Derivational paradigmatic models put to test on some non-canonical phenomena

A growing body of work discusses the benefits of a paradigmatic description of derivational morphology (Bochner, 1993; Van Marle, 1985; Bauer, 1997; Štekauer, 2014; Hathout & Namer, 2018, 2019) and in particular in the analysis of several non-canonical constructions (Corbett, 2010) . The aim of this talk is to highlight the features that characterize and distinguish the paradigmatic models of derivational morphology by putting to test four of them on a variety of non-canonical phenomena including (*i*) form-meaning discrepancies, a minimum prerequisite they must meet; (*ii*) defectiveness, suppletion and *n*-uplets which are difficult to capture by WFRs; (*iii*) The models are also compared on their ability to account for the paradigmatic dimension of these phenomena including the explicit representation of derivational families, the distinction between abstract and concrete paradigms, and paradigm generalization.

1. Data.   The models will be compared on the French paradigm (1) proposed by Bonami & Strnadová (2019). The paradigm contains three families aligned semantically. The semantic contrasts between the lexemes included in the aligned cells of each family are identical: the first column contain location nouns, the second one nouns denoting people whose activity is related to the location, the third one, relational adjectives of the location nouns, and the fourth one, verbs that denote the action of (metaphorically) moving something in the location. On the formal level, the lexemes of the last three columns are formed by concurrent processes (*-ant/-ier*; *-al/-aire*; *-iser/en-/in-*). The paradigm also contains suppletive forms in the second and third families. Finally, the third family contains a doublet composed of *emprisonner* formed on the location noun *prison* and *incarcérer* formed on the same stem as the relational adjective *carcéral*.

(1)

commerce 'commerce'	*commerçant* 'shopkeeper'	*commercial* 'commercal'	*commercialiser* 'to commercialize'
école 'school'	*écolier* 'schoolboy'	*scolaire* 'educational'	*scolariser* 'to send to school'
prison 'prison'	*prisonnier* 'inmate'		*emprisonner* 'to imprison'
		carcéral 'of prison'	*incarcérer* 'to imprison'

2. Models.   in this talk, we compare the analysis of (1) in four models: Construction Morphology (CxM) of Booij (2010), the Cumulative Patterns (CP) of Bochner (1993) (B93), the Paradigmatic Systems (PS) of Bonami & Strnadová (2019) (BS19) and ParaDis of Namer & Hathout (2020). **CxM** is based on three devices: a multiple inheritance hierarchy; construction schemas which describe the internal structure of lexemes; second-order schemas that describe indirect relations between lexemes of the same complexity (Booij & Masini, 2015). In CxM, the description of morphological families is not central. It can be achieved by extended second-order schemas between fully instantiated constructions. **B93**'s model introduces two structures, the cumulative sets (CSs) which are sets of lexical items that belong to the same morphological family and the CPs which are sets of schemas that generalize the relations that hold in a collection of CSs or CPs. The model is said to be cumulative because any subset of a CS (resp. a CP) is itself a CS (resp. a CP). **BS19** define PSs as a set of semantically aligned families of the same size. The family members are sets of lexical items. Alignment of families consists in

superposing the elements which present the same contrasts of meaning with the other members of their families. In **ParaDis**, the paradigmatic representation is distributed over three levels of representation in order to enable a separate description of formal, categorical and semantic regularities that exist in the paradigm. The three levels of representation contain paradigms that are in correspondence with the paradigms of a fourth level, the morphological level. At all levels, concrete paradigms are superpositions of families. Abstract paradigms define graphs. The graphs defined by the abstract formal, categorical and semantic in correspondence with a morphological paradigm may be different in shape and size.

3. Discrepancies. In the first family in (1) *commercialiser* presents form-meaning discrepancy because its meaning directly depends on that of the noun *commerce* ('to put in commerce') and its form is formed by suffixing *-iser* to the form of the adjective *commercial*. **CxM** can account for this asymmetry by means of the formal and semantic indexes (2) in order to dissociate the relations of form and meaning within the paradigm. **B93** describes this partial family by means of the CS {*commerce, commercial, commercialiser*} and accounts for the mismatch in the same way as CxM. In **BS19**, all paradigms are considered to be complete graphs, semantically and formally. BS19's paradigms being morpho-semantic, the discrepancy is just ignored. In **ParaDis**, the morphological paradigm which contains *commercialiser* is in correspondence with a formal (abstract) paradigm and a semantic one. These paradigms define graphs of different shapes. The formal one is a complete graph but the semantic one is not because the meanings of *commercialiser* and *commercial* are not directly connected.

(2) $< [\text{kɔmers}]_{Ni} \leftrightarrow [\text{SEM}]_i > \approx \ < [[X]_{Ni} \text{-jal}]_{Aj} \leftrightarrow [\text{of } [\text{SEM}]_i]_j > \approx$
$< [[Y]_{Aj} \text{-iz}]_{Vk} \leftrightarrow [\text{to put in } [\text{SEM}]_i]_k >$

4. Defectiveness, suppletion, n-uplets. the families of *école* and *prison* are DEFECTIVE because they lack relational adjectives derived from the noun, unlike *commerce~commercial*. in **CxM**, **B93** and **ParaDis**, defective families are not distinguished from the other ones. on the other hand, defectiveness is explicitly represented by empty sets in **BS19**. Moreover, Bonami & Strnadová (2019) point out that the gaps are correlated with the presence in these families of relational adjectives formed on suppletive stems: *scolaire, carcéral*. In **CxM** and **B93**, SUPPLETIVE STEMS are described by means of independent variables (*X* and *Y-aire* standing for *école* and *scolaire*). In CxM, the second order schema $X \approx Y\text{-aire}$ generalizes the more specific schema $X \approx X\text{-aire}$ used for example to describe the *déficit~déficitaire* derivation. In **BS19**, suppletives form are variations which have no particular representation in the SPs. In **ParaDis**, the family of *école* (resp. *prison*) is a lexical family made up of two formally homogeneous morphological families: {*école, écolier*} and {*scolaire, scolariser*}. Both families are in correspondence with one and the same semantic family. In **CxM**, the DOUBLET *emprisonner/incarcérer* is represented by a second-order schema where the two constructions share the same semantic representation. Alternatively, one could use two partially redundant second-order schemas: *prison≈prisonnier≈carcéral≈emprisonner* and *prison≈prisonnier≈carcéral≈incarcérer*. **B93** may account for the doublet in the same way with either a CS of 5 lexemes or two partially redundant CSs with 4 lexemes each. In **BS19**, the doublet is described by a set of two lexemes which represents one member of the family. In **ParaDis**, doublet emerge from (*i*) the superposition of two formally homogeneous morphological paradigms: one contains [*prison, prisonnier, emprisonner*], and the other [*carcéral, incarcérer*], and (*ii*) the fact that the two lexemes are in correspondence with a same cell in the same semantic paradigm.

5. Paradigms and generalizations. Although second-order schemas are minimal abstract paradigms defined as associations of (two) construction schemas of the same complexity, strictly speaking, CxM does not contain paradigms. These can be described with generalized second order schemas between more than two constructions that may have different complexity. For example, paradigm (1) can be described as in (3) with a generic second-order schema which generalizes the constructions and relations of the family of *commerce* to the formal variations present in the three families of the paradigm. (3) is an abstract paradigm. CxM does not have a device to superpose families into concrete ones. On the other hand, CxM is redundant: (3) is complemented with more specific schemas between subsets of words of the families of (1). It therefore provides accurate representations of all the"local" relations along with a global description of the paradigm in which these are embedded. However, the local and global descriptions are not formally connected in the inheritance hierarchy because the schemas are of different sizes. A description of (1) in the same vein as B93's (4,5,6) is also possible in CxM.

(3) $< [X]_{Ni} \leftrightarrow [SEM]_i > \approx < [[X]_{Ni} \text{-suff1}]_{Nj} \leftrightarrow [\text{whose activity is related to } [SEM]_i]_j > \approx$
$< [[Y] \text{-suff2}]_{Ak} \leftrightarrow [\text{of } [SEM]_i]_k > \approx < [Z]_{Vl} \leftrightarrow [\text{to put in } [SEM]_i]_l >$

In **B93**, CPs describe abstract paradigms but the model cannot represent the superposition of CSs or CPs. Simplicity being the main objective of the B93, CPs stay close to the data. To this aim, each of the three families of paradigm (1) is described by a specific CP (4,5,6) which "locally" generalizes only one families.

(4)   [X,N,Z], [X-ã,N,'whose activity is related to Z'], [X-jal,A, 'of Z'], [X-jaliz,V,'to put in Z']

(5)   [X,N,Z], [X-je,N,'whose activity is related to Z'], [Y-ɛʁ,A,'of Z'], [Y-aʁiz,V,'to put in Z']

(6)   [X,N,Z], [X-je,N,'whose activity is related to Z'], [Y-al,A,'of Z'], [ã-X,V,'to put in Z'], [ɛ̃-Y,V,'to put in Z']

In **BS19**, the situation is reversed. SPs are concrete paradigms made up of aligned families but the model does not explicitlt include abstract paradigms. Families are aligned according to meaning contrasts only regardless of their formal variations (stem suppletion or affix competition). It is also possible to align families of different sizes by adding empty sets (resp. putting several lexemes in a single set) in order for them to fit into larger (resp. smaller) paradigms as in (7). This makes BS19 a very flexible model.

(7)   {commerce} {commerçant} {commercial} {commercialiser}
{école}   {écolier}   {scolaire}   {scolariser}
{prison}   {prisonnier} {carcéral}   {emprisonner, incarcérer}

ParaDis is more complete than the three previous models because formal, categorical and semantic regularities are described separately and then mapped into the morphological level. Each level of representation contains families and paradigms. Paradigms are superpositions of families with identical contrasts and are therefore totally homogeneous at the three levels of representation (formal, categorical and semantic). The morphological level contains two sorts of paradigms: (*i*) homogeneous morphological paradigms in correspondence with a single paradigm in each of the three levels of representation; (*ii*) derivational paradigms which are superpositions of morphological paradigms. The latter account for particular generalizations like the identity of the semantic contrasts in (1). The analysis of (1) involves five morphological paradigms highlighted in (8) with different colors. The five morphological paradigms are in correspondence with five distinct formal paradigm. On the other hand, all five are in correspondence with one categorical paradigm and one semantic paradigm which accounts for the identity of the meaning contrasts in the three families of (1). In sum, the analysis of (1) in ParaDis unfolds all the specific regularities it contains and then reconstructs the full paradigm by superposing the unfolded morphological paradigms.

(8)

commerce	commerçant	commercial	commercialiser
école	écolier		
		scolaire	scolariser
prison	prisonnier		emprisonner
		carcéral	incarcérer

6. Conclusion. All the models considered in this study account in a more or less precise way for the non canonical phenomena illustrated by (1). However, we have seen that (*i*) CxM does not provide explicit representations of paradigms; that (*ii*) only abstract paradigms can be described in B93 and that the paradigmatic structure is in large part determined by the formal variations; that (*iii*) in contrast, BS19 contains only concrete paradigms structured according to meaning contrasts and the formal variations are secondary; that (*iv*) ParaDis gives a precise account of the paradigmatic regularities by separating and articulating the description at four levels (formal, categorial, semantic and morphological). We have also seen that ParaDis is the only model that provides both concrete and abstract paradigms.

References

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

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

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 & 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.

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

Hathout, Nabil & Fiammetta Namer. 2018. Defining paradigms in word formation: concepts, data and experiments. *Lingue e Linguaggio* 17(2). 151–154.

Hathout, Nabil & Fiammetta Namer. 2019. Paradigms in word formation: what are we up to? *Morphology* 29(2). 153–165.

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

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

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