyond these examples, the previous affirmative answer is based on a set of correspondences between inflectional and derivational objects and regularities. Basically, inflectional morphology is concerned with the relations between the inflected forms of a lexeme, whereas derivational morphology deals with the regular relations of form and meaning between morphologically related lexemes, that is, between lexemes that belong to the same MORPHOLOGICAL FAMILY. The notion of morphological family, aka word family, is well-known in morphology (Haspelmath & Sims, 2010). It can be defined as a set of lexemes connected by derivational relations. A part of the morphological family of the verb *laver* is presented in (2). In this family, *lavage*, *laveur*, *laveuse* and *lavable* derive from *laver*, and *lavabilité* and *inlavable* from *lavable*.

(2) *laver*, *lavage* 'washing', *laveur* 'male washer', *laveuse* 'female washer', *lavable* 'washable', *lavabilité* 'washability', *inlavable* 'unwashable', ...

The analogy between lexemes in inflection and morphological families in derivation is however not entirely accurate because the size of the morphological families is highly variable since derivation is not obligatory. As a result, the interpredictability between inflected forms cannot be fully transposed to morphological families. For example, if we know that the family of *herbe* 'weed' contains the noun *désherbage* 'weeding', we cannot predict the existence of derivatives such as *herbacée* 'herbaceous' or *herboristerie* 'herbalist shop' and vice versa. However, the correspondence could be restored if we consider that, just as the paradigm of a lexeme is defined by its inflectional class, derivational paradigms can be defined by an equivalent of the latter, namely a DERIVATIONAL CLASS, that is, a set of morphological relations that form a dense network (Štekauer, 2014; Fradin, 2020). In the following, we call PARADIGMATIC FAMILY a subset of a morphological family identified by the derivational class of a paradigm. One consequence of this definition is that the paradigmatic families that belong to the same derivational class have the same size and that the lexemes they contain are connected by the same relations. Moreover, these lexemes are strongly related and

5

highly interpredictable. For example, if we know that a family contains an action noun ending in *-age* such as *lavage*, we can predict that it also contains a dynamic verb *laver*, a masculine agent noun *laveur*, a feminine action noun *laveuse* and a possibility adjective *lavable*. Paradigmatic families could also be seen as sets of lexemes denoting entities which participate in situations similar to the ones considered by Fillmore (1976) in frame semantics as proposed by Sanacore et al. (2021).

2.4 Content-based derivational paradigms

Paradigms describe regularities and generalizations that can be semantic, formal or both. However, in order to achieve a better convergence with inflectional morphology, many authors like Štekauer (2014) argue that derivational paradigms are determined by the semantic content of the lexemes they contain and that the possible formal variations they display are secondary. This view, also adopted by Antoniova & Štekauer (2015) and Bonami & Strnadová (2019), is illustrated in Table 3. In this example, the lexemes in each family are interpredictable because the existence of an agent (Agent_N) implies that of a related dynamic Verb and a related action noun (Action_N). These families all instantiate the same abstract situation called *action network* by Roché (2017) and Fradin (2021). On the other hand, they display several formal variations. Their action nouns are suffixed in *-age* (*lavage*), *-ation* (*formation*) and *-ment* (*lancement*) or formed by conversion (*danse*), and their agent nouns are suffixed in *-eur* (*laveur*) and *-ateur* (*formateur*).

Verb	Action_N	Agent_N
laver	*lavage*	*laveur*
'to wash'	'whashing'	'washer'
saler	*salage*	*saleur*
'to salt'	'salting'	'salter'
former	*formation*	*formateur*
'to train'	'training'	'trainer'
fonder	*fondation*	*fondateur*
'to found'	'foundation'	'founder'
lancer	*lancement*	*lanceur*
'to lauch'	'launching'	'launcher'
ronfler	*ronflement*	*ronfleur*
'to snore'	'snoring'	'snorer'
danser	*danse*	*danseur*
'to dance'	'dance'	'dancer'
voler	*vol*	*voleur*
'to steal'	'theft'	'thief'

Table 3: Another example of French derivational paradigm, adapted from Bonami & Strnadová (2019).

3 Existing frameworks and models

ParaDis has two distinctive features that set it apart from other paradigmatic morphological models: (*i*) paradigms are generalized to all levels of representation; (*ii*) the levels of representation are

autonomous. These features place ParaDis at the crossroads of two grammatical traditions: it is both a paradigmatic model of derivational morphology and a model that belongs to the family of approaches based on independent and interconnected levels of representation like Parallel Architecture (Jackendoff, 2002, 2007, 2009). Although Parallel Architecture (PA) is usually presented from a syntactic point of view, it is very close to ParaDis in that semantics, syntax, and phonology are treated in parallel as autonomous generative devices. PA is presented in Section 3.1. The following sections present five models where derivational morphology can be described paradigmatically and with which ParaDis shares some of its characteristics. All five can handle inflection and derivation in a uniform manner but nevertheless differ on the way families and paradigms are represented and on the nature of the relations that exist between family members.

3.1   Parallel Architecture

PA is a theoretical framework developed over the last two decades by Jackendoff (2002, 2007, 2009). It covers all areas of grammar. PA is a successor of Sadock's (1991) Autolexical Syntax, whose central idea is that the components of the grammar are autonomous specifications which state conditions of sentence well-formedness. It is based on the association between three independent modules: semantic, syntactic and phonological. Semantics, syntax and phonology generate parallel representations which do not derive from one another as each module is hierarchically equivalent to all the others. The connection between the modules is realized by lexical items, i.e. small-scale interface rules. A lexical item (which may correspond to any unit or construct: sentence, lexeme, etc.) is well-formed if it satisfies all the constraints of all the modules.

Lexical items can be illustrated by a ternary diagram as in (3), borrowed from Jackendoff & Audring (2020b, 11), which shows how indexes are used to connect the three modules. More precisely, the role of indexes is to identify lexical items by means of three sets of properties: semantic, syntactic and phonological. For example, the constant index 4 in (3) identifies a verb by the phonological form /dəvawr/ and the concept DEVOUR. Likewise, the lexical item identified by index 5 is syntactically a transitive VP whose head is indexed by 4; its semantic-conceptual structure involves an agent X and a patient Y; phonologically, it is realized as a sequence obtained by concatenating an unspecified segment $/\cdots/$ to the right of /dəvawr/. The segment $/\cdots/$ is the phonological property of an abstract lexical item identified by a variable index y, which semantically is the patient Y and syntactically is a noun phrase (NP).

(3)   Phonology: $/\text{dəvawr}_4 \cdots_y/_5$
      Syntax: $[_{VP}\ V_4\ NP_y\ ]_5$
      Semantics: $[\text{DEVOUR}_4(\text{AGENT:X, PATIENT:Y}_y)]_5$

PA is close to Construction Grammars (Fillmore, 1968) with the difference that form-meaning correlation is not required: the description of a lexical item may not be specified at the three levels. In the case of an idiomatic construction like *chew the fat* 'converse idly' in (4), borrowed from Jackendoff & Audring (2020b, 95), the lexical items identified by 17, 18 and 19 have no semantic realization. Only lexical item 16 has one. The possibility to dispense with full form-meaning correspondence gives PA great flexibility. PA has recently been enriched with a morphological component, Relational Morphology (Jackendoff & Audring, 2018, 2020b,a), presented in Section 3.5.

(4)   Phonology: $/\text{tʃu}_{17}\ \text{ðə}_{18}\ \text{fæt}_{19}/_{16}$
      Syntax: $[_{VP}\ V_{17}\ [_{NP}\ Det_{18}\ N_{19}]_{20}\ ]_{16}$
      Semantics: $[\text{CONVERSE(X)}]_{16}$

3.2 Cumulative patterns

Bochner (1993) is one of the first to propose a paradigmatic model of derivation, in line with Jackendoff's (1975) *Lexical Relatedness Morphology*. It is based on the assumption that derivational rules are not directed and on the notion of CUMULATIVE SET (CS). A CS is defined as a set of correlated units within a morphological family such as (5) borrowed from (Bochner, 1993, 70), or (6). In Bochner's model, like in those of many others (Stump, 1991; Koenig, 1999; Spencer, 2013), lexemes and inflected forms are not treated separately. For example, a CS may contain both a plural form like *causes* and a derivative like *causal*. CSs are said to be "cumulative" because any subsets of (5) or (6) is a CS.

(5) {*cause, causal, causality*}

(6) a. {*laver, lavage, laveur*}
     b. {*saler, salage, saleur*}

In a CS, words are connected by undirected relations. For instance, the relation between *laver* and *lavage* in (6a) and between *saler* and *salage* in (6b) are instances of schema (7a). Similarly, the relation between *laver* and *laveur* and between *saler* and *saleur* are described by schema (7b). In addition, in Bochner's (1993) model, *lavage* and *laveur* are connected by a redundant morphological relation described by schema (7c).

(7) a. $\begin{bmatrix} /X/ \\ V \\ \text{`Z'} \end{bmatrix} \leftrightarrow \begin{bmatrix} /X \text{ -age}/ \\ N \\ \text{`act of Z'} \end{bmatrix}$

     b. $\begin{bmatrix} /X/ \\ V \\ \text{`Z'} \end{bmatrix} \leftrightarrow \begin{bmatrix} /X \text{ -eur}/ \\ N \\ \text{`he who Z'} \end{bmatrix}$

     c. $\begin{bmatrix} /X \text{ -age}/ \\ N \\ \text{`act of Z'} \end{bmatrix} \leftrightarrow \begin{bmatrix} /X \text{ -eur}/ \\ N \\ \text{`he who Z'} \end{bmatrix}$

(8) $\left\{ \begin{bmatrix} /X/ \\ V \\ \text{`Z'} \end{bmatrix}, \begin{bmatrix} /X \text{ -age}/ \\ N \\ \text{`act of Z'} \end{bmatrix}, \begin{bmatrix} /X \text{ -eur}/ \\ N \\ \text{`he who Z'} \end{bmatrix} \right\}$

With these three relations, each of the two CSs in (6) becomes a complete undirected graph which could be described as an instance of an abstract set of patterns like (8). This set of patterns is called a CUMULATIVE PATTERN (CP). The patterns in a CP are instantiated by the words of the corresponding CSs (e.g., (8) is instantiated by the two CSs in (6)). A CP can be seen as an abstract paradigm and its patterns as the cells of this paradigm. Moreover, just like CSs, CPs are cumulative as any sub-set of a CP is itself a CP. CPs account for:

1. the correlations that exist between the formal, categorical and semantic properties of the lexemes, e.g., the fact that nouns denoting actions may end in *-age*, as in the second pattern in (8);

2. the fact that formal, categorical and semantic properties may be shared by several lexemes within a CS. Sharing is described by means of variables like X and Z;

3. the fact that formal, categorical and semantic properties may be shared by several lexemes that hold the same position in the instances of a CP. For example, all lexemes that match the second pattern in (8) are action nouns suffixed in *-age* and all the ones that appear in the third position are agent nouns suffixed in *-eur*.

The above examples show that Bochner's goal of "simplicity" is fully achieved with a framework where paradigms and derivational families are explicitly represented.

3.3 Paradigmatic systems

In line with Bochner (1993), Bonami & Strnadová (2019) propose a model in which inflectional and derivational paradigms are represented in exactly the same way. The model includes a set of concepts and tools that extend the inflectional paradigms to derivation such as the family defined as a complete graph (9). An important contribution of this work is the explicit description of how families align within paradigms (10).

(9) A MORPHOLOGICAL FAMILY is a tuple $F = (w_1, ..., w_n)$ such that any member w_i of the family is morphologically related to any other member of w_j. A morphological family F is COMPLETE if there exists no larger morphological family that contains all members of F. A morphological family is PARTIAL if it is not complete.

(10) Given two ordered pairs of morphologically related words (w_1, w_2) and (w_3, w_4), we say that the two pairs are ALIGNED if the same content relation holds between them: there is some content relation R_c such that $R_c(w_1, w_2)$ and $R_c(w_3, w_4)$. We call R_c the ALIGNING RELATION.

With this second definition, Bonami & Strnadová (2019) assume that morphological paradigms are structured by content-based contrasts, following Štekauer (2014). They then define the paradigmatic systems, i.e. the morphological paradigms, as sets of fully aligned families (11). For instance, in the paradigmatic system in Table 3, the action nouns in the second column are all in the same semantic relations with the other members of their families. One consequence of this definition is that the families that make up a paradigmatic system are all of the same size.

(11) A PARADIGMATIC SYSTEM is any set of morphological families of the same arity $P = \{(w_1^1, ..., w_n^1), ..., (w_1^m, ..., w_n^m)\}$ such that for any strictly positive $i, j \leqslant n$, all of $(w_i^1, w_j^1), ..., (w_i^m, w_j^m)$ are aligned pairwise by the same aligning relation.

Besides, Bonami & Strnadová (2019) consider that paradigm cells contain sets of units (either lexemes or inflected forms) and not single units in order to account for non-canonical phenomena such as defectiveness, rival overabundant forms in inflection or n-uplets in derivation. Defectiveness is described by cells containing empty sets and n-uplets by cells containing sets of n units. In this way, the size of the families in a paradigm remains constant. In addition, Bonami & Strnadová (2019) use implicative entropy to show that the interpredictability between the inflected forms of a lexeme is very similar to the interpredictability between the lexemes in a morphological family. Notice that the paradigmatic systems defined by Bonami & Strnadová (2019) correspond to the macro-paradigms proposed by Carstairs (1987, p. 69) for inflection (12). A macro-paradigm is a superposition of inflectional Paradigms$_2$ that realize Paradigms$_1$ which display predictable formal variations.

9

(12) A MACRO-PARADIGM consists of: (a) any two or more similar paradigms whose inflectional differences either can be accounted for phonologically, or else correlate consistently with differences in semantic or lexically determined syntactic properties (like gender); or (b) any paradigm which cannot be thus combined with other paradigm(s).

3.4 Construction morphology

Construction Morphology (CxM) is another framework that could be used for the paradigmatic description of morphology. CxM is developed by Booij (2010) for both inflection and derivation. The description of inflectional morphology in CxM is discussed in detail in (Booij, 2017). CxM is theoretically based on HPSG's Hierarchical Lexicon (Pollard & Sag, 1994) and is very similar to other models based on the Hierarchical Lexicon like TUHL (see below). In CxM, constructions represent both lexemes and word formation processes. Processes are represented by constructional schemas, that is, abstract structures which associate the formal and semantic properties of existing complex words, and that state how new words can be formed. For example, schema (13) describes how the form (on the left part of the arrow) and the meaning (on the right part of the arrow) of French agent nouns suffixed in -*eur* are related.

(13) $< [[x]_{Vi} \text{ eur}]_{Nj} \leftrightarrow [\text{he who } [\text{SEM}]_i]_j >$

(13) indicates that the form of the agent noun is obtained by suffixing -*eur* to the form $[x]_{Vi}$ of a verb and that its meaning is defined by the paraphrase 'he who $[\text{SEM}]_i$' where $[\text{SEM}]_i$ is the meaning of that same verb. As in PA, indexes i and j are used to identify the properties of the verb (index i) and of the derivative (index j). Similarly, the association of the form and the meaning of the base verb could be represented by a trivial construction like (14). As we can see, (i) meaning and form are specified separately in CxM; (ii) they are explicitly associated with a double arrow; (iii) the indexing system allows the derived lexemes to be connected to their base.

(14) $< [x]_{Vi} \leftrightarrow [\text{SEM}]_i >$

In CxM, the constructions are part of a hierarchical lexicon, i.e. a multiple inheritance network which can be used for instance to account for the competition between the suffixes that form the action nouns of the second column of Table 3. These constructions, described by schemas (15a), (15b) and (15c), are sub-types of a generic construction (16) with an abstract suffix represented by the variable *suf*. The inheritance relations are illustrated in (17).

(15) a. $< [[x]_{Vi} \text{ age}]_{Nj} \leftrightarrow [\text{act of } [\text{SEM}]_i]_j >$
      b. $< [[x]_{Vi} \text{ ment}]_{Nj} \leftrightarrow [\text{act of } [\text{SEM}]_i]_j >$
      c. $< [[x]_{Vi} \text{ ation}]_{Nj} \leftrightarrow [\text{act of } [\text{SEM}]_i]_j >$

(16) $< [[x]_{Vi} \text{ } suf]_{Nj} \leftrightarrow [\text{act of } [\text{SEM}]_i]_j >$

(17) $< [[x]_{Vi} \text{ } suf]_{Nj} \leftrightarrow [\text{act of } [\text{SEM}]_i]_j >$

$$(15a) \quad (15b) \quad (15c)$$
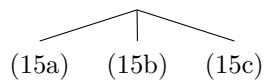
Booij & Masini (2015) add to CxM the possibility to define sets of paradigmatically related construction schemas called SECOND ORDER SCHEMAS and represented by the $\approx$ symbol. Second-order schemas are designed to represent indirect derivational relations or as Booij & Masini (2015,

51) put it, "state the formal and semantic correlation between two classes of words with the same degree of morphological complexity." Second-order schemas transcribe in CxM the notion of non-oriented WFR found in many lexeme-based approaches, such as (Haspelmath & Sims, 2010, 50). They can be used to represent abstract paradigms as illustrated in (18). (18) is a representation of a three cells abstract paradigm whose instances include the first six families in Table 3.

(18) $< [x]_{Vi} \leftrightarrow [\text{SEM}]_i > \approx < [[x]_{Vi}\text{ -eur}]_{Nj} \leftrightarrow [\text{he who }[\text{SEM}]_i]_j > \approx$
$< [[x]_{Vi}\ suf]_{Nk} \leftrightarrow [\text{act of }[\text{SEM}]_i]_k >$

3.5 Relational morphology

Relational Morphology (RM) is a morphological implementation of PA proposed by Jackendoff & Audring (2020b). It parallelizes the morphological descriptions into three levels of representation: phonological, morphosyntactic and semantic. Like PA, RM uses constant and variable indexes to identify the objects involved in the morphological description and to map the representations in the three levels on one another. On the other hand, RM is very close to CxM, in particular because all the objects (words, phrases, schemas, etc.) are constructions. For example, (19) describes the formation of French action nouns in -age where the variable index y identifies an abstract object (i.e. a pattern) having three properties: it is a masculine deverbal noun; its form ends in /ɑʒ/; it describes an action (ACTION). The base verb, identified by the variable index x is a predicate (PRED) formally unspecified /···/. (19) includes a third object indexed by 80, namely a nominalizing affix (NZR) whose form is /ɑʒ/.

(19) Semantics: $[\text{ACTION}(\text{PRED}_x)]_y$
Morphosyntax: $[_{Nmas}\ V_x\ \text{NZR}_{80}]_y$
Phonology: $/\cdots_x\ \text{ɑʒ}_{80}/_y$

As in CxM, the use of indexes allows for a simple and natural representation of discrepancies. However, the two models have some differences. One of them is that affixes are conceived in RM as units in their own right and are identified by their own index. For example in (19), index 80 identifies the nominalization suffix -age.

Like CxM, RM has an inheritance system. For example, (19) inherits from the abstract deverbal nominalization schema (20). The object identified by 80 in (19) is an instance of the one identified by the variable index w in (20) and the exponent /ɑʒ/ is an instance of the unspecified sequence /···/. (20) also generalizes the suffixations in -ment and -ion and the formation of verb-to-noun conversion because /···/ can also be instantiated by a phonologically null sequence.

(20) Semantics: $[\text{ACTION}(\text{PRED}_x)]_y$
Morphosyntax: $[_N\ V_x\ \text{NZR}_w]_y$
Phonology: $/\cdots_x\ \cdots_w/_y$

RM also has an equivalent of CxM's second-order schemas called SISTER SCHEMAS. For example, the three sister schemas in (21) are analogous to the second-order schema (18). However, unlike CxM, RM uses a specific index class to identify lexical items shared by sister schemas. For example, the index α specifies that the three sister schemas in (21) share the lexical item x. In this example, α expresses sorority whereas x, y and z identify the individual schemas (we arbitrarily chose index 78 to represent the suffix -eur).

(21)   a.   b.   c.

Semantics: $\mathrm{PRED}_{\alpha,x}$ $[\mathrm{AGENT}(\mathrm{PRED}_{\alpha,x})]_y$ $[\mathrm{ACTION}(\mathrm{PRED}_{\alpha,x})]_z$

Morphosyntax: V_x $[_N\ \mathrm{V}_x\ \mathrm{AGT}_{78}]_y$ $[_N\ \mathrm{V}_x\ \mathrm{NZR}_w]_z$

Phonology: $/\cdots/_x$ $/\cdots_x\ \text{œʁ}_{78}/_y$ $/\cdots_x\cdots_w/_z$

3.6 Typed Underspecified Hierarchical Lexicon

Historically, CxM was preceded by a very similar model, TUHL, proposed by Koenig (1999). Like CxM and other frameworks (Krieger & Nerbonne, 1993; Riehemann, 1998), TUHL is based on the Hierarchical Lexicon of Pollard & Sag (1994). It is designed for both inflection and derivation. In TUHL, all linguistic signs, including the lexemes and their inflected forms, are represented as types organized in a multi-inheritance hierarchy illustrated in Figure 1. These types can also be described by means of ATTRIBUTE-VALUE MATRICES (AVM) such as (22) for the type *laveur*. In this AVM, the value of the path μ-STRUCT|DGHTR is the description of the morphological structure of *laveur*.
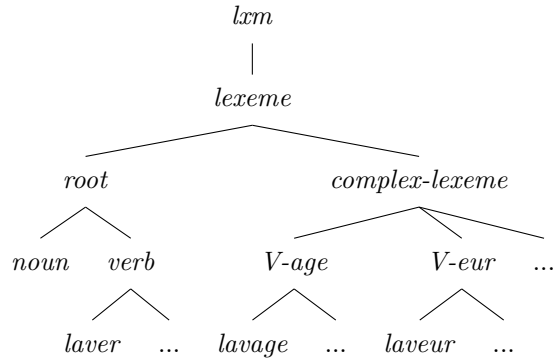
Figure 1: Excerpt of a THUL type hierarchy that contains the lexemes of the morphological family of *laver*. *lxm* represents a linguistic sign. This hierarchy is adapted from (Koenig, 1999, p. 93)

$$(22)\quad \begin{bmatrix} \textit{laveur} \\ \text{PHON} & \begin{bmatrix} \text{FORM}\ \boxed{1}\ \oplus\ /\text{œʁ}/ \end{bmatrix} \\ \text{CAT} & n \\ \text{CONT} & \begin{bmatrix} \textit{agent-of} \\ \text{ARG}\ \boxed{2} \end{bmatrix} \\ \mu\text{-STRUCT}\,|\,\text{DGHTR} & \begin{bmatrix} \textit{laver} \\ \text{PHON}\ \boxed{1}\ /\text{lav}/ \\ \text{CAT} & v \\ \text{CONT}\ \boxed{2}\ \textit{'wash'} \end{bmatrix} \end{bmatrix}$$

TUHL does not provide an explicit representation of morphological families nor derivational paradigms. However, we can extend the existing hierarchy with an additional type *family* and list its members in the value of a MEMBERS attribute as in the representation of the morphological family

of *laver* (23). The *family-laver* type inherits from the more abstract type *paradigm-V-Nage-Neur*, which represents a three cell paradigm (24).

(23)
$$
\begin{bmatrix}
\textit{family-laver} \\[2pt]
\text{MEMBERS} \left\langle \boxed{3}
\begin{bmatrix}
\textit{laver} \\
\text{PHON } \boxed{1} \ /\text{lav}/ \\
\text{CAT } \quad v \\
\text{CONT } \boxed{2} \ \text{'wash'}
\end{bmatrix},
\begin{bmatrix}
\textit{lavage} \\
\text{PHON} \qquad \boxed{1} \oplus /a\textipa{Z}/ \\
\text{CAT} \qquad n \\[4pt]
\text{CONT} \qquad \begin{bmatrix} \textit{act-of} \\ \text{ARG } \boxed{2} \end{bmatrix} \\[6pt]
\mu\text{-STR} \,|\, \text{DGHTR } \boxed{3}
\end{bmatrix},
\begin{bmatrix}
\textit{laveur} \\
\text{PHON} \qquad \boxed{1} \oplus /\text{œ}\textrm{ʁ}/ \\
\text{CAT} \qquad n \\[4pt]
\text{CONT} \qquad \begin{bmatrix} \textit{agent-of} \\ \text{ARG } \boxed{2} \end{bmatrix} \\[6pt]
\mu\text{-STR} \,|\, \text{DGHTR } \boxed{3}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

(24)
$$
\begin{bmatrix}
\textit{paradigm-V-Nage-Neur} \\[2pt]
\text{MEMBERS} \left\langle \boxed{3}
\begin{bmatrix}
\textit{verb} \\
\text{PHON } \boxed{1} \\
\text{CAT } \quad v \\
\text{CONT } \boxed{2}
\end{bmatrix},
\begin{bmatrix}
\textit{N-age} \\
\text{PHON} \qquad \boxed{1} \oplus /a\textipa{Z}/ \\
\text{CAT} \qquad n \\[4pt]
\text{CONT} \qquad \begin{bmatrix} \textit{act-of} \\ \text{ARG } \boxed{2} \end{bmatrix} \\[6pt]
\mu\text{-STR} \,|\, \text{DGHTR } \boxed{3}
\end{bmatrix},
\begin{bmatrix}
\textit{N-eur} \\
\text{PHON} \qquad \boxed{1} \oplus /\text{œ}\textrm{ʁ}/ \\
\text{CAT} \qquad n \\[4pt]
\text{CONT} \qquad \begin{bmatrix} \textit{agent-of} \\ \text{ARG } \boxed{2} \end{bmatrix} \\[6pt]
\mu\text{-STR} \,|\, \text{DGHTR } \boxed{3}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

4 ParaDis

ParaDis is very close to the models proposed by Bochner (1993) and Bonami & Strnadová (2019). Its central idea is that all morphologically relevant regularities are paradigmatic, including the formal, categorical and semantic ones. As a result, families and paradigms are extended to these levels of representation. In other words, we consider that the formal, categorical and semantic properties of the morphologically related lexemes are organized paradigmatically. Moreover, lexemes too form paradigms which belong to a fourth level that we call the MORPHOLOGICAL LEVEL. In short, ParaDis has three levels of representation (formal, categorical, semantic) and one (morphological) level of structuring; all four contain families and paradigms.

4.1 Formal description

We first give a formal description of the model where we present the objects and operations introduced and used in the remainder of Section 4. The starting point is the lexicon. It is seen as a graph, i.e. as a set of lexemes connected by lexical relations. More precisely, we are interested here in the subpart of the lexicon determined by morphological relations, i.e. regular relations of form and meaning between pairs of lexemes.[1] This subpart, we call MORPHOLOGICAL LEVEL, is constituted by the morphological relations and the lexemes they connect. The morphological relations define a graph on the set of lexemes. The connected components of this graph are the MORPHOLOGICAL FAMILIES.

Within these morphological families there are connected subgraphs consisting of edges that describe exactly the same form and meaning relations. These subgraphs are PARADIGMATIC FAMILIES. Superposition of paradigmatic families that have the same form and meaning relations forms MORPHOLOGICAL PARADIGMS. Superposition can be defined in two ways: (1) a morphological paradigm can be defined as a set of paradigmatic families equipped with an alignment relation defined on their

---

[1]We omit the categorical dimension in this first part of the description.

members such that the form and meaning contrasts of the aligned lexemes with the other members of their families are identical; (2) equivalently, paradigmatic families can be ordered into tuples of n lexemes so that the lexemes in i-th position in all the families of the paradigm are aligned (i.e., have the same form and meaning contrasts with the j-th member of their families, for any j such that $1 \leq j \leq n$ and $j \neq i$). This second definition amounts to representing paradigms as tables and families as rows in these tables. Note that these tables are only one possible form of presentation of the paradigms. On the other hand, morphological paradigms are lexical objects just as lexemes and lexeme families. As a consequence, they are not part of the grammar.

The columns of the tables constitute series of lexemes (or MORPHOLOGICAL SERIES). A series of lexemes is a set of lexemes aligned with one another within a morphological paradigm, or in other words, a set of lexemes of the same rank in the aligned paradigmatic families, when represented as tuples. Morphological paradigms are more constrained in ParaDis than in (Bonami & Strnadová, 2019) because we require the aligned relations to correspond to the same contrasts of form and the same contrasts of meaning. In other words, ParaDis' morphological paradigms are homogeneous in form and meaning.

LEXEMES consist of a phonological representation, a categorical representation and a semantic representation. Each of them belongs to a distinct dimension we call LEVEL OF REPRESENTATION. The representations (i.e. the objects) in these levels are connected by regular relations of contrast, i.e. relations of contrast found between several objects in the same level. These relations form graphs. The connected components of these graphs define families of objects in the corresponding level of representation, namely FORMAL FAMILIES, CATEGORICAL FAMILIES and SEMANTIC FAMILIES. Networks of identical contrasts define superposable subfamilies equivalent to paradigmatic families. Superposition of these subfamilies yields FORMAL PARADIGMS, CATEGORICAL PARADIGMS and SEMANTIC PARADIGMS. Similarly, FORMAL SERIES (i.e., series of formal objects), CATEGORICAL SERIES and SEMANTIC SERIES can be defined from the formal, categorical and semantic paradigms respectively. Families and series are objects of the model in their own right, as are paradigms and lexemes.

In ParaDis, the formal, categorical and semantic levels of representation are distinguished dimensions of a lexical description. On the other hand, the morphological level is not a simple combination of these levels since it is needed to describe certain morphological phenomena that involve constraints on lexemes, families of lexemes, and paradigms of lexemes or that refer to them. The morphological level fulfills two functions within the model. On the one hand, it is the level where the contradictory requirements from the three other levels of representation are arbitrated. On the other hand, it is the level where trade-offs are recorded.

Morphological paradigms can be arbitrarily superposed to form DERIVATIONAL PARADIGMS. A superposition of morphological paradigms can be defined as a set of morphological paradigms equipped with a superposition relation defined on their morphological series. A derivational paradigm is therefore a set of tuples of morphological series that can be represented as a table of morphological series. We call the columns of this table DERIVATIONAL SERIES. An important feature of derivational series is that they may contain empty positions. More precisely, the i-th position in a derivational series is empty when the corresponding morphological series does not exist in the i-th superposed morphological paradigm. Superposition relations are subject to the following constraints: (1) at least one series of each morphological paradigm is superposed on another series of another morphological paradigm (i.e., a superposed morphological paradigm cannot be disconnected from all the other ones); (2) a morphological series cannot be part of more than one derivational series; (3) all morphological series are superposed in the derivational paradigm; (4) a derivational

series contains at least one morphological series (i.e., derivational series cannot be empty).

4.2 Independent levels

A key contribution of lexematic morphology is the separation of the formal, categorical and semantic levels and their complete independence. This independence can be understood in two ways.[2] (*i*) Independence can be local and then follows directly from the notion of lexeme: the formal, categorical and semantic properties of each lexeme, each paradigmatic family and each morphological paradigm are described separately and independently. (*ii*) Independence is global when each level of representation does its thing and when the association is between complete representations of the three levels. This idea has also been proposed by Koenig (1999, 155) to account for the inflection of Breton endocentric compounds.

In line with PA, independence of the levels of representation in ParaDis is both local and global. More specifically, representations at one level account for regularities that are specific to that particular level and inaccessible to the other ones. For example, the morphophonological conditions that bear on the size of the derived lexemes do not depend on the meaning of these lexemes and are therefore only relevant at the formal level. Conversely, the fact that a state is a property that an entity can acquire is independent of the form of the adjective that denotes this state and of the form of the inchoative predicate that expresses the acquisition of this state.

4.3 Paradigm superposition

In ParaDis, morphological paradigms can be superposed just as families are in the paradigms. Superposed morphological paradigms form derivational paradigms. Superposition allows for generalizations and analyses that are more flexible and complex than the ones described by individual morphological paradigms, and that can be adjusted to the needs of the analysis, namely to the linguists' intuition. For example, the paradigm presented in Table 3 is a semantically motivated superposition of four homogeneous morphological paradigms, the first containing the families of *laver* and *saler*, the second the families of *former* and *fonder*, the third the families of *lancer* and *ronfler* and the fourth the families of *danser* and *voler*. Derivational paradigms correspond to the macro-paradigms proposed by Carstairs (1987) for inflection (12).

As we saw in Section 4.1, the superposition of morphological paradigms within derivational paradigms is subject to four constraints. However, these constraints are weak enough to allow in theory for the description of any kind of semantic or formal regularities. In practice, superposition is only used to state semantically motivated generalizations where formal variations are ignored,[3] as illustrated in Figure 3.

In Figure 2, the families in *-age* and *-eur* are separated from those in *-ation* and *-ateur* and those in *-ment* and *-eur*. The morphological paradigms they form are identified by their morphological series: the first one, that of the families of *laver* and *saler*, is made up of the morphological series M1, M2 and M3; in other words, we have MP1 = (M1, M2, M3); similarly, the second morphological paradigm is MP2 = (M4, M5, M6), the third is MP3 = (M7, M8, M9) and the fourth is MP4 = (M10, M11, M12). The superposition results in a derivational paradigm that could be defined as a triplet of quadruplets: DP1 = ((M1, M4, M7, M10), (M2, M5, M8, M11), (M3, M6, M9, M12)).

---

[2]We thank the anonymous reviewer who provided us this explicit formulation.

[3]They can be used to ignore categorical variations, for example, in order to account for the fact that French derivatives in *-able* can be formed from verbs or nouns (Hathout et al., 2003).

As a consequence, the series in the same position in each quadruplet are aligned with each other as shown in Figure 3. DP1 corresponds exactly to the paradigm of Table 3.

$$
\text{MP1} =
\begin{array}{|l|l|l|}
\hline
\text{M1} & \text{M2} & \text{M3} \\
\hline
\textit{laver} & \textit{lavage} & \textit{laveur} \\
\textit{saler} & \textit{salage} & \textit{saleur} \\
\hline
\end{array}
\qquad
\text{MP2} =
\begin{array}{|l|l|l|}
\hline
\text{M4} & \text{M5} & \text{M6} \\
\hline
\textit{former} & \textit{formation} & \textit{formateur} \\
\textit{fonder} & \textit{fondation} & \textit{fondateur} \\
\hline
\end{array}
$$

$$
\text{MP3} =
\begin{array}{|l|l|l|}
\hline
\text{M7} & \text{M8} & \text{M9} \\
\hline
\textit{lancer} & \textit{lancement} & \textit{lanceur} \\
\textit{ronfler} & \textit{ronflement} & \textit{ronfleur} \\
\hline
\end{array}
\qquad
\text{MP4} =
\begin{array}{|l|l|l|}
\hline
\text{M10} & \text{M11} & \text{M12} \\
\hline
\textit{danser} & \textit{danse} & \textit{danseur} \\
\textit{voler} & \textit{vol} & \textit{voleur} \\
\hline
\end{array}
$$

Figure 2: Four morphological paradigms. The paradigms are identified by the morphological series listed in the headers of the tables.

$$
\text{DP1} =
\begin{array}{|l|l|l|}
\hline
\text{M1} & \text{M2} & \text{M3} \\
\text{M4} & \text{M5} & \text{M6} \\
\text{M7} & \text{M8} & \text{M9} \\
\text{M10} & \text{M11} & \text{M12} \\
\hline
\end{array}
$$

Figure 3: Derivational paradigm resulting from the superposition of four morphological paradigms.

4.4   Correspondence

The formal, categorical and semantic levels of representation are independent without being totally disconnected from each other. A system of CORRESPONDENCEs connects them to the morphological level in order to ensures the cohesion of the different parts of the model. Formally, correspondences are binary relations that connect a cell in the morphological paradigm to a cell in a paradigm of another level of representation. ParaDis imposes two conditions on correspondences. (*i*) Each cell in the morphological level must be in correspondence with one and only one cell in each of the three other levels of representation. In other words, every lexeme in the morphological level is in correspondence with one representation in the three other levels (formal, categorical and semantic). For example, the lexeme *laveur* of paradigm MP1 is in correspondence with its formal representation /lavœʁ/ which is an object of the formal level of representation, with a categorical representation N which is located in the categorical level and with a semantic representation 'he who washes' which belongs to the semantic level. Conversely, these three representations are associated through this lexeme. (*ii*) The cells in a morphological series are all related to cells that belong to the same series in each of the three levels. In other words, if x_1 and x_2 are two cells of the same morphological series, such that x_1 is related to y_1 in one of the levels of representation and x_2 is related to y_2 in the same level, then y_1 and y_2 must belong to the same series in that level. One consequence of this constraint is that the number of series in a morphological paradigm is greater than or equal to the largest number of series in the three paradigms with which it is in correspondence. Correspondences also concern families, series and paradigms: every morphological family (resp. series, resp. paradigm) is in correspondence with a family (resp. series, resp. paradigm) in the three other levels. For example, the family (*laver*, *lavage*, *laveur*) is in correspondence with the formal family (/lav/, /lavaʒ/, /lavœʁ/), with the categorical family (V, N) and with the semantic family

('to wash', 'act of washing', 'he who washes'). As a consequence, the description of a morphological paradigm involves a formal paradigm, a categorical paradigm and a semantic paradigm. On the other hand, there is no direct connection between the formal, categorical and semantic levels, which makes them fully independent (see Figure 4).

ParaDis and PA have similar architectures, but their conception present a notable difference, namely the way in which the three levels of representation are interconnected. In PA, all the levels are interconnected. In ParaDis, correspondences are made through a morphological level. This also distinguishes ParaDis from other paradigm-based models, in particular from the ones of Bochner (1993) and Bonami & Strnadová (2019).

Figure 4 provides a graphical representation of the structure of the four paradigms involved in the analysis of MP1, and of the correspondences that connect the formal, categorical and semantic paradigms to the morphological one. The figure shows that the graphs defined by the different levels are not necessarily identical and may have different shape and size.[4] These graphs are subject to three constraints: (*i*) they must be connected graphs otherwise there would be no paradigm; (*ii*) every vertex in a morphological graph must correspond to exactly one formal vertex, one categorical vertex and one semantic vertex; (*iii*) all the vertices in a morphological graph must be connected to vertices of the same formal graph, the same categorical graph and the same semantic graph (i.e. a morphological paradigm must be connected to exactly one formal paradigm, one categorical paradigm and one semantic paradigm). On the other hand, it is not necessary that all edges in the morphological graph are related to an edge in the other three graphs. For example, the edge *lavage*:*laveur* in the morphological graph is not in correspondence with any edge in the formal graph. Conversely, no constraints other than connectedness affects the graphs of the other three levels. These graphs may contain vertices and edges without morphological correspondents. The above constraint (*ii*) ensures that any lexeme has a description in each of the three levels of representation. Because constraint (*ii*) only imposes the connectedness of the graph, ParaDis is able to account for all the discrepancies and idiosyncrasies that exist in the derivational lexicon of a language like French. An example of non isomorphic families and paradigms is provided by the morphological family (*déisme* 'deism', *déiste* 'deist'). Semantically, this family is similar to (*fétiche* 'fetish', *fétichisme* 'fetishism', *fétichiste* 'fetishist'), *fétichisme* being the 'belief in the power of the fetishes', and a *fétichiste* being a 'believer in the power of the fetishes'. This family is therefore in correspondence with the semantic family ('power of the fetishes', 'belief in the power of the fetishes', 'believer in the power of the fetishes'). Similarly, (*déisme*, *déiste*) is in correspondence with the semantic family ('existence of God', 'belief in the existence of God', 'believer in the existence of God'), with one difference: the first member of the semantic family is not in correspondence with a lexeme of the morphological family. Its presence in the family is however semantically motivated by the fact that a concept 'believer in *X*' refers *de facto* to a concept '*X*'. This example illustrates the situation where a morphological family is in correspondence with a semantic family in which some elements have no lexical realization. Conversely, a vertex or an edge at the formal, categorical or semantic level may have more than one correspondent in the morphological graph, as in the French paradigms of toponyms, i.e. names of country (C), demonyms, i.e. names of inhabitants (I) and names of languages (L) (Roché, 2008, 2017; Molinier, 2018; Schalchli & Boyé, 2018). For example, the morphological family (*Italie*C 'Italy', *Italien*I 'Italian person', *italien*L 'Italian language') is in correspondence with the formal family (/itali/, /italjɛ̃/) where /italjɛ̃/ is connected to both *Italien*I and *italien*L; similarly, the formal edge /itali/:/italjɛ̃/ corresponds to the morphological

---

[4]An alternative view of these connected graphs could be to have weighted complete graphs at all four levels with uncorrelated weights.

edges *Italie*$_C$:*Italien*$_I$ and *Italie*$_C$:*italien*$_L$. Likewise, in Figure 4, the V:N edge in the categorical graph is connected to the *laver*:*lavage* and *laver*:*laveur* edges of the *laver* family. Another example is provided by the morphological family (*danser*, *danse*, *danseur*), part of paradigm MP4, where the verb and the action noun have the same stem /dɑ̃s/. In this family, the edge *danser*:*danse* does not have a formal correspondent (see Section 4.9).
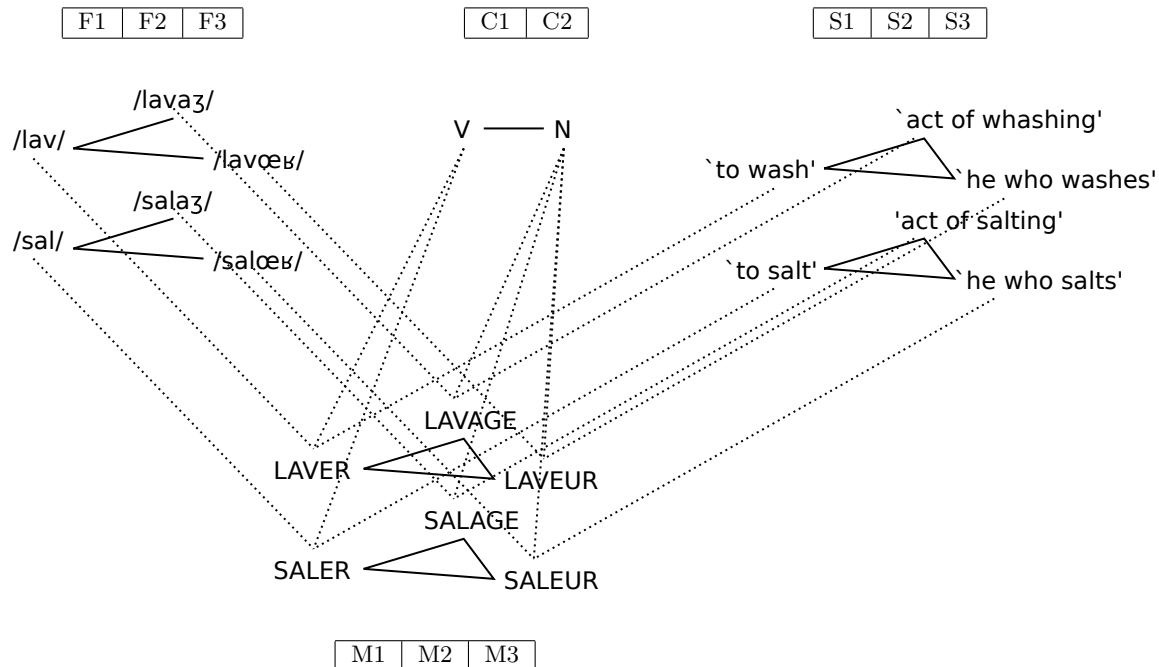

Figure 4: Schematic representation of the morphological paradigm MP1 = (M1, M2, M3), formal paradigm FP1 = (F1, F2, F3), categorical paradigm CP1 = (C1, C2) and semantic paradigm SP1 = (S1, S2, S3). The morphological paradigm is in correspondence with the other three. Paradigms are superpositions of families. Families are represented as connected graphs. The series that make up the paradigms at each level are listed above or below the graphs.

4.5   Formal paradigm

Before describing the formal paradigms themselves, we need to determine what objects do their cells contain. The answer rests on the assumption that the formal representations and the formal relations involved in the description of derivational regularities are based on the inflectional organization of the lexicon. In particular, we consider that inflectional regularities emerge from the lexemes and their themes, i.e. their stems, as proposed by Aronoff (1994); Bonami & Boyé (2002); Boyé (2011) and that these stems can also be used to describe the formal relation between lexemes. Therefore, we here suppose that the cells of a formal paradigm contain stems of the lexemes they represent. One consequence of this assumption is that the same stem is involved in all the formal relations that exist between a lexeme and the other lexemes of its family since a cell can only contain one representation. We also make the additional assumption that the stems that make up the

F1	F2	F3
$/f_1/$	$/f_1\text{-a}ʒ/$	$/f_1\text{-œʁ}/$
/lav/ (F1,1)	/lavaʒ/ (F2,1)	/lavœʁ/ (F3,1)
/sal/ (F1,2)	/salaʒ/ (F2,2)	/salœʁ/ (F3,2)

Table 4: Formal paradigm FP1 in correspondence with MP1. The upper part (first line after the header) describes the abstract formal paradigm and the lower part (second and third lines) the concrete formal paradigm. The cells in the concrete paradigm are identified by a unique label of the form Fi,j where i is the index of the formal series and j the index of the formal family the cell belongs to.

formal paradigms are the ones that bring out the most prominent formal regularities in the lexicon. The emergence of the inflectional stems and their selection for the description of derivational regularities are issues beyond the scope of this article. In what follows, we simply hypothesize that the stems contained in the cells of the formal paradigms allow for the description of the derivational relations we are interested in.

We can now return to the description of the formal paradigms and their connection to the morphological paradigms. A formal paradigm consists of two parts: a concrete formal paradigm, and an abstract one. The CONCRETE FORMAL PARADIGM is a superposition of formal families that can be represented in a table. For example, the concrete formal paradigm in correspondence with MP1 is represented by the last two lines of Table 4. As we just indicated, its cells contain stems, e.g., /lav/ for the verb *laver*, /lavaʒ/ for the noun *lavage* and /lavœʁ/ for the noun *laveur* in the second line of the table. These cells are identified by two coordinates. The first is the index of the series which contains the cell and the second is the index of its family. For example, the second cell of the formal paradigm in Table 4 (/lavaʒ/) is labeled F2,1 because it belongs to the formal series F2 and to the formal family 1 (/lav/, /lavaʒ/, /lavœʁ/). More generally, we assume that within every level of representation, every family and every series has a unique identifier.

The first line after the header in Table 4 represents the ABSTRACT FORMAL PARADIGM. It describes the relations between the cells of the concrete formal paradigm and corresponds to the implicative structures used by Blevins (2006) and Bonami & Beniamine (2016). For example the forms in F2 can be obtained by adding the suffix /aʒ/ to the forms in F1 (e.g., /lav/ → /lavaʒ/). Similarly, the forms in F3 can be constructed from the ones in F1 by adding the suffix /œʁ/. On the other hand, the relation between the forms in /aʒ/ and in /œʁ/ has not been included in the abstract paradigm because it is not regular enough as illustrated in (25). The presence of a form ending in /aʒ/ in a family (*esclavage*, *laitage*) is not predictive of the presence in the same family of a form ending in /œʁ/ because /aʒ/ is also the exponent of suffixation processes that derive state nouns (*esclavage*) and nouns of collections (*laitage*). Conversely, a formal family may contain a form in /œʁ/ (*ronfleur*, *diffuseur*) but no form in /aʒ/ because action nouns can be formed by conversion or by means of suffixation processes with other exponents such as /mã/ (*ronflement*) or /jɔ̃/ (*diffusion*). The abstract paradigm therefore defines a graph represented as in Figure 5.

(25)   a. *lait* 'milk', *laitage* 'dairy', *laitier* 'milk man'
        b. *esclave* 'slave', *esclavage* 'slavery'

19

c. *ronfler, ronflement, ronfleur*

d. *diffuser* 'broadcast', *diffusion* 'broadcasting', *diffuseur* 'broadcaster'
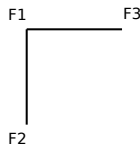

Figure 5: Graph defined by the abstract formal paradigm FP1.

4.6 Categorical paradigm

There is not yet a clear consensus in derivational morphology on the nature and content of the categorical descriptions of the lexemes (Anderson, 1992; Aronoff, 1994; Fradin, 2003). Most authors consider that they contain a grammatical category, e.g. verb, noun, adjective, adverb in English. This assumption has three consequences: (*i*) there are only a small number of categorical paradigms; (*ii*) categorical paradigms are trivial because they contain only one family each; (*iii*) the co-occurrence of categories in derivational families is regular and predictable. If a family contains a noun, then it probably contains an adjective (e.g. its relational adjective); if it contains a verb, then it normally contains a noun (e.g. the corresponding action noun); if it contains an adjective, then it also contains its quality noun; and so on. We therefore propose to consider that the categorical families form complete graphs.

C1	C2
c_1	c_2
V	N
(C1,1)	(C2,1)

Table 5: Categorical paradigm in correspondence with MP1. The upper part (first line after the header) describes the abstract categorical paradigm and the lower part (second line) the concrete categorical paradigm. The cells in the concrete paradigm are identified by a unique label of the form Ci,j.

The families of MP1 consist of one verb and two nouns. The corresponding categorical paradigm CP1 = (C1, C2) has therefore only two categories: verb and noun. All the nouns in the morphological paradigm MP1, (i.e. *lavage*, *laveur*, *salage*, *saleur*) are in correspondence with the same cell (C2,1) in the categorical paradigm. More generally, the same categorical representation can be shared by several members of the same morphological family, but also by members of different families that may belong to the same morphological paradigm or to different ones (see Section 4.9). The categorical paradigm in correspondence with MP1 can be represented as in Table 5. The abstract paradigm defines a graph that can be represented as in Figure 6. The example shows that correspondences may connect paradigms that neither contain the same number of families nor the same number of series. Note that categorical properties could also be described by means of feature representations. However, we prefer a graph-based description for homogeneity reasons.

20

The paradigms of all levels of representation can thereby be described by means of the same type of structure.
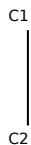
C1

|

C2

Figure 6: Graph defined by the abstract formal paradigm CP1.

The categorical level is the least complex level of representation. However, it is a level of representation in its own right. It would be difficult to merge it with the formal level (as it is done for example in CxM) or with the semantic level, because it is required for the description of some conversion relations. Conversion may connect lexemes which have the same form, (almost) the same semantic content and only differ in their categorical properties, e.g., the verb *voler* 'to steal' and its action noun *vol* 'theft'. The categorical level is also involved for instance in the constraints that in French impose syncretism between demonyms ($Italien_N$ 'inhabitant of Italy'), their relational adjective ($italien_A$ 'of Italians') and the relational adjective of corresponding toponym ($italien_A$ 'of Italy'). It also plays a key role in the prototypical realization of the pragmatic functions of reference, predication and modification as nouns, verbs and adjectives respectively (Croft, 1991) and in categorical transpositions (Kleiber, 1984; Roché, 2006), such as noun → relational adjective, adjective → quality noun, verb → action noun.

4.7 Semantic paradigm

The semantic paradigm that corresponds to MP1 is presented in Table 6. Its concrete paradigm contains semantic representations described by their relations with the meaning of the other lexemes in the family. For example, the meaning in cell S2,1 is 'act of washing' defined with respect to the meaning 'to wash' (S1,1). It is also the 'act performed by a washer', defined with respect to 'washer' (S3,1). These two definitions describe the same concept. Likewise, the meaning in S3,1 is defined with respect to the meaning in S1,1 ('he who washes') and in S2,1 ('he who performs a washing').

The semantic paradigm SP1 = (S1, S2, S3) includes the meanings of the families in paradigms MP1, MP2, MP3 and MP4 since these meanings are in exactly the same relations. These families instantiate the abstract paradigm described in the first line in Table 6. This abstract paradigm defines a complete graph (Figure 7) because the meanings in S2 and S3 can be cross-defined.
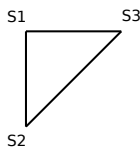
S1   S3

S2

Figure 7: Graph defined by the abstract semantic paradigm SP1.

Semantic paradigms can be conceived as descriptions of bundles of meanings similar to the semantic frames of Fillmore (1976). Furthermore, as discussed in Section 4.4, they are independent of the lexicalization of the meanings they contain.

S1	S2	S3
's_1'	'act of s_1' 'act performed by s_3'	'he who s_1' 'he who performs s_2'
'to wash' (S1,1)	'act of washing' 'act performed by a washer' (S2,1)	'he who washes' 'he who performs a washing' (S3,1)
'to salt' (S1,2)	'act of salting' 'act performed by a salter' (S2,2)	'he who salts' 'he who performs a salting' (S3,2)
'to train' (S1,3)	'act of training' 'act performed by a trainer' (S2,3)	'he who trains' 'he who performs a training' (S3,3)
'to found' (S1,4)	'act of founding' 'act performed by a founder' (S2,4)	'he who founds' 'he who performs a foundation' (S3,4)
'to launch' (S1,5)	'act of launching' 'act performed by a launcher' (S2,5)	'he who launches' 'he who performs a launching' (S3,5)
'to snore' (S1,6)	'act of snoring' 'act performed by a snorer' (S2,6)	'he who snores' 'he who performs a snoring' (S3,6)
'to dance' (S1,7)	'act of dancing' 'act performed by a dancer' (S2,7)	'he who dances' 'he who performs a dance' (S3,7)
'to steal' (S1,8)	'act of stealing' 'act performed by a thief' (S2,8)	'he who steals' 'he who performs a theft' (S3,8)

Table 6: Semantic paradigm in correspondence with MP1, MP2, MP3 and MP4. The upper part (first line after the header) describes the abstract semantic paradigm and the lower part (second and following lines) the concrete semantic paradigm. The cells in the concrete paradigm are identified by a unique label of the form Si,j.

4.8   Formalization of correspondences

In order to complete the representation of MP1, we need to indicate which formal, categorical and semantic cells are in correspondence with each cell in the morphological paradigm. In Table 7, the three parts of the expression in the third line of each cell reflect the three levels of the lexeme it contains. For example, the table states that the morphological cell M1,1 is in correspondence with the formal cell F1,1, with the categorical cell C1,1 and with the semantic cell S1,1.

4.9   The other three morphological paradigms

The other morphological paradigms in Figure 2 are analyzed in exactly the same way as MP1. MP2 = (M4, M5, M6) is in correspondence with a formal paradigm FP2 = (F4, F5, F6) described

M1	M2	M3
laver	*lavage*	*laveur*
(M1,1)	(M2,1)	(M3,1)
F1,1/C1,1/S1,1	F2,1/C2,1/S2,1	F3,1/C2,1/S3,1
saler	*salage*	*saleur*
(M1,2)	(M2,2)	(M3,2)
F1,2/C1,1/S1,2	F2,2/C2,1/S2,2	F3,2/C2,1/S3,2

Table 7: Full description of MP1. Cells are identified by unique labels Mi,j. The third line in each cell indicates the formal, categorical and semantic correspondents of that cell.

in Table 8. It is slightly different from FP1 because the forms in /asjɔ̃/ and in /atœʁ/ are fully interpredictable (Boyé, 2011, p. 50): the /asjɔ̃/ ending characterizes action nouns coined on the supine stem of verbs borrowed from Latin whose agentive derivatives end in /atœʁ/ (Bonami et al., 2009). In the abstract formal paradigm of FP2, the description of the indirect relation between the two derivatives involves a phonological sequence g which represents the stem common to the two forms. FP2 defines a complete graph represented in Figure 8. MP2 is in correspondence with the same categorical and semantic descriptions as MP1, namely CP1 and SP1. It is another illustration of how representations can be shared thanks to the independence of the levels of representations.

F4	F5	F6
$/f_4/$	$/f_4$-asjɔ̃$/$	$/f_4$-atœʁ$/$
	$\exists g,\ /f_5/=/g\text{-asjɔ̃}/\ \wedge\ /f_6/=/g\text{-atœʁ}/$	$\exists g,\ /f_5/=/g\text{-asjɔ̃}/\ \wedge\ /f_6/=/g\text{-atœʁ}/$
/fɔʁm/	/fɔʁmasjɔ̃/	/fɔʁmatœʁ/
(F4,3)	(F5,3)	(F6,3)
/fɔ̃d/	/fɔ̃dasjɔ̃/	/fɔ̃datœʁ/
(F4,4)	(F5,4)	(F6,4)

Table 8: Formal paradigm in correspondence with MP2. The cells of the concrete paradigm contain the formal representations of the lexemes *former*, *formation*, *formateur* (line 2) and *fonder*, *fondation*, *fondateur* (line 3).
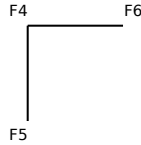
Figure 8: Graph defined by the abstract formal paradigm FP2.

The analysis of MP3 = (M7, M8, M9) also involves an additional formal paradigm, FP3 = (F7, F8, F9), similar to FP1, as shown in Table 9. The connected graph defined by FP3 is represented in Figure 9. MP3 is in correspondence with the same categorical (CP1) and semantic paradigms (SP1) as MP1.

Table 10 presents the formal paradigm FP4 = (F10, F11) involved in the analysis of MP4 =

F7	F8	F9
/f₇/	/f₇-əmã/	/f₇-œʁ/
/lãs/ (F7,5)	/lãsəmã/ (F8,5)	/lãsœʁ/ (F9,5)
/ʁɔ̃fl/ (F7,6)	/ʁɔ̃fləmã/ (F8,6)	/ʁɔ̃flœʁ/ (F9,6)

Table 9: Formal paradigm in correspondence with MP3. The cells of the concrete paradigm contain the formal representations of the lexemes *lancer, lancement, lanceur* (line 2) and *ronfler, ronflement, ronfleur* (line 3).
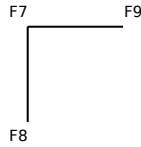

Figure 9: Graph defined by the abstract formal paradigm FP3.

(M10, M11, M12). The paradigm has only two series because the verbs and the action nouns share the same stem and are represented by the same cells in F10. Therefore, in the families of MP4, the cells in M10 and M12 are in correspondence with the same cell in F10. The graph defined by the abstract formal paradigm is represented in Figure 10. Like the other three morphological paradigms, MP4 is in correspondence with CP1 and SP1. As mentioned above, the analysis of the families of verb-based converted nouns (e.g., *danse ← danser*) shows the benefits of a separate and independent description of the formal, categorical and semantic properties. ParaDis explicitly represents the fact that conversion does not involve any formal modification: the verb and its action noun share the same stem (/dãs/), and are in correspondence with the same cell of the formal paradigm.

F10	F11
/f₁₀/	/f₁₀-œʁ/
/dãs/ (F10,7)	/dãsœʁ/ (F11,7)
/vɔl/ (F10,8)	/vɔlœʁ/ (F11,8)

Table 10: Formal paradigm in correspondence with MP4. The cells of the concrete paradigm contain the formal representations of the lexemes *danser, danse* in (F10,7), *danseur* in (F11,7), *voler, vol* in (F10,8) and *voleur* in (F11,8).
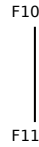
```
F10
 |
 |
 |
F11
```

Figure 10: Graph defined by the abstract formal paradigm FP4.

4.10 Constraints

In addition to the objects and structures we just presented, ParaDis includes a set of constraints which we alluded to several times (see also Section 7). Their detailed description will be the subject of a future publication.

ParaDis' constraints are similar to the ones used in Optimality Theory (Prince & Smolensky, 1993). They are gradable and may not be satisfied. On the other hand, they can combine and create *gang effects* in order to overcome stronger constraints. They may apply to objects of a single level, like the dissimilative constraints (OCP, McCarthy (1986)) which penalize the forms that contain successions of similar sounds. Constraints can also be placed on the correspondences between several levels, like the faithfulness constraint proposed in (Corbin, 2001; Hathout, 2009, 2011) which penalizes the derivatives with suppletive or allomorphic forms.

4.11 To sum up

The separate description of the formal, categorical and semantic paradigms in ParaDis allows the morphological paradigms MP1, MP2, MP3 and MP4 to be in correspondence with the same categorical and semantic paradigms. This reflects the intuition that the difference between the four morphological paradigms are only formal. The superposition of these paradigms described in Figure 3 gives priority to the semantic (and categorical) regularities over the formal ones. This fully accounts for the analysis of Bonami & Strnadová (2019), and also captures the intuition that lexemes are first defined by their meaning: *lavage* is the action noun of *laver* before being a derivative in *-age*. Its place in the lexicon is determined by its semantic properties. More generally, semantic properties organize the lexicon as a whole. As a result, it is semantic paradigms that define the structure of derivational paradigms. The prominence of the semantic regularities also reinforces the similarity with inflection where morphosyntactic regularities are primary and determine the structure of inflectional paradigms, and where formal variations divide the lexicon into inflectional classes.

ParaDis is also flexible and open enough to be used to describe inflectional paradigms at only a small additional cost. In this transposition, inflectional classes take the place of morphological paradigms. They are superpositions of lexemes exactly as morphological paradigms are superpositions of morphological families. The morphosyntactic properties of word forms are contained in morphosyntactic paradigms. As with categorical paradigms, each morphosyntactic paradigm contains a single family which lists the combinations of morphosyntactic features of the lexemes of a given inflectional class. The semantic paradigms are trivial and contain a single series since all inflected forms of a lexeme have the same lexical meaning.

5 Non canonical phenomena

Canonicity is a notion introduced by Corbett (2003, 2007) for inflection and extended to derivation in (Corbett, 2010). A morphological phenomenon is considered to be canonical if it conforms to a hypothetical ideal that serves as a theoretical benchmark for the characterization and comparison of the morphology of the languages of the world. The ideal is reached when a morphological system is formally transparent, structurally regular and directly interpretable, i.e. minimally ambiguous. For derivation, Corbett (2010, p. 142) proposes two principles (26).

(26)   Principle I: Canonical derived words have clear indicators of their synchronic status.
       Principle II: Canonical derived words are fully distinct from their base.

Hathout & Namer (2014a) implement these principles as a categorization of non-canonical constructions in terms of over- and under-marking. Bonami & Strnadová (2019), for their part, analyze several non-canonical derivational phenomena within their paradigmatic model. We saw in Section 3 that the cells of the paradigms they propose are sets of lexemes allowing the representation of defectiveness by empty sets and of overabundance by sets containing more than one word. In the following, we review some non-canonical configurations listed by Walther (2013, p. 186) in order to identify their possible equivalents in French derivational morphology and describe them in ParaDis.

M31	M32	M33	M34
cerise	*cerisier*	*cerisaie*	*ceriseraie*
fraise	*fraisier*	*fraisaie*	*fraiseraie*
amande	*amandier*	*amandaie*	*amanderaie*

M35	M36	M37	M38
myrtille	*myrtille*	*myrtillaie*	*myrtilleraie*

M39	M40	M41
abricot	*abricotier*	*abricoteraie*

M31	M32	M33	M34
M35	M36	M37	M38
M39	M40		M41

Figure 11: Excerpt from the derivational paradigm of French names of fruits, plants and plantations (Roché, 2011). The fruit names are *cerise* 'cherry', *fraise* 'strawberry', *amande* 'almond', *myrtille* 'blueberry' and *abricot* 'apricot'; plant names are derived from fruit names by suffixation in *-ier*; plantation names can be formed by suffixation in *-aie* or *-eraie*. The derivational paradigm results from the superposition of three morphological paradigms. The tables on the left describe morphological paradigms and the one on the right how they are superposed.

The paradigm of French names of fruits, trees and plantations in Figure 11 has been proposed by Roché (2011) to illustrate the contribution of derivational families and series to the analysis of several non-canonical phenomena. It can be described in ParaDis as a derivational paradigm resulting from the superposition of three morphological paradigms presented in the left part of Figure 11. These paradigms are in correspondence with a one-cell categorical paradigm which only contains the nominal category. They are also in correspondence with a single semantic paradigm composed of three series: a series of fruit names, a series of plant names and a series of plantation names ('s_{31}', 'plant that produces s_{31}' = 's_{32}', 'plantation of s_{32}'). On the other hand, they are in correspondence with three different formal paradigms: the first one is composed of four series forming the abstract paradigm ($/\int/$, $/\int$-je$/$, $/\int$-ɛ$/$, $/\int$-ərɛ$/$); the second one has three series forming

the abstract paradigm ($/\!\int\!/$, $/\!\int\!$-ɛ/, $/\!\int\!$-ərɛ/); the third one is also a three series paradigm but a different one: ($/\!\int\!/$, $/\!\int\!$-je/, $/\!\int\!$-ərɛ/). In Figure 11, the nouns in *-aie* and *-eraie* belong to distinct morphological series even if they refer to the same entities and are in correspondence with the same cells in the semantic paradigm. For example, *cerisaie* and *ceriseraie* are in correspondence with two different formal descriptions (/səʁizɛ/ and /səʁizəʁɛ/ respectively) but the same meaning ('cherry orchard').

Syncretism. A paradigm is said to be syncretic when it contains several cells with identical forms such as *myrtille*, denoting both a fruit (M35) and a plant (M36) in paradigm MP6 = (M35, M36, M37, M38). In this case, a single cell /miʁtij/ in the formal paradigm is in correspondence with two lexemes *myrtille* in MP6. Syncretism may also result from conversion as discussed in Section 4. The analysis in ParaDis of the two phenomena is identical: a single cell in the formal paradigm is in correspondence with several lexemes in the morphological paradigm.

Defectivness. A paradigm is defective when it contains empty cells. This is the case of the derivational paradigm in Figure 11 where the derived noun in *-aie* is missing in the family of *abricot*. This gap is revealed by the superposition of the three morphological paradigms that make up the derivational paradigm; there is no empty cell in MP7 = (M39, M40, M41). ParaDis accounts for the intuition that the void is only perceptible in comparison with the other paradigms MP7 is aligned with and that it only shows up at the derivational level. In other words, defects only concern the derivational paradigms.

Overabundance and n-uplets. Doublets (and more generally n-uplets) are another very frequent phenomenon in derivational morphology. For example, in the first paradigm of Figure 11, the names of plantation can be coined by suffixation in *-aie* on the stem of either the name of fruit, or the name of tree (with an alteration of /je/ in contact with /ɛ/). We consider that these two forms belong to distinct cells in the morphological paradigm because they co-occur on a regular basis. Therefore, these n-uplets are not treated as derivational equivalents of overabundance, because the latter presupposes that the co-occurring forms have identical contents. The case of doublets such as *rançonnage*, *rançonnement* in (27) is different (Bonami & Strnadová, 2019). Our analysis of these nouns is based on the existence of two different morphological paradigms within the same derivational paradigm (Figure 3). More specifically, *rançonnage* belongs to the family (*rançonner*, *rançonneur*, *rançonnage*), part of the morphological paradigm MP1, *rançonnement* belongs to the family (*rançonner*, *rançonneur*, *rançonnement*), part of MP3 and these two morphological families are in correspondence with the same semantic family. In this way, the analysis accounts for the fact that the co-occurrence between *rançonnage* and *rançonnement* is not predictable. It also reflects that this doublet results from the competition between the suffixations in *-age* and in *-ment*.

(27)   *rançonner* 'to rob', *rançonneur* 'robber', *rançonnage* 'robbery', *rançonnement* 'robbery'

Stem suppletion. Suppletive lexemes such as *carcéral* in the family of *prison* or *scolaire* in the family of *école* pose similar problems as shown in Table 11, borrowed from Bonami & Strnadová (2019). The three families in Table 11 correspond to a single semantic paradigm consisting of 4 series whose relations can be described using the abstract paradigm ('s_{51}', 'he who is in s_{51}', 'of s_{51}', 'to put in s_{51}').

commerce 'commerce'	commerçant 'storekeeper'	commercial 'commercial'	commercialiser 'to commercialize'
école 'school'	écolier 'schoolboy'	scolaire 'of school'	scolariser 'to send to school'
prison 'jail'	prisonnier 'inmate'	carcéral 'of prison'	emprisonner incarcérer 'to jail'

Table 11: Paradigm with stem suppletion (Bonami & Strnadová, 2019). The families of *école* and *prison* are compared to that of *commerce*. The first column contains location nouns, the second one, nouns denoting persons involved in activities taking place in this location, the third one, the relational adjectives of the nouns in column 1, and the last column, causative verbs expressing change of location. One part of the family of *école* is formed on the stem /ekɔl/ (*école*, *écolier*) and the other part on the stem /skɔl/ (*scolaire*, *scolariser*). The family of *prison* contains two causative verbs 'to jail', one coined on the stem /prizɔn/ (*emprisonner*) and the other on the stem /caʁseʁ/ (*incarcérer*), also used to form the adjective *carcéral*.

The analysis of the paradigm of Table 11 in ParaDis separates the families of *école* and *prison* into several formally homogeneous subfamilies: they are split in five morphological families, each belonging to a different morphological paradigm. These families and paradigms are presented in the left part of Figure 12. Their superposition, described in the right part, yields what Roché (2009) and Hathout (2011) call lexical families. More specifically, a LEXICAL FAMILY can be formally defined as the union of all the morphological families which are in correspondence with the same semantic family. According to this definition, the lexical family of *école* is {*école*, *écolier*, *scolaire*, *scolariser*}, the morphological families (*école*, *écolier*) and (*scolaire*, *scolariser*) being in correspondence with the same semantic family ('school', 'schoolboy', 'of school', 'to send to school'). Likewise, the lexical family of *prison* is {*prison*, *prisonnier*, *emprisonner*, *carcéral*, *incarcérer*}. It results from the union of morphological families (*prison*, *prisonnier*, *emprisonner*) and (*carcéral*, *incarcérer*) both related to the same semantic family ('jail', 'inmate', 'of prison', 'to jail'). These families are said to be lexical because from a lexical point of view *scolaire* is the relational adjective of *école* just as *commercial* is the relational adjective of *commerce*. The lexical families we just defined are different from the macro-families proposed by Bonami & Strnadová (2019) since a macro-family is the union of two or more families that are in complementary distribution. For example, *emprisonner* and *incarcérer* do not belong to the same macro-family but belong to the same lexical family.

The analysis of Figures 11 and 12 show how ParaDis differs from the model of Bonami & Strnadová (2019) in the treatment of non-canonical phenomena. In (Bonami & Strnadová, 2019), they are analyzed at the level of the family. These are made up of sets of lexical units in order to guarantee that the families of the same paradigm are all of the same size. In contrast, overabundance and gaps in ParaDis are not within the scope of families nor morphological paradigms, but only concern derivational paradigms. In this way, ParaDis reflects the intuition that a gap is only detectable in comparison with more complete data. Furthermore, ParaDis differentiates between occasional *n*-uplets and regular (i.e. predictable) overabundance, like the co-occurrence of names of plantation in *-aie* and *-eraie*. In contrast, this distinction cannot be described in the model of Bonami & Strnadová (2019) which only has one mechanism to account for both types of overabundance, namely the use of sets of lexemes.

M51	M52	M53	M54
commerce	commerçant	commercial	commercialiser

M55	M56
école	écolier

M57	M58
scolaire	scolariser

M59	M60	M61
prison	prisonnier	emprisonner

M62	M63
carcéral	incarcérer

M51	M52	M53	M54
M55	M56		
		M57	M58
M59	M60		M61
		M62	M63

Figure 12: Analysis of the paradigm of Table 11. The right-hand side shows how the superposition of the five morphological paradigms creates composite lexical families whose members belong to several distinct morphological paradigms.

6 Parasynthetic constructions

We have just seen how ParaDis analyzes a range of canonical and non-canonical constructions. Other phenomena may illustrate the contribution of paradigms even more clearly, especially the ones that include derivatives that cannot be analyzed with respect to their bases only because their formation involves several words in their derivational families or indirect relations. These phenomena will be called PARADIGMATIC PHENOMENA. They include the formation of English parasynthetic adjectives prefixed in *inter-* like *international*.

family	familial	interfamilial
nation	national	international
fistula	fistulous	interfistulous
filament	filamentous	interfilamentous
lexeme	lexemic	interlexemic
ocean	oceanic	interoceanic
module	modular	intermodular
capsule	capsular	intercapsular

Table 12: English adjectives prefixed in *inter-*. Column 2 contains the relational adjectives of the nouns in column 1. Column 3 contains adjectives denoting a relation between several entities referred to by the nouns in column 1.

Table 12 presents a sample of English adjectives prefixed in *inter-*, their base nouns and the relational adjectives of these nouns. The relational adjectives are suffixed in *-al*, *-ous*, *-ic* and *-ar*. The rivalry between these affixes does not seem to be driven by some semantic differentiation (Aronoff & Lindsay, 2016). Neither can we consider that they occupy morphophonological "niches." For example, *-ar* may have a preference for the nouns ending in /ul/ (*modular*) but these nouns can also be the base of adjectives in *-ous* (*fistulous*). In other words, the suffix of these relational adjective seems not to be predictable from the properties of the base noun.

The meaning of the adjectives prefixed in *inter-* in Table 12 can be paraphrased by 'which

connects several instances of *E*' where *E* stands for the entity denoted by the noun in the first column. For example, an *interoceanic channel* is a 'channel that connects several oceans'. The noun *ocean* is therefore the base of *interoceanic*. In addition to the prefix *inter-*, *interoceanic* has an suffix *-ic* which happens to be the exponent of the relational adjective *oceanic*. More generally, Table 12 shows that in a family the relational and the prefixed adjectives always have the same suffix. This allows the form of the prefixed adjective to be obtained by prefixing *inter-* to the form of the relational adjective. The formation of the prefixed adjectives is therefore a paradigmatic phenomenon which involves two members of their derivational families: their meaning is derived from that of the noun and their form from that of the relational adjective of that noun.

M71	M72	M73
family	familial	interfamilial
nation	national	international

M74	M75	M76
fistula	fistulous	interfistulous
filament	filamentous	interfilamentous

M77	M78	M79
lexeme	lexemic	interlexemic
ocean	oceanic	interoceanic

M80	M81	M82
module	modular	intermodular
capsule	capsular	intercapsular

M71	M72	M73
M74	M75	M76
M77	M78	M79
M80	M81	M82

Figure 13: Analysis of the paradigm of Table 15. The derivational paradigm of the adjectives prefixed in *inter-* results from the superposition (on the right side) of the four morphological paradigms listed on the left side.

6.1 Parasynthetic constructions in ParaDis

The analysis of these prefixed adjectives in ParaDis is straightforward. The eight families divide into the four homogeneous morphological paradigms listed on the left side of Figure 13. These paradigms are then superposed and yield a derivational paradigm as stated in the right side of the figure. The four morphological paradigms are in correspondence with a single trivial categorical paradigm which contains a single family composed of two categories: N and A. They are also in correspondence with a single semantic paradigm since the semantic relations within their families are all identical. This semantic paradigm SP4 = (S71,S72,S73) is presented in Table 13. The abstract semantic paradigm in the first line of the table shows that there is no direct connection between the meanings of the relational adjective and of the prefixed adjective. As a result, SP4 forms a connected but not complete graph (Figure 14). On the other hand, the formal paradigms in correspondence with the four morphological paradigms are all different. For instance, MP8 = (M71, M72, M73) is in correspondence with the formal paradigm FP8 = (F71, F72, F73) presented in Table 14. Unlike SP4 in Figure 14, the formal paradigm defines a complete graph (Figure 15) because the forms of the prefixed adjectives can be coined on the forms of the nouns and of the

relational adjectives. The other three morphological paradigms MP9 = (M74, M75, M76), MP10 = (M77, M78, M79) and MP11 = (M80, M81, M82) are in correspondence with similar formal paradigms that only differ from FP8 by the suffix.

S71	S72	S73
's_{71}'	'related to s_{71}'	'between several s_{71}'
'family'	'related to the family'	'between several families'
'nation'	'related to the nation'	'between several nations'
'fistula'	'related to the fistula'	'between several fistulae'
'filament'	'related to the filament'	'between several filaments'
'lexeme'	'related to the lexeme'	'between several lexemes'
'ocean'	'related to the ocean'	'between several oceans'
'module'	'related to the module'	'between several modules'
'capsule'	'related to the capsule'	'between several capsules'

Table 13: Semantic paradigm in correspondence with the morphological paradigms of the adjectives prefixed in *inter-*. The first line after the header describes the abstract semantic paradigm. The following lines form the concrete semantic paradigm.
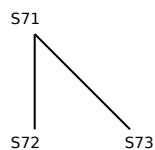

Figure 14: Graph defined by the abstract semantic paradigm SP4 = (S71, S72, S73) presented in Table 13.

Figures 14 and 15 show that the graphs defined by the abstract formal and semantic paradigms are not isomorphic which makes the data in Table 12 difficult to analyze in the morphological frameworks where form and meaning are not treated as independent levels of representation.

F71	F72	F73
$/f_{71}/$	$/f_{71}\text{-əl}/$	$/\text{mtɚ-}f_{71}\text{-əl}/$ $= /\text{mtɚ-}f_{72}/$
fæmɪli	fəmɪljəl	mtɚfəmɪljəl
neɪʃən	næʃənəl	mtɚnæʃənəl

Table 14: Formal paradigm FP8. The first line after the header presents the abstract formal paradigm and the next two lines the concrete formal paradigm.

The superposition of the four morphological paradigms brings out a formal regularity, namely the fact that the adjectives in the series M73, M76, M79 and M82 all begin in *inter-*. The superposition induces a superposition of the formal paradigms illustrated in Table 15 which shows that the four formal series describe the same prefixation in /mtɚ/.
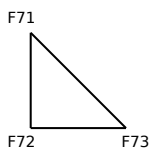
Figure 15: Graph defined by the abstract formal paradigm FP8 presented in Table 14.

(abstract) FP8	$/f_{71}/$	$/f_{72}/ = /f_{71}\text{-əl}/$	$/f_{73}/ = /\text{mtɚ-}f_{72}/$
(abstract) FP9	$/f_{74}/$	$/f_{75}/ = /f_{74}\text{-ɪk}/$	$/f_{76}/ = /\text{mtɚ-}f_{75}/$
(abstract) FP10	$/f_{77}/$	$/f_{78}/ = /f_{77}\text{-əs}/$	$/f_{79}/ = /\text{mtɚ-}f_{78}/$
(abstract) FP11	$/f_{80}/$	$/f_{81}/ = /f_{80}\text{-ər}/$	$/f_{82}/ = /\text{mtɚ-}f_{81}/$

Table 15: Superposition of the abstract formal paradigms of the adjectives prefixed in *inter-*. The cells in columns 2 and 3 are in the same relations in all the rows of the table.

6.2 Parasynthetic constructions in Bochner's (1993) Cumulative Patterns

The analyses of paradigmatic phenomena in ParaDis and with CPs are similar. For example, the families of the adjectives beginning in *inter-* and ending in *-al* can be represented by the cumulative pattern (28) which precisely corresponds to the morphological paradigm MP8. The analysis is based on two contributions of CPs: (*i*) they can describe derivational relations between any members of a derivational family; (*ii*) they connect all the members of a family at the same time. Patterns similar to (28) can be defined for the adjectives that have the three other endings (*-ous*, *-ic*, *-ar*). These four CPs can be generalized into a more abstract one (29) where the ending is represented by a variable SUFF. The only difference between (29) and the superposition of the formal paradigms in Table 15 is that (29) does not state explicitly that the form of the prefixed adjective is coined on that of the relational adjective.

$$(28) \quad \left\{ \begin{bmatrix} /X/ \\ N \\ \text{'Z'} \end{bmatrix}, \begin{bmatrix} /X\text{-al}/ \\ A \\ \text{'relative to Z'} \end{bmatrix}, \begin{bmatrix} /\text{inter-}X\text{-al}/ \\ A \\ \text{'between several Z'} \end{bmatrix} \right\}$$

$$(29) \quad \left\{ \begin{bmatrix} /X/ \\ N \\ \text{'Z'} \end{bmatrix}, \begin{bmatrix} /X\text{-SUFF}/ \\ A \\ \text{'relative to Z'} \end{bmatrix}, \begin{bmatrix} /\text{inter-}X\text{-SUFF}/ \\ A \\ \text{'between several Z'} \end{bmatrix} \right\}$$

This example shows that ParaDis and CPs basically have the same descriptive power with one difference: CPs being complete graphs of lexemes (where the correspondences are fixed once and for all), they cannot explicitly describe form-meaning discrepancies nor structural differences between non isomorphic formal and semantic abstract paradigms. In contrast, this is possible in ParaDis, thanks to the indirect correspondences between formal and semantic paradigms.

6.3 Parasynthetic constructions in Construction Morphology

Parasynthetic phenomena are analyzed in CxM by means of second order constructions and an inheritance hierarchy. The paradigm in Table 12 is described by (30) where the first construction represents the noun, the second its relational adjective and the third the adjective prefixed in

32

inter-. This construction is complemented by inheritance relations (31) and (32) which state that constructions (32a), (32b), (32c) and (32d) are subtypes of the more generic construction (31) where the exponent is represented by the variable *suff*. Inheritance relations account for the superposition of morphological paradigms. The analysis shows that inheritance and second order constructions give CxM the same expressive power as ParaDis.

(30) $< [x]_{Ni} \leftrightarrow [\text{SEM}]_i > \approx < [[x]_{Ni}\ \textit{suff}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j > \approx$
$< [\text{inter-} [[x]_{Ni}]_{Aj}]_{Ak} \leftrightarrow [\text{between several } [\text{SEM}]_i]_k >$

(31) $< [[x]_{Ni}\ \textit{suff}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j >$

(32)   a. $< [[x]_{Ni}\ \text{-al}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j >$
   b. $< [[x]_{Ni}\ \text{-ous}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j >$
   c. $< [[x]_{Ni}\ \text{-ic}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j >$
   d. $< [[x]_{Ni}\ \text{-ar}]_{Aj} \leftrightarrow [\text{relative to } [\text{SEM}]_i]_j >$

6.4   Parasynthetic constructions in RM

The analysis of the adjectives prefixed in *inter-* in RM is similar to their analysis in CxM. It is based on an inheritance hierarchy and sister schemas. More specifically, the three sister schemas in (33) describe the three series of the derivational paradigm of Figure 13. (33c) expresses the fact that the meaning of the prefixed adjective is defined as a function of the meaning of the noun (33a) and that its form is based on that of the relational adjective instantiating schema (33b) and identified by the variable index y. The exponent on the relational adjectives is represented by a variable REL identified by w. On the other hand, the four suffixation schemas in (34) inherit their properties from (33b);[5] in addition, each of them instantiates w to a constant index that identifies its suffix: 101 for *-al*, etc.

(33)   a.   b.   c.
Sem: $\text{ENTITY}_{\alpha,x}$ $[\text{PERTAINING-TO}(\text{ENTITY}_{\alpha,x})]_y$ $[\text{BETWEEN-SEVERAL}(\text{ENTITY}_{\alpha,x})]_z$
MSyn: N_x $[_A\ N_x\ \text{REL}_w]_y$ $[_A\ \text{INTER}_{83}\ [_A\ N_x\ \text{REL}_w]_y]_z$
Phon: $/\cdots/_x$ $/\cdots_y\ \cdots_w/_y$ $/\text{int}\textipa{@}_{83}\ \cdots_y/_z$

(34)   a. Semantics: $[\text{PERTAINING-TO}(\text{ENTITY}_x)]_y$
        Morphosyntax: $[_A\ N_x\ \text{REL}_{101}]_y$
        Phonology: $/\ldots_x\ \textipa{@}\text{l}_{101}/_y$
     b. Semantics: $[\text{PERTAINING-TO}(\text{ENTITY}_x)]_y$
        Morphosyntax: $[_A\ N_x\ \text{REL}_{102}]_y$
        Phonology: $/\ldots_x\ \textipa{@}\text{s}_{102}/_y$
     c. Semantics: $[\text{PERTAINING-TO}(\text{ENTITY}_x)]_y$
        Morphosyntax: $[_A\ N_x\ \text{REL}_{103}]_y$
        Phonology: $/\ldots_x\ \text{ɪk}_{103}/_y$
     d. Semantics: $[\text{PERTAINING-TO}(\text{ENTITY}_x)]_y$
        Morphosyntax: $[_A\ N_x\ \text{REL}_{104}]_y$
        Phonology: $/\ldots_x\ \textipa{@}\text{r}_{104}/_y$

---

[5]The scope of a variable index is limited to the schema or sister schemas where it occurs. For instance, the variable index y identifies the same lexical item in (33b) and (33c).

6.5 Parasynthetic constructions in TUHL

$$
\begin{bmatrix}
interoceanic \\
\text{PHON} \quad \langle /\text{int}\mathscr{x}/ \oplus \boxed{4} \rangle \\
\text{CAT} \quad a \\
\text{CONT} \quad \begin{bmatrix} btw\text{-}several\text{-}rel \\ \text{ARG} \boxed{2} \end{bmatrix} \\
\mu\text{-STRUCT} \begin{bmatrix}
\text{PHON-DGHTR} \begin{bmatrix}
oceanic \\
\text{PHON} \quad \boxed{4} \langle \boxed{1} \oplus /\text{ik}/ \rangle \\
\text{CAT} \quad a \\
\text{CONT} \begin{bmatrix} related\text{-}to \\ \text{ARG} \boxed{2} \end{bmatrix} \\
\mu\text{-STRUCT} \begin{bmatrix}
\text{PHON-DGHTR} \boxed{3} \begin{bmatrix} ocean \\ \text{PHON} \boxed{1}/\text{o}\upsilon\text{ʃən}/ \\ \text{CAT} \quad n \\ \text{CONT} \boxed{2} \text{ 'ocean'} \end{bmatrix} \\
\text{CONT-DGHTR} \boxed{3}
\end{bmatrix}
\end{bmatrix} \\
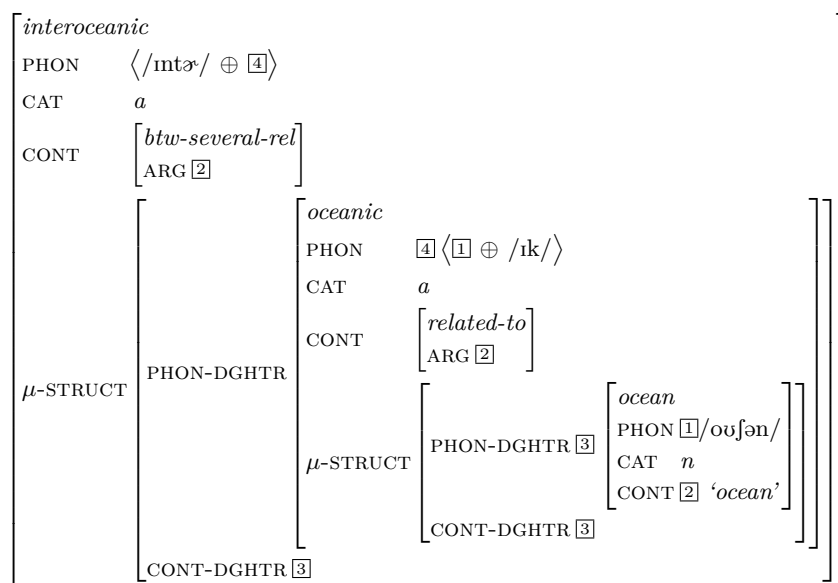\text{CONT-DGHTR} \boxed{3}
\end{bmatrix}
\end{bmatrix}
$$

Figure 16: AVM of *interoceanic*. Its formal base *oceanic* and its semantic base *ocean* are distinguished by the PHON-DGHTR and CONT-DGHTR features in μ-STRUCT.

The analysis of the adjectives prefixed in *inter-* in TUHL too is similar to their analysis in CxM. The issue is to account for the fact that an adjective like *interoceanic* has two "bases", a formal one, namely the relational adjective *oceanic*, and a semantic one, namely the noun *ocean*. In order to represent this mixed formation we divide the structural information given by μ-STRUCT into two features, PHON-DGHTR for the formal base and CONT-DGHTR for the semantic base.[6] By default, the two features have the same value. Figure 16 illustrates their use for the analysis of *interoceanic*. When described within its morphological family (Figure 17), the analysis of *interoceanic* becomes simpler because it directly refers to the representations of *ocean* and *oceanic*.

As in CxM, the description of the morphological paradigm of Table 12 is based on an inheritance hierarchy and an abstract paradigm. The hierarchy (Figure 18) contains types for the relational adjectives suffixed in *-al*, *-ous*, *-ic* and *-ar* and for the corresponding adjectives prefixed in *inter-*. These types inherit from the more abstract one *rel-adj-suff* (resp. *btw-adj-inter-suff*) where the exponent is unspecified. The second part of the description is an abstract paradigm (Figure 19) that generalizes the family in Figure 17.

---

[6]Technically, the mismatch between formal and semantic properties could be described without decomposing DGHTR, by coindexing the paths PHON and μ-STRUCT|DGHTR|PHON, on the one hand, and the paths CONT|ARGS and μ-STRUCT|DGHTR|CONT on the other. However, splitting DGHTR in two features makes the dissociation explicit and enables us to state that *interoceanic* is constructed from two distinct lexemes.

$$
\begin{bmatrix}
\textit{family-ocean} \\[4pt]
\text{MEMBERS}\ \Big\langle\ \boxed{3}\begin{bmatrix}\textit{ocean}\\ \text{PHON } \boxed{1}\ /\text{oʊʃən}/\\ \text{CAT } \quad n\\ \text{CONT } \boxed{2}\ \text{`ocean'}\end{bmatrix},\ \boxed{5}\begin{bmatrix}\textit{oceanic}\\ \text{PHON}\quad \boxed{4}\big\langle \boxed{1}\oplus/\text{ɪk}/\big\rangle\\ \text{CAT}\quad a\\ \text{CONT}\quad \begin{bmatrix}\textit{related-to}\\ \text{ARG } \boxed{2}\end{bmatrix}\\ \mu\text{-STRUCT}\begin{bmatrix}\text{PHON-DGHTR } \boxed{3}\\ \text{CONT-DGHTR } \boxed{3}\end{bmatrix}\end{bmatrix},\ \begin{bmatrix}\textit{interoceanic}\\ \text{PHON}\quad \big\langle/\text{ɪntɚ}/\oplus\boxed{4}\big\rangle\\ \text{CAT}\quad a\\ \text{CONT}\quad \begin{bmatrix}\textit{btw-several-rel}\\ \text{ARG } \boxed{2}\end{bmatrix}\\ \mu\text{-STRUCT}\begin{bmatrix}\text{PHON-DGHTR } \boxed{5}\\ \text{CONT-DGHTR } \boxed{3}\end{bmatrix}\end{bmatrix}\ \Big\rangle
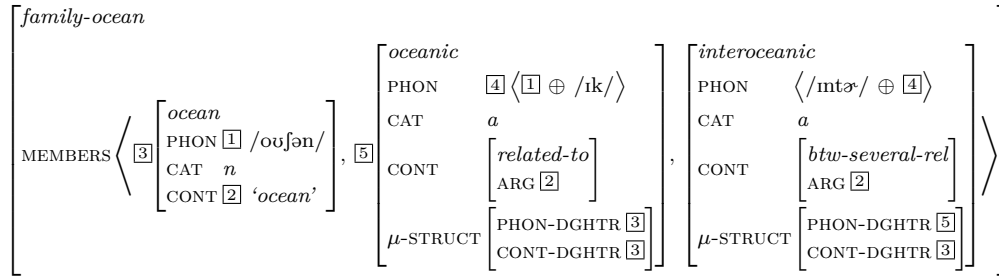\end{bmatrix}
$$

Figure 17: The family (*ocean*, *oceanic*, *interoceanic*).
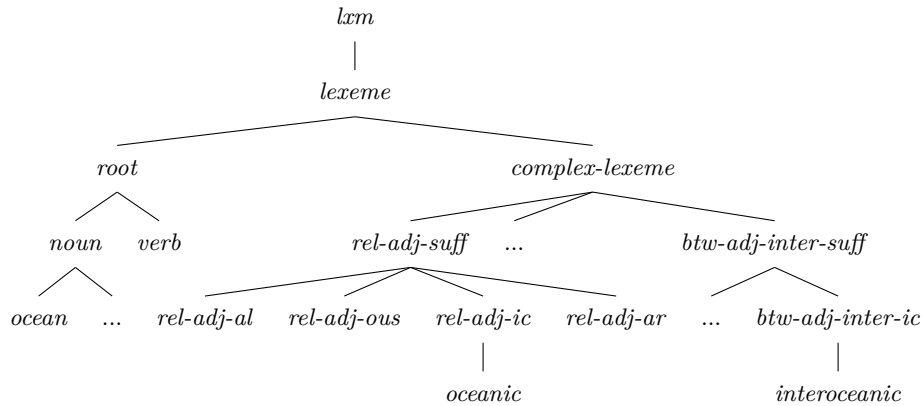
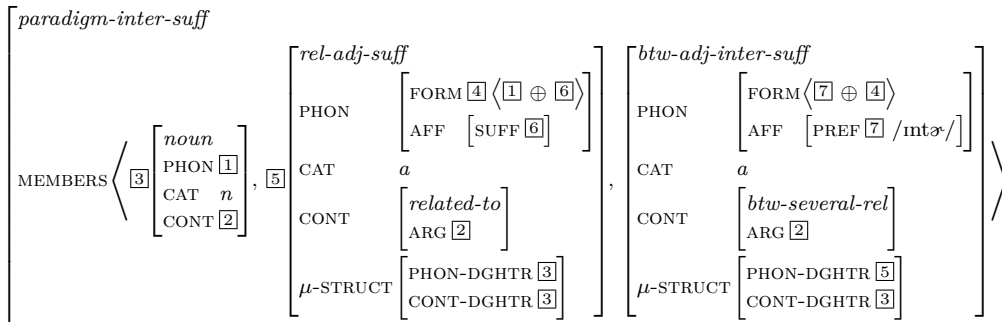Figure 18: Except of the hierarchy which contains the members of the morphological family of *ocean*.

$$
\begin{bmatrix}
\textit{paradigm-inter-suff} \\[4pt]
\text{MEMBERS}\ \Big\langle\ \boxed{3}\begin{bmatrix}\textit{noun}\\ \text{PHON } \boxed{1}\\ \text{CAT } \quad n\\ \text{CONT } \boxed{2}\end{bmatrix},\ \boxed{5}\begin{bmatrix}\textit{rel-adj-suff}\\ \text{PHON}\quad \begin{bmatrix}\text{FORM } \boxed{4}\big\langle\boxed{1}\oplus\boxed{6}\big\rangle\\ \text{AFF}\quad \begin{bmatrix}\text{SUFF } \boxed{6}\end{bmatrix}\end{bmatrix}\\ \text{CAT}\quad a\\ \text{CONT}\quad \begin{bmatrix}\textit{related-to}\\ \text{ARG } \boxed{2}\end{bmatrix}\\ \mu\text{-STRUCT}\begin{bmatrix}\text{PHON-DGHTR } \boxed{3}\\ \text{CONT-DGHTR } \boxed{3}\end{bmatrix}\end{bmatrix},\ \begin{bmatrix}\textit{btw-adj-inter-suff}\\ \text{PHON}\quad \begin{bmatrix}\text{FORM}\big\langle\boxed{7}\oplus\boxed{4}\big\rangle\\ \text{AFF}\quad \begin{bmatrix}\text{PREF } \boxed{7}\ /\text{ɪntɚ}/\end{bmatrix}\end{bmatrix}\\ \text{CAT}\quad a\\ \text{CONT}\quad \begin{bmatrix}\textit{btw-several-rel}\\ \text{ARG } \boxed{2}\end{bmatrix}\\ \mu\text{-STRUCT}\begin{bmatrix}\text{PHON-DGHTR } \boxed{5}\\ \text{CONT-DGHTR } \boxed{3}\end{bmatrix}\end{bmatrix}\ \Big\rangle
\end{bmatrix}
$$

Figure 19: Abstract paradigm of the adjectives prefixed in *inter-*.

7 Discussion

Overall, there is nothing ParaDis can do that at least one of the other models we presented cannot do. It differs from them in that it combines some of their strengths and their most useful features and improves the distribution of information and mechanisms. However, it does not allow for the

analysis of compounds unlike CxM, RM and TUHL.

Explicit separation of the levels of representation. Having explicitly separate levels of representation is one feature that distinguishes ParaDis from CP, Bonami & Strnadová's (2019) model, CxM and TUHL. It allows for a precise description of regularities that may be specific to one level of representation or that may concern several levels, including the morphological one. The separation also makes it possible to mutualize some of the families and paradigms which intervene in the description of these regularities (e.g., a semantic family may be in correspondence with several morphological ones).

Constraints. The separation of levels also allows the definition of constraints specific to a particular level of representation, i.e., which only apply to objects and relations in that level. Constraints can also be placed on the correspondences between the morphological level and the other ones. These constraints do not exist in other models even if CxM, RM and TUHL offer the possibility of imposing some of them through inheritance with however a notable difference: the constraints imposed through inheritance are rigid and mandatory, whereas in ParaDis they are gradable and can be violated. Their primary function is to preserve the regularities that exist in the lexicon. For example, it is by means of constraints that syncretism is maintained in the French morphological families of toponyms.

Paradigmatic families. Unlike other models, ParaDis distinguishes between the usual morphological families and the paradigmatic families which are subsets of morphological families that align to form morphological paradigms. This distinction improves the usual definition of word families, in particular the one proposed in (Bonami & Strnadová, 2019) which only indicates that derivational families can be partial, and the one of Bochner (1993), where the size of CSs is delimited by an evaluation metric.

Superposition of the morphological paradigms. The possibility offered by ParaDis to superpose homogeneous morphological paradigms determines a method of analysis that standardizes the description of the morphological phenomena. In particular, it helps linguists make explicit the variations they want to ignore. The superposition mechanism is also flexible (as it is hardly constrained compared to the solution adopted in the model of Bonami & Strnadová (2019) in which abstractions are "hard-coded") and intuitive (compared to models that use inheritance in which generalizations are represented by means of additional types). Superposition is not available in the other models where abstractions is accounted for by means of different mechanisms, e.g., inheritance in TUHL and CxM or cells containing sets of lexemes in (Bonami & Strnadová, 2019).

Parasynthetic constructions. By design, ParaDis has a homogeneous architecture that accounts for form-meaning discrepancies without any additional device. It shares this characteristic with Bochner's (1993) CPs with however a more explicit description of the contribution of the different levels of representation. The comparison of the analysis of parasynthetic constructions in ParaDis and in CxM, RM and TUHL highlights the need to generalize the second-order patterns in CxM and the sister relations in RM and to extend the type system and modify the structure of the complex lexemes in TUHL.

8   Conclusion

In this paper, we showed that the paradigmatic organization of derivational morphology can be captured in a simple, flexible and precise way. Our proposal is based on two notions: family and paradigm. Within this framework, we proposed ParaDis, a model whose main characteristics are: (*i*) A description of lexemes and morphological paradigms at four levels, all structured in the same way. These levels include the three classical ones: form, category and meaning. The fourth, called morphological level, plays the role of a trading floor where of the other three can interact. (*ii*) A description of the morphological regularities distributed on the four levels. This results in a dissociated representation that is both joint at the morphological level and totally independent at the other three. In this way, the analysis of a phenomenon adjusts closely to its level-specific regularities.

We also argued that ParaDis is suitable for both inflection and derivation and is therefore one more step towards a unified treatment of morphology. At the same time, it accounts for a large number of derivational processes known to be problematic in other theoretical models. It is able to catch the regularity of non canonical phenomena and to analyze a diversity of constructions in a uniform manner because it possesses the degrees of freedom needed to deal with most form-meaning discrepancies. In addition, the superposition of homogeneous morphological paradigms creates derivational paradigms that account for intuitive generalizations. We have also seen throughout the article that whenever a morphological phenomenon can be described by a table, its paradigmatic analysis at all levels of representation is, so to speak, mechanical and as simple as the design of canonical WFRs.

ParaDis sets up a framework for the study of many unresolved issues. The first is the characterization of the structure of paradigmatic families: how do they fit within the morphological families, and how are their boundaries delimited? Another issue is the actual integration of inflectional paradigms within the model. The answers to these questions will most likely be based on an extensive approach (Hathout et al., 2003), namely the construction of a large number of derivational paradigms in order to observe and identify the structural regularities they display. This is the purpose of the ongoing Demonext project (Namer et al., 2019; Hathout et al., 2020) whose objectives include the construction of a paradigmatic version of the Demonette database, a lexical resource designed for the description of word formation in French (Hathout & Namer, 2014b; Namer & Hathout, 2020; Hathout et al., 2020).

9   Acknowledgments

Declarations

References

Ackerman, Farrell, James P Blevins & Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In James P Blevins & Juliette Blevins (eds.), *Analogy in grammar: Form and acquisition*, 54–81. Oxford: Oxford University Press.

Anderson, Stephen R. 1982. Where's morphology. *Linguistic Inquiry* 13. 571–612.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge, UK: Cambridge University Press.

Antoniova, Vesna & Pavol Štekauer. 2015. Derivational paradigms within selected conceptual fields – contrastive research. *Facta Universitatis, Series: Linguistics and Literature* 13(2). 61–75.

Aronoff, Mark. 1994. *Morphology by itself. Stem and inflectional classes* (Linguistic Inquiry Monographs). Cambridge, MA: MIT Press.

Aronoff, Mark & Mark Lindsay. 2016. Competition and the lexicon. In Annibale Elia, Claudio Iacobino & Miriam Voghera (eds.), *"Livelli di analisi e fenomeni di interfaccia." Atti del xlvii congresso internazionale della società di linguistica italiana*, 39–52. Roma: Bulzoni.

Baerman, Matthew, Greville G. Corbett & Dustan Brown. 2010. *Defective paradigms: Missing forms and what they tell us*. Oxford: Oxford University Press.

Bauer, Laurie. 1997. Derivational paradigms. In *Yearbook of morphology 1996*, 243–256. Springer.

Beniamine, Sacha. 2018. *Classifications flexionnelles. Étude quantitative des structures de paradigmes*: Univeristé Paris Diderot Thèse de doctorat.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573.

Blevins, James P. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press.

Blevins, James P., Farrell Ackerman & Robert Malouf. 2016. Morphology as an adaptive discriminative system. In Daniel Siddiki & Heidi Harley (eds.), *Morphological metatheory*, 271–302. Amsterdam: John Benjamins.

Bochner, Harry. 1993. *Simplicity in generative morphology*. Berlin & New-York: Mouton de Gruyter.

Bonami, Olivier. 2014. *La structure fine des paradigmes de flexion*. Paris: Université Paris Diderot Habilitation à diriger des recherches.

Bonami, Olivier & Sacha Beniamine. 2016. Joint predictiveness in inflectional paradigms. *Word Structure* 9(2). 156–182.

Bonami, Olivier & Gilles Boyé. 2002. Suppletion and dependency in inflectional morphology. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 conference*. Stanford: CSLI Publications.

Bonami, Olivier, Gilles Boyé & Françoise Kerleroux. 2009. L'allomorphie radicale et la relation flexion-construction. In Bernard Fradin, Françoise Kerleroux & Marc Plénat (eds.), *Aperçus de morphologie du français*, 103–125. Saint-Denis: Presses universitaires de Vincennes.

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

Booij, Geert. 2010. *Construction morphology.* Oxford: Oxford University Press.

Booij, Geert. 2017. Inheritance and motivation in second orders schemata in construction morphology. In Nikolas Gisborne & Andrew Hippisley (eds.), *Defaults in morphological theory*, 18–39. Oxford: Oxford University Press.

Booij, Geert & Francesca Masini. 2015. The role of second order schemas in the construction of complex words. In Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds.), *Semantics of complex words*, vol. 47, 47–66. Heidelberg: Springer.

Boyé, Gilles. 2011. Régularité et classes flexionnelles dans la conjugaison du français. In *Des unités morphologiques au lexique*, 41–68. Paris: Hermès Science-Lavoisier.

Boyé, Gilles & Gauvain Schalchli. 2016. The status of paradigms. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*, chap. 9. Cambridge University Press.

Boyé, Gilles & Gauvain Schalchli. 2019. Realistic data and paradigms: The Paradigms Cell Finding Problem. *Morphology* 29(2). 199–248.

Carstairs, Andrew. 1987. *Allomorphy in inflexion.* London: Croom Helm.

Carstairs-McCarthy, Andrew. 1994. Inflection classes, gender, and the principle of contrast. *Language* 737–788.

Corbett, Greville G. 2003. Agreement: the range of the phenomenon and the principles of the surrey database of agreement. *Transactions of the philological society* 101. 155–202.

Corbett, Greville G. 2007. Canonical typology, suppletion and possible words. *Language* 83. 8–42.

Corbett, Greville G. 2010. Canonical derivational morphology. *Word Structure* 3(2). 141–155.

Corbin, Danielle. 2001. Préfixe et suffixes : du sens aux catégories. *Journal of French Language Studies* 11(1). 41–69.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information.* Chicago, IL: University of Chicago Press.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & Robert Harms (eds.), *Universals in linguistic theory*, 1–88. New-York: Holt, Rinehart, and Winston.

Fillmore, Charles J. 1976. Frame semantics and the nature of language. In *Annals of the new york academy of sciences: Conference on the origin and development of language and speech.* 20–32.

Fradin, Bernard. 2003. *Nouvelles approches en morphologie.* Paris: PUF.

Fradin, Bernard. 2020. Characterizing derivational paradigms. In Jesús Fernández-Domínguez, Alexandra Bagasheva & Cristina Lara-Clares (eds.), *Paradigmatic relations in word formation*, 49–84. Brill.

Fradin, Bernard. 2021. Caractériser les paradigmes dérivationnels. *Verbum* XLIII(1). 149–178.

Haspelmath, Martin. 1996. Word-class-changing inflection and morphological theory. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1995*, 43–66. Dordrecht: Kluwer Academic Publishers.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology, second edition.* London / New York: Routledge.

Hathout, Nabil. 2009. *Contributions à la description de la structure morphologique du lexique et à l'approche extensive en morphologie.* Toulouse: Université de Toulouse 2 - Le Mirail Habilitation à diriger des recherches.

Hathout, Nabil. 2011. Une approche topologique de la construction des mots : propositions théoriques et application à la préfixation en *anti-*. In *Des unités morphologiques au lexique*, 251–318. Paris: Hermès Science-Lavoisier.

Hathout, Nabil & Fiammetta Namer. 2014a. Discrepancy between form and meaning in word formation: the case of over- and under-marking in French. In Franz Rainer, Wolfgang U. Dressler, Francesco Gardani & Hans Christian Luschützky (eds.), *Morphology and meaning*, 177–190. Amsterdam: John Benjamins.

Hathout, Nabil & Fiammetta Namer. 2014b. Démonette, a French derivational morpho-semantic network. *Linguistic Issues in Language Technology* 11(5). 125–168.

Hathout, Nabil, Marc Plénat & Ludovic Tanguy. 2003. Enquête sur les dérivés en *-able*. *Cahiers de grammaire* 28. 49–90.

Hathout, Nabil, Franck Sajous, Basilio Calderone & Fiammetta Namer. 2020. Glawinette: a linguistically motivated derivational description of French acquired from GLAWI. In *Proceedings of the twelfth international conference on language resources and evaluation (LREC 2020)*. 3870–3878. Marseille.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51(3). 639–671.

Jackendoff, Raymond. 2002. *Foundations of language.* Oxford: Oxford University Press.

Jackendoff, Raymond. 2007. A parallel architecture perspective on language processing. *Brain Research* 1146. 2–22.

Jackendoff, Raymond. 2009. *Meaning and the lexicon: The parallel architecture 1975-2010.* Oxford: Oxford University Press.

Jackendoff, Raymond & Jenny Audring. 2018. Relational morphology in the parallel architecture. In Jenny Audring & Francesca Masini (eds.), *The oxford handbook of morphological theory*, 390–408. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199668984.013.33.

Jackendoff, Raymond & Jenny Audring. 2020a. Relational morphology: A cousin of construction grammar. *Frontiers in Psychology* 11-2241. 1–12.

Jackendoff, Raymond & Jenny Audring. 2020b. *The texture of the lexicon: Relational morphology and the parallel architecture.* Oxford: Oxford University Press.

Kleiber, Georges. 1984. Dénomination et relations dénominatives. *Langages* 76. 77–94.

Koenig, Jean-Pierre. 1999. *Lexical relations.* Stanford, CA: CSLI Publications.

Krieger, Hans-Ulrich & John Nerbonne. 1993. Feature-based inheritance networks for computational lexicons. In *Inheritance, defaults and the lexicon*, Cambridge: Cambridge University Press.

Körtvélyessy, Lívia, Alexandra Bagasheva & Pavol Štekauer (eds.). 2020. *Derivational networks across languages.* Walter de Gruyter.

Matthews, Peter Hugoe. 1972. *Inflectional morphology: A theoretical study based on aspects of latin verb conjugation.* Cambridge: Cambridge University Press.

McCarthy, John J. 1986. Ocp effets: Gemination and antigemination. *Linguistic Inquiry* 17(2). 207–263.

Molinier, Christian. 2018. Éthniques et gentilés. formes et propriétés respectives. *Le Français Moderne* 2018(2).

Namer, Fiammetta. 2021. Adjectifs dénominaux, interprétation multiple et indices morpho-sémantiques. *Verbum* XLIII(1). 119–147.

Namer, Fiammetta, Lucie Barque, Olivier Bonami, Pauline Haas, Nabil Hathout & Delphine Tribout. 2019. Demonette2 - une base de données dérivationnelles du français à grande échelle : premiers résultats. In *Actes de la 26ᵉ conférence annuelle sur le traitement automatique des langues naturelles (taln-2003)taln.* 233–243. Toulouse.

Namer, Fiammetta & Nabil Hathout. 2020. ParaDis and Démonette – from theory to resources for derivational paradigms. *The Prague Bulletin of Mathematical Linguistics* 114. 5–33.

Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase structure grammar.* Chicago, IL: The University of Chicago Press.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical report 2 Rutgers University Center for Cognitive Science New Brunswick, NJ.

Riehemann, Susanne Z. 1998. Type-based derivational morphology. *The Journal of Comparative Germanic Linguistics* 2(1). 49–77.

Roché, Michel. 2011. Pression lexicale et contraintes phonologiques dans la dérivation en *-aie* du français. *Linguistica* 51. 299–315.

Roché, Michel. 2017. Les familles dérivationnelles : comment ça marche ? Manuscrit.

Roché, Michel. 2006. Comment les adjectifs sont sémantiquement construits. *Cahiers de grammaire* 30. 373–387.

Roché, Michel. 2008. Structuration du lexique et principe d'économie : Le cas des ethniques. In Jacques Durand, Benoît Habert & Bernard Laks (eds.), *Actes du congrès mondial de linguistique française (cmlf-2008)*. 1571–1585. Paris: ILF.

Roché, Michel. 2009. Pour une morphologie *lexicale*. In *La morphologie lexicale est-elle possible ?*, vol. 17 (Mémoires de la Société de Linguistique, Nouvelle Série), 65–87. Leuven: Éditions Peeters.

Roché, Michel. 2011. Quel traitement unifié pour les dérivations en *-isme* et en *-iste* ? In Roché et al. (2011) 69–143.

Roché, Michel. 2017. Un exemple de réseau constructionnel : ethniques, toponymes, gentilés. Manuscript.

Roché, Michel, Gilles Boyé, Nabil Hathout, Stéphanie Lignon & Marc Plénat. 2011. *Des unités morphologiques au lexique.* Paris: Hermès Science-Lavoisier.

Sadock, Jerrold. 1991. *Autolexical syntax: A theory of parallel grammatical representations.* Chicago: University of Chicago press.

Sanacore, Daniele, Nabil Hathout & Fiammetta Namer. 2021. Frame-like structure for morphosemantic description. *Verbum* XLIII(1). 179–194.

Schalchli, Gauvain & Gilles Boyé. 2018. Paradigms and syncretism in derivation: the case of ethnics in French. *Lingue e linguaggio* 2018(2). 197–216.

Spencer, Andrew. 2013. *Lexical relatedness.* Oxford: Oxford University Press.

Štekauer, Pavol. 2015. The delimitation of derivation and inflection. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, vol. 1, 218–235. Berlin: Mouton de Gruyter.

Strnadová, Jana. 2014. *Les réseaux adjectivaux : sur la grammaire des adjectifs dénominaux en français*: Université Paris Diderot / Univerzita Karlova V Praze Thèse de doctorat.

Stump, Gregory & Raphael A Finkel. 2013. *Morphological typology: From word to paradigm*, vol. 138. Cambridge University Press.

Stump, Gregory T. 1991. A paradigm-based theory of morphosemantic mismatches. *Language* 675–725.

Stump, Gregory T. 2001. *Inflectional morphology. a theory of paradigm structure.* Cambridge, Royaume-Uni: Cambridge University Press.

Stump, Gregory T. 2005. Word-formation and inflectional morphology. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 49–72. Dordrecht: Springer.

Van Marle, Jaap. 1985. *On the paradigmatic dimension of morphological creativity.* Dordrecht: Foris.

Walther, Géraldine. 2013. *Sur la canonicité en morphologie : perspective empirique, formelle et computationnelle*: Univeristé Paris 7 Denis Diderot Thèse de doctorat.

Wunderlich, Dieter & Ray Fabri. 1995. Minimalist morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft* 14(2). 236–294.

Wurzel, Wolfgang Ulrich. 1989. *Inflectional morphology and naturalness*. Dordrecht: Reidel.

Štekauer, Pavol. 2014. Derivational paradigms. In Rochelle Lieber & Pavol Štekauer (eds.), *The oxford handbook of derivational morphology*, 354–369. Oxford: Oxford, Oxford University Press.