The MoNoPoli Database:
Extragrammatical and Subverted Processes in French Words
Based on Proper Names of Politicians

Mathilde Huguin

UMR 7118 ATILF, Université de Lorraine & CNRS – Nancy, France

Abstract

In this paper we present our method to build a derivational database of French deanthroponyms, which we call MoNoPoli for ***Mo**ts construits sur **No**ms propres de personnalités **Poli**tiques* ('complex words based on politician proper names'). MoNoPoli contains 6,545 complex words amounting to a total of 55,030 tokens and includes almost only neologistic forms. The Web is the only conceivable resource for collecting them: it alone gives massive access to discourse genres that contain neologisms. To feed the database, a program automatically generates the set of all possible derived words. Generated forms are then used as queries on the Web. Attested forms are kept with their context. This method provides a potential solution to collect data that cannot be found elsewhere. Finally, this article describes some of the remarkable results obtained with the analysis of the deanthroponyms of MoNoPoli. We show that the original nature of our data is reflected both in the use of new extragrammatical patterns (e.g., X*istan* secretion pattern) and in the subversion of grammatical processes (e.g., foreign suffixation X*ix*).

1. Introduction

We study French deanthroponyms, i.e., words morphologically built on proper names (Schweickard, 1992; Leroy, 2008) that refer to contemporary political figures, henceforth PPN 'politician proper name'. These data have the particularity of being absent from most existing French corpora since they are neologisms and often exhibit the properties of nonce-formations. According to Bauer (1983) and Dal and Namer (2018), nonce-formations are words deemed to be new by their creators and used

intentionally to meet an immediate need in a given context. In (1)[1] the forms *Macronite* ('Macron-itis'), *Balkanycratie* ('Balkany-cracy') and *hollandistan* ('holland-istan') are intentionally used by writers to express their aversion to the referents of the PPNs or their political ideas/actions.

(1)   a. *Un nouveau cas de* **Macronite** *aiguë était signalé en France.* (*Emmanuel Macron*)
          'A new case of acute Macron-itis was reported in France.'
       b. *La* **Balkanycratie** *est très éloignée de la démocratie !* (*Patrick Balkany*)
          'Balkany-cracy is very far from democracy!'
       c. *Bienvenue en* **hollandistan**, *pays en sous-développement.* (*François Hollande*)
          'Welcome to holland-istan, an underdeveloped country.'

As these complex words are most often neologistic nonce-formations, the constitution of the corpus required the elaboration of a specific methodology which we describe in this paper. This methodology consists of two steps. First, we automatically generate hypothetical deanthroponym *candidates* (e.g., *hollandien*, 'holland-ian', *lepenphobe*, 'lepen-phobic') using the 89 PPNs and 90 suffixes of Huguin (2018, 2021).[2] We then look up these candidates in context, using the Web.

We also present some of the results we obtained focusing on five types of atypical constructions which characterize deanthroponyms. To present these atypical constructions we use the typology of the *Natural Morphology* (Dressler and Barbaresi, 1994; Dressler, 2000; Kilani-Schoch and Dressler, 2005). Our reasoning for choosing Kilani-Schoch and Dressler (2005)'s typology is that it takes into account recreational (i.e., extragrammaticality, marginality) processes which are preponderant in the deanthroponymic lexicon. Indeed, even if the notion of nonce-formation does not overlap with that of extragrammaticality, we will see that extragrammatical deanthroponyms are typically nonce-formations. Deanthroponyms often deviate from prototypical morphology to meet playful needs (see also Zwicky and Pullum, 1987).

Our paper is structured as follows. We present the method used to build MoNoPoli in Section 2. We present some of the results that emerged from the analysis of our data in Section 3. The morphological processes identified in MoNoPoli are spread over the whole scale of *Natural Morphology*: from the most regular French processes (e.g., suffixation) to the most irregular ones (e.g., blending). We show that the original nature of our data is reflected both in the use of new extragrammatical patterns and in the subversion of grammatical processes. Finally, the Section 4 contains the summary of this paper.

---

[1]Each example of deanthroponym is provided with its context; the base PPN is indicated between brackets (*First Name Last Name*) and the stem/suffix boundary is marked by a hyphen in the English translation.

[2]These lists are available online at `https://apps.atilf.fr/homepages/mhuguin/these/documents/`.

2. Methodology

This section presents each of the steps taken to build MoNoPoli. First, we explain why the Web is the most methodologically adequate resource to contain deanthroponyms (Section 2.1). Then, we discuss two types of possible strategies to collect them and evaluate their theoretical implications (Section 2.2). Finally, we present the architecture of the program that allowed their generation (Section 2.3) and the results of our collection (Section 2.4).

2.1. Where to find data

In order to know where to look for deanthroponyms, one must ask about their characteristics, including their degree of institutionalization (Bauer, 2000; Hohenhaus, 2005; Lipka, 2007) or their discursive function.

The list of 89 PPNs used to generate candidates contains names of politicians who have held a prominent position in French politics (Presidents, Ministers, leaders of political party, etc.) since 1981 (e.g., Jacques Chirac, Emmanuel Macron). The choice of this date is motivated by the need to limit the data to be studied, on the one hand, and, on the other, the willingness to work on contemporary data. The aim is to study words built on recent PPNs, which have not been impacted by time and whose meaning has not become opaque (Bauer, 2000). As we have selected names of current personalities, we expect that the words based on PPNs are recent creations, i.e., neologisms (Štekauer, 2002; Kerremans, 2015). The list of PPNs is also intended to be representative of the French political landscape at the time the study was launched. The number of personalities selected from each of the French political parties is proportional to the number of seats the parties have in the French National Assembly. As in the National Assembly, women are under-represented in our list (27% women, 73% men). Finally, the choice to limit ourselves to anthroponyms whose referents are French is pragmatic. This facilitates the interpretation of the words built on these names, which can be based on the actions of the personalities.

Given that these politicians make decisions that directly impact the French people, one can assume that deanthroponyms formed on their names will occur in puns, jokes, criticisms or claims. Hence we can except that the complex words we are going to find will occupy argumentative or humorous functions. Therefore they display the characteristics of nonce-formations (Hohenhaus, 2005, 2015) as in (2). In (2a), alongside their morphological creations *royaliste*$_N$ ('follower of Ségolène Royal') and *montebourgeois*$_N$ ('follower of Arnaud Montebourg'), the writer inserts a meta-discursive comment in brackets: "I don't know if that's how you say it". The deanthroponym *hollandophobe*$_{Adj}$ ('hollando-phobic') (2b) appears in a sequence that contains several terms of the same series (X*phobe*$_{Adj}$), which Tanguy (2012, p. 104) calls suffixal outbursts. Comments and outbursts are among the structures that Dal and Namer (2018) have coined (meta)discursive and that often overlap with nonce-formations.

(2)   a. *« Perdre la raison », un blog militant. Longtemps* **royaliste***, maintenant* **mon-
           tebourgeois** (*je ne sais pas si ça se dit comme ça*). (*Ségolène Royal*, *Arnaud
           Montebourg*)
           "'Lose his mind", a militant blog. Long time royal-ist, now montebourg-ian
           (I don't know if that's how you say it).'
      b. *Il y a de quoi devenir phobe :* **hollandephobe**, **vallsphobe**, **taubiraphobe**,
           belkacemophobe*, gauchophobe, antifaphobe, imamophobe.* (*François Hollande*,
           Manuel Valls, *Christiane Taubira*, *Najat Vallaud-Belkacem*)
           'There is enough to come phobic: hollande-phobic, valls-phobic, taubira-
           phobic, belkacemo-phobic, lefto-phobic, antifa-phobic, imamo-phobic.'

Neologisms are more frequent in opinion genres than in information genres (Gérard,
2018). They indeed tend to be more frequently attested in less formal—or even
satirical—contexts. In order to maximize our chances of obtaining them, we should
look for resources where speakers/writers will be able to express themselves freely,
and where they will be able to reach a wide audience. Social networks, forums and
blogs, which are genres specific to the Web (Dal and Namer, 2015), provide such free-
dom and audience. To build up our corpus, we used the Web as a resource since it
alone provides access to these discursive genres in real time.

Plénat et al. (2002), Lüdeling et al. (2007), Fradin et al. (2008) and Dal and Namer
(2015) among others, have shown that the Web is useful for collecting contextualized
lexical scarcities. Since search engines are constantly performing new indexing, they
provide access not only to forms that have been recorded for a long time but also to
very recent coinages. To automatically and massively explore the content of the Web,
we used a Web scraping program[3] to query the Bing search engine. Our approach
can be described as hypothetico-deductive (Tanguy, 2012, p. 101): we first generated
a list of deanthroponym candidates and then searched for contexts containing the
members of this list on the Web.

2.2. Possible strategies

The hypothetical deanthroponyms used as queries are built from PPNs (3a) by
means of suffixes (3b). When generating candidates from such inputs, two strategies
can be adopted. Each strategy has its theoretical implications.

(3)   a. *Jacques Chirac*, *Christine Lagarde*, *Jean-Marie Le Pen*, *Emmanuel Macron*, *Jean-
           Luc Mélenchon*, *Nadine Morano*, *Nicolas Sarkozy*, *Christiane Taubira*, [...]
      b. *able, ade, ais, al, ard, erie, esque, eur, ien, ification, ine, iser, isme, issime, iste
           isterie, itude, ix, logue, mètre, oïa, ose, ou, phage, phile, phobe, thon, us* [...]

---

[3]The program we use is provided by the company Data Observer. Data Observer (`www.data-observer.
com`) is a start-up specialized in the collection, processing and analysis of textual data from the Web.

The first strategy, which we call *minimal* strategy, consists in generating only morphologically well-formed candidates. They respect a set of wellformedness morphophonological constraints (Roché and Plénat, 2014) as in *Optimality Theory* (Prince and Smolensky, 1993). For instance, this strategy leads to build exclusively *taubirie* /tobiʁi/ ('taubir-land') from the inputs *Taubira* /tobiʁa/ (from *Christiane Taubira*) and *-ie* /i/ so as to:

(i)   avoid the hiatus /ai/ (/tobiʁa/ + /i/ = */tobiʁai/) proscribed by the markedness constraint *ʜɪᴀᴛᴜs (Plénat and Roché, 2001),

(ii) tend towards the trisyllabic optimal, required according to the size constraints (Plénat, 2009).

The objective of this method is to model the repair strategies instinctively implemented by speakers—and assumed by the linguist—to obtain an *optimal* derivative. This first strategy therefore assumes that speakers/writers always (unconsciously) apply the phonotactic constraints and/or that we are only interested in well-formed deanthroponyms.

With the *maximal* strategy, all possible forms are generated, regardless of their adequacy to wellformedness principles. This strategy corresponds to the hypothesis that a speaker/writer may ignore morphophonological constraints of wellformedness, especially in a situation of spontaneous written expression. For example, the sequence /ʁaʁi/ from /tobiʁaʁi/, which corresponds to the attested form *Taubirarie* ('taubirar-land') from (4a) entails that the derivative violates the constraints of faithfulness, size as well as the *Obligatory Contour Principle* (Goldsmith, 1976). Faced with such attested examples, we opt for the *maximal* strategy. Moreover, the hierarchy of phonological constraints is not known. We regularly observe several output variants of a construction process, as the derivatives of (4) attest. The variants in the output of a morphological construction process are due to the idiosyncratic ordering of constraints as shown by Roché (2010).

(4)   a.   *Mais où sommes-nous ? En France ? Ou* **Taubirarie** *?* (*Christiane Taubira*)
           'But where are we? In France? Or Taubirar-land?'

      b.   *Vous vous foutez de qui en* **Taubirie** *?* (*Christiane Taubira*)
           'You do not care who in Taubir-land?'

      c.   *Il risque très peu en* **Taubirasie**... *no problemo.* (*Christiane Taubira*)
           'He risks very little in Taubiras-land ... no problemo.'

In sum, we choose to generate as many forms as possible using PPN stems or variants thereof and a list of suffixes. (5) is an excerpt from the set of graphical forms obtained from the PPN *François Bayrou* and the French suffix *-able*.

(5)   *françoisbayrouable, françoisbayroussable, françoisbayrousable, françoisbayroutable, françoisbayroulable, françoisbayroudable, françoisbayrouzable, françoisbayrounable, françoisbayrourable françoisbayrolable, bayrouable, bayroussable, bayrousable,*

> *bayroutable, bayroulable, bayroudable, bayrouzable, bayrounable, bayrourable, bayro-lable, françoisable, françoissable* [...]

In terms of costs and benefits, the maximal strategy produces much more noise than the minimal strategy. The higher the number of queries, the higher the noise. Nevertheless, the noise is an inconvenience that has a lesser impact than the dearth of results from the minimal strategy. Noisy results can be filtered out, whereas the lack of data cannot be compensated. In addition, this strategy allows us to collect unexpected formal creations, and, consequently, nonce-formations and extravagant formations that the minimal strategy does not allow us to obtain because it obeys morphological standards.

2.3. Generating derived forms

We run our candidate generation program on all the graphical forms that realize each of the 89 PPNs in our list and all 90 suffixes from our set (i.e., derivational suffixes or a verbal inflectional endings). PPNs are indeed realized in different forms, at least 3 (the *first name*, the *last name*, the *full name*), and up to 6, which we call sub-names and present in Table 1.

Sub-name	Example	Derived form	Gloss
Last name	Strauss-Kahn	*strausskahnité*	'strausskahn-ity'
First name	Dominique	*dominiqueur*	'dominiqu-er'
Full name	Dominique Strauss-Kahn	*dominiquestrauss--kahnien*	'dominiquestrauss kahn-ian'
Last name 1st part	Strauss	*straussophile*	'strausso-phile'
Last name 2nd part	Kahn	*kahnisation*	'kahn-ization'
Acronym	DSK	*dskie*	'dsk-land'

Table 1. Sub-names from the PPN Dominique Strauss-Kahn

The sub-names of a PPN are coreferential names that are used both autonomously in syntax and as bases in derivation. Unlike what happens with lexeme stems (Bonami et al., 2009), derivation rules do not impose constraints on sub-names, which is an additional argument for choosing the maximal strategy. We have shown (Huguin, 2018; Lignon et al., 2019) that the selection of a sub-name depends on sociolinguistic or extralinguistic conditionings such as the gender of the referent: e.g., the *firstname* is more often used if the referent is a woman (6).

(6)   a.   *Le Figaro, merci de défendre la langue française. On a déjà assez à faire pour lutter contre la* **najatisation** *de l'enseignement !* (*Najat Vallaud-Belkacem*)

'Le Figaro, thank you for defending the French language. We are already busy enough fighting against the najat-ization of education!'

b. *Le clientélisme et le* **clémentinisme** *se rejoignent.* (*Clémentine Autain*)

'Clientelism and clémentin-ism come together.'

The program inputs and outputs are sequences of characters. These graphical forms encode morphophonological phenomena as well as purely orthographic variations. The program generates all possible tuples formed by a *stem* of sub-name and a *suffix*. For each tuple, the outputs of the program correspond to a set of possible derived words that we name *Candidates*. Each *Candidate* is obtained by concatenating ($\oplus$) the form of a sub-name $\textit{Stem}_i^{n\,\prime}$ and a suffix $\textit{Suff}_j$ (7a) ($0 < j \leq 90$). For a given PPN_i ($0 < i \leq 89$), the symbol $\textit{Stem}_i^{n\,\prime}$ corresponds to the *stem* of one of its sub-names n ($0 < n \leq 6$), or consists of a variation of this *stem* (7b) produced by one of the 36 $\mathfrak{R}$ rules of the program.

(7) a. $\textit{Candidate} = \textit{Stem}_i^{n\,\prime} \oplus \textit{Suff}_j$

  b. $\begin{cases} \textit{Stem}_i^{n\,\prime} = \textit{Stem}_i^{n} \\ \textit{Stem}_i^{n\,\prime} = \mathfrak{R}(\textit{Stem}_i^{n\,\prime}) \end{cases}$

Each of them selects two arguments: $\textit{Stem}_i^{n\,\prime}$ and $\textit{Suff}_j$. Rules are organized in four blocks, cf. Figure 1. The rules in the block $\mathcal{B}1$ remove a graphical sequence from $\textit{Stem}_i^{n\,\prime}$. The rules of $\mathcal{B}2$ add a graphical sequence to $\textit{Stem}_i^{n\,\prime}$. The rules of $\mathcal{B}3$ operate graphical substitutions. When relevant, the rules are the graphical transcriptions of morphophonological rules: truncation for $\mathcal{B}1$, epenthesis for $\mathcal{B}2$ and allomorphy for $\mathcal{B}3$. Finally, the $\mathfrak{R}36$ rule of $\mathcal{B}4$ concatenates the inputs $\textit{Stem}_i^{n\,\prime}$ and $\textit{Suff}_j$.
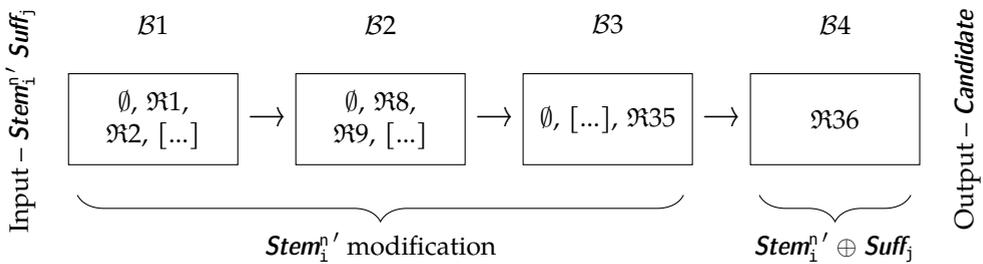

Figure 1. Rule combinations

The program is oriented and acyclic. In each block, the rules are in complementary distribution. The values of $\textit{Stem}_i^{n\,\prime}$ et $\textit{Suff}_j$ constrain which blocks ans which rules can be activated. This organization leads to 136 possible rule combinations. Each *stem*/*suffix* input explores the 136 combinations, but a *Candidate* is only produced when

the conditions of application of all the rules of the combination are met. Otherwise, the program tries the next combination.

Let us take the example of the input *taubira*/*ique*.

- If we follow the first possible combination, we apply the null rule ($\emptyset$) in each of the blocks $B1$, $B2$ and $B3$. Then rule $\mathfrak{R}36$ of concatenation in $B4$ gives the **Candidate** (8).
- In contrast, the second combination using $\mathfrak{R}1$ in $B1$ will be discarded by the program. $\mathfrak{R}1$ corresponds to deleting the final *e* of a **stem**, hence it cannot be applied to *taubira*. The next rule in $B1$, i.e., $\mathfrak{R}2$, can be applied, since the input *taubira*/*ique* satisfies the conditions for application of the $\mathfrak{R}2$ truncation rule: *taubira* ends with a vowel and *ique* begins with a vowel. $\mathfrak{R}2$ deletes the vowel *a* at the end of **stem**, to produce *taubir* (9a). Then, the null rules apply in $B2$ and $B3$. The output of $\mathfrak{R}2$ is given as input to $\mathfrak{R}36$ in $B4$ to generate the **Candidate** (9b).
- Testing all rule combinations exhaustively will eventually produce all other candidates: e.g., with epenthesis (10) and (11).

(8) $\mathfrak{R}36$(*taubira, ique*) = *taubiraique*
(9)   a. $\mathfrak{R}2$(*taubira, ique*) = (*taubir, ique*)
     b. $\mathfrak{R}36$(*taubir, ique*) = *taubirique*
(10) $\mathfrak{R}8$(*taubira, ique*) = (*taubirat, ique*)
    $\mathfrak{R}36$(*taubirat, ique*) = *taubiratique*
(11) $\mathfrak{R}9$(*taubira, ique*) = (*taubiras, ique*)
    $\mathfrak{R}36$(*taubiras, ique*) = *taubirasique*

2.4. Data collection and annotation

The program produces 110,658 candidate forms, and each is used as a query, i.e., submitted to Bing. The set of attested deanthroponyms, their contexts, the URLs, and the number of pages associated with each query are saved in a tabulated file. A manual post-processing is then applied. It consists, for example, in deleting the entries of candidates homographs to attested lexemes with another meaning. For instance, *hollandais* ('holland-ese') is derived from *François Hollande* (12) but more often refers to the inhabitants of Holland.

(12) *Dans le cercle des* **hollandais**, *certains émettent l'hypothèse d'une absence du président sortant dans la course présidentielle.* (*François Hollande*)
    'In the circle of holland-ese, some speculate that the outgoing president will not be in the presidential race.'

The database we obtain contains 6,545 different deanthroponyms, for a total of 55,030 tokens. This corpus contains 3,830 complex words whose formation mode were expected as they were explicitly generated by the program. But Bing's indexing process accidentally brought back a significant amount of unexpected forms: 40% of the deanthroponyms harvested are not part of our candidate list. For instance, we

obtained occurrences of the prefixed noun *anti alliot-marisme*$_N$ ('anti-alliot-marism')
(13a) and of the compound adjective *chiraco-raffarinesque*$_{Adj}$ ('chiraco-raffarinian') (13b),
looking for attestations of the candidates **marisme** and **raffarinesque**.

(13)   a. *Est-ce qu'une vague d'***anti alliot-marisme** *peut déferler sur la circonscription*
            [...] ? (Michèle Alliot-Marie)
            'Can a wave of anti-alliot-marism sweep through the riding [...]?'
        b. *La majorité parlementaire* **chiraco-raffarinesque** *supprime une bonne partie des*
            moyens financiers [...]. (Jacques Chirac, Jean-Pierre Raffarin)
            'The chiraco-raffarinian parliamentary majority suppresses a good part of
            the financial means [...].'

Each entry in the database describes an occurrence of one of the 6,545 deanthro-
ponyms collected. This description is decomposed into a hundred or so features
which describe, for each deanthroponym: its morphological properties such as its
pattern (*Xade* for *peillonnade* in Table 2), its category and its morphophonological
characteristics. Other properties do not result directly from the observation of the
deanthroponyms but from our own analysis and are absent from Table 2. These in-
clude, for example, the semantic category of derivatives. The reader will be able to
find the database and an explanation of each feature online via the Open Resources
and TOols for LANGuage (ORTOLANG) platform.[4]

PPN	Derivative in context	Pattern	POS	Phonology
Vincent Peillon	*une nouvelle* **peillonnade** *: la ren-trée en août !* 'a new peillon-ade: back to school in August!'	*Xade*	Nc	/pejonad/
Rama Yade	*Matignon redoute la* **ramayadisa-tion** *de Vallaud-Belkacem* 'Matignon fears the ramayad-ization of Vallaud-Belkacem'	*Xisation*	Nc	/ʁamajadizasjɔ̃/
François Hollande	*le chimpanzé à cul rose* **homo hol-landus** 'the pink ass chimpanzee homo holland-us'	*homoXus*	Nc	/omoolɑ̃dys/

Table 2. Exerpt of MoNoPoli

---

[4]The database is available at this address: https://www.ortolang.fr/market/corpora/monopoli. Mo-
NoPoʟɪ is distributed under CC-by 4.0 license.

3. Analysis of deanthroponyms: some remarkable results

Unsurprisingly, the PPNs most often used as bases in MoNoPoli (14) correspond to the most prominent public figures. They have held a more important position (President *vs* Member of Parliament) or have been involved in high-profile events (laws, judicial scandals). The 13 PPNs in (14) are the bases of 50% of the deanthroponyms in our corpus. Jacques Chirac (15) was President of the French Republic. Dominique Strauss-Kahn (16) was implicated in scandals (sexual assault and rape). The PPNs of these referents are typically used to create nonce-formations since the referents are subject to controversy.

(14)   *Dominique Strauss-Kahn, Marine Le Pen, Emmanuel Macron, Manuel Valls, Jean-Luc Mélenchon, François Mitterrand, Christiane Taubira, Ségolène Royal, François Bayrou, Lionel Jospin, François Hollande, Jacques Chirac, Nicolas Sarkozy*

(15)   a. *On dit "arrête de* **chiraquer**" *pour dire arrête de faire des bêtises.* (*Jacques Chirac*)
           'We say "stop chiraqu-ing" to say stop doing stupid things.'
       b. *Heureux français, pensez à votre chance : vous n'avez rien à craindre du* **sarkoz-ium**, *du* **ségolénium**, *du* **chiraquium**, *mais méfiez-vous quand même du* **lepénium**. (*Nicolas Sarkozy*, *Ségolène Royal*, *Jacques Chirac*, *Jean-Marie Le Pen*)
           'Happy French people, think about your luck: you have nothing to fear from sarkoz-ium, segolen-ium, chiraqu-ium, but beware of lepen-ium.'

(16)   a. *Enfumages sans feux : après l'éruption mentale de* **viol-kahnisme** *sulfureux présumé, retour au volcanisme réel.* (*Dominique Strauss-Kahn*)
           'Smoke and mirrors without fire: after the mental eruption of presumed sulphurous rape-kahnism, back to real volcanism.'
       b. *Les socialos vont lancer le* **dskthon** *car faut aider cet homme qui aurait voulu diriger la France dans la plus grande hypocrisie.* (*Dominique Strauss-Kahn*)
           'The socialos are going to launch the dsk-thon because it is necessary to help this man who would have wanted to lead France in the biggest hypocrisy.'

The nonce-formations are identifiable thanks to the meta-discursive signals and the discursive processes (see Section 2.1) but also sometimes thanks to the morphological processes used. 10% of the deanthroponyms of MoNoPoli are produced by extragrammatical processes (Kilani-Schoch and Dressler, 2005; Fradin et al., 2009). As a reminder, according to Fradin et al. (2009) a process belongs to the extragrammatical domain if:

(17)   a. it is conscious;
       b. it is located at the bottom of a scale of typological prototypicity of the processes (see (18) below);
       c. it is not productive (see Dal and Namer, 2016);
       d. it interacts with the different modules of grammar (e.g., syntax).

As Fradin (2003) points out, the reason these criteria are not precise is that extragrammatical morphology is defined in the negative: it is broadly what grammatical

morphology is not. Kilani-Schoch and Dressler (2005, p. 94) give seven criteria that
define grammatical morphology:

(18)   a. the rules of grammatical morphology manipulate form and meaning si-
           multaneously;
        b. they are applied in a regular and predictable way;
        c. a morphological rule applies to a basic distinct class of bases (notably dis-
           tinct by its part of speech);
        d. the meaning change in rule operations is additional;
        e. morphological rules produce new lexemes, distinct from their base;
        f. the grammatical morphology is the object of an unconscious knowledge;
        g. the grammatical morphology is productive.

We provide an overview of the morphological diversity of the content of MoNoP-
oli in Table 3.

Process & Frequency		Example & Gloss
Grammatical processes	90%	*hollandifier*$_V$ 'holland-ify'
Derivation	58%	*jospinerie*$_N$ 'jospin-ery'
Suffixation	51%	*chiraquiste*$_N$ 'chiraqu-ist'
Prefixation	5%	*ex-bovétiste*$_{Adj}$ 'ex-bovét-ist'
Conversion	2%	*bayrouer*$_V$ 'to-bayrou'
Compounding	32%	*lepénisme-mégretisme*$_N$ 'lepenism-megretism'
Extragrammatical processes	10%	*aubrython*$_N$ 'aubry-thon'

Table 3. Overview of the processes of MoNoPoli
according to Kilani-Schoch and Dressler (2005)'s typology

The Table 3 leads to some comments concerning the methodology and the pro-
cesses attested in MoNoPoli.
- The most frequent types of constructs are formed by suffixation which is ex-
  pected since candidates are generated thanks to 90 suffixes (cf. Section 1 and
  Section 2.3).
- For the same reason, the compounds systematically contain at least one suffixed
  deanthroponym and MoNoPoli does not contain any deanthroponym with only
  a prefix, i.e., without the simultaneous presence of a derivational suffix (e.g,
  -isme in *anti-aubrisme* 'anti-aubr-ism') or a verbal inflectional ending (e.g., *-er* in
  re-macroner 'to re-macron').

- Derivation by conversion is the least represented,[5] which is certainly an effect of our methodology. Indeed, we only looked for forms corresponding to infinitive verbs (no noun or adjective).
- The rather small proportion of extragrammatical formations is also the result of the methodology used to form the deanthroponym candidates. Indeed, non-concatenative processes have little chance of being collected with our strategy.

In the following, we show that PPNs are privileged bases for original constructions on several levels. On the one hand, they serve as a bases for processes classified as extragrammatical in French (Section 3.1). On the other hand, they also sometimes lead to the subversion of more classical, i.e., more regular processes (Section 3.2).

3.1. Extragrammatical processes

MoNoPoli contains 662 deanthroponyms constructed by extragrammatical processes (see Table 3). More than half (54%) of the extragrammatical forms are hapaxic. This number confirms indirectly that they are original creations. The extragrammatical formations are indeed guided by a will to stand out from the crowd. These are often constructions whose *form* is atypical and rare, which indicates the occasional character of the deanthroponym. For instance, (19) contains a hapax legomenon: *taubirier*$_V$. It is a blend (Bat-El, 1996; Fradin, 2000; Fradin et al., 2009) built on the bases *Christiane Taubira* and *marier* ('to marry'). It means 'get married thanks to Christiane Taubira' (Christiane Taubira is the Minister of Justice who legalized gay marriage in France). This occasional character is often confirmed by the analysis of the context which presents a remarkable discursive structure (e.g., rhymes, chiasms or meta-discursive comments, see Section 2.1). For example, the two contexts of the blends in (20) are built on the model of a dictionary definition with a generic concept (*genus proximum*) and its specific features (*differentia specifica*).

(19)   *Il a eu son heure de gloire en organisant un faux mariage d'homos il y a quelques années.*
         Maintenant qu'Hollande et Taubira ont fait passer leur loi en force il peut en **Taubirier**
         à la pelle chaque semaine. (*Christiane Taubira*)
         'He had his moment of glory by organizing a fake gay wedding a few years ago. Now that Hollande and Taubira have rammed through their law, he can Taubirier a lot every week.'

(20)   a. **Moranogastrique***: (n.c. masc) Dispositif médical visant à couper l'appétit. Dis-*
             positif lourd, le moranogastrique est proposé en dernier recours et ne saurait être
             utilisé comme simple coupe faim. Parmi ses nombreux effets secondaires, on note
             qu'il a souvent tendance à provoquer nausée et diarrhée. (*Nadine Morano*)
             'Moranogastric: (n.c. masc) Medical device designed to suppress appetite.
             A heavy device, the moranogastric is proposed as a last resort and should

---

[5]The database contains 128 converted verbs (1,045 tokens). 55 PPNs out of the 89 bases used give a converted verb.

not be used as a simple appetite suppressant. Among its many side effects, it often tends to cause nausea and diarrhea.'

   b. **Duflotaïne***: Psychotrope hallucinogène entraînant une propension à renier ses convictions et des troubles du comportement. (Cécile Duflot)*
   'Duflotaine: Hallucinogenic psychotropic drug leading to a propensity to deny one's convictions and behavioral disorders.'

In the following, we focus on two extragrammatical patterns present in MoNoPoli which have not been described in French. Section 3.1.1 is devoted to the pattern X*oX*suff exemplified in (21a) and Section 3.1.2 to the pattern X*istan* exemplified in (21b).

(21)   a. *Imaginez que Lionel Jospin ait au second tour le projet de fondre toute la gauche plurielle dans un seul mouvement* **jospino-jospinien***. Si tel était son intention, je peux vous assurer d'entrée de jeu qu'il a perdu les élections. (Lionel Jospin)*
   'Imagine that Lionel Jospin had a plan in the second round to merge the entire plural left into a single jospino-jospinian movement. If this was his intention, I can assure you from the outset that he lost the election.'

   b. Nous regrettons aussi qu'il n'y ait pas eu de rubriques sur le **Mitterrandistan** ou le **Jospinistan**. (*François Mitterrand*, *Lionel Jospin*)
   ' We also regret that there were no sections on Mitterrand-istan or Jospin-istan.'

3.1.1. Reduplication

MoNoPoli contains 23 deanthroponyms that instantiate the pattern X*oX*suff as complex words of (22) where X is the stem of a sub-name of PPN. In (22a) *suff* is *-iste* ('-ist'), in (22b) and (21a) it is *-ien* ('-ian').

(22)   a. *Il s'explique dans un entretien à paraître ce mardi dans les éditions mayennaises d'Ouest-France : sur le plan départemental, c'est la motion la droite forte qui est arrivée largement en tête. Les militants ont choisi la ligne dure,* **sarkozo-sarkozyste***. (Nicolas Sarkozy)*
   'He explains himself in an interview to appear this Tuesday in the Mayenne editions of Ouest-France: on the departmental level, it is the motion of the hard right which arrived largely in head. The militants chose the hard line, sarkozo-sarkozist.'

   b. *En témoigne le remaniement ministériel, qui fait la part plus que belle aux* **chiraco-chiraquiens** *: il s'apparente à la formation de la tortue, notent les bons observateurs de la vie politique. (Jacques Chirac)*
   'This is evidenced by the ministerial reshuffle, which gives the lion's share to the chiraco-chiraquians: it is similar to the formation of the turtle, remark astute observers of the political sphere.'

At first sight, the process used to obtain these derivatives is compounding. One could see in *sarkozo-sarkozyste*$_\text{Adj}$ and *chiraco-chiraquien*$_\text{N}$ a particular case of the compound X*oYsuff*. For example, the demonym *franco-canadien*$_\text{Adj/Nc}$ ('French and Canadian') is compounded from the bases X *français*$_\text{Adj/Nc}$, which appears truncated, et Y *canadien*$_\text{Adj/Nc}$ (Dal and Amiot, 2008). As in *franco-canadien*$_\text{Adj/Nc}$, we find in *sarkozo-sarkyste* the intercalary vowel /o/ typically associated with compounds, and more specifically *learned* compounds, i.e., including the stem of a lexeme inherited from Greek or Latin. In *franco-canadien*$_\text{Adj/Nc}$, the vowel can be thought of as subverted, in that it does not mark the learned character of a stem but constitutes an iconic marker of compounding (Dal and Amiot, 2008).

Contrary to the compounds which are constituted of two distinct bases, in the deanthroponyms of (22), it is always the same (truncated) stem of X which is used in each element. This observation constitutes a first obstacle, of a formal nature, to the analysis of *sarkozo-sarkozyste*$_\text{Adj}$ and *chiraco-chiraquien*$_\text{N}$ as compounds. The second obstacle is semantic. The meanings of deanthroponyms of (22) do not correspond to coordination, subordination or apposition if we follow the tripartite classification of compounds of Scalise et al. (2005), for example. Semantically, *sarkozo-sarkozyste* (22a) means 'very/typically/exclusively sarkozist' and *chiraco-chiraquiens* (22b) are 'very/typically/exclusively chiraquian people'. This meaning consists of an amplification or an exaggeration of the property denoted by the base, and this semantic value is attested in cases of reduplication. The deanthroponyms of (22) are therefore not compounds but exhibit characteristics of reduplication.

This intensifying and restrictive polarization of reduplication is notably attested in syntax. In some syntactic reduplications, called *Identical Constituent Compounding* (Hohenhaus, 2004) or *Contrastive Reduplication* (Ghomeshi et al., 2004), the reduplicant allows a meta-discursive comment on the other term. In (23) the reduplication *mad mad* means 'very/completely mad'. When these are reduplicated nouns the associated paraphrase may be 'exclusively NOUN': e.g., *I want salad salad* means 'only/exclusively salad' (not *tuna salad* or *mixed salad*).

(23) She's mad […] Not **mad mad**, but, you know. Out of control.
        (Hohenhaus, 2007, p. 26)

These paraphrases have semantic values equivalent to those we have determined for the complex words of (22). In syntax, as in morphology, it is a matter of either intensifying a property or restricting the designated referential class to referents which possess prototypical properties of the class (Kleiber, 1990). Intensification and restriction are well known semantic values for reduplication (Moravcsik, 1978) and attested in many languages (e.g., English, Italian, French, Turkish). In French, reduplication is used exclusively for evaluative purposes. We therefore analyze the deanthroponyms X*oXsuff* as resulting from reduplication; more specifically, this is partial pre-reduplication since the reduplicant X*o* is on the left and does not use the entire

phonological material of the base. Finally, we should add that the identified process is not specific to anthroponymic bases since it also applies to ethnic adjectives such as *français*$_{Adj}$ which gives *franco-français*$_{Adj}$ 'very/typically/exclusively French' in (24).

(24) *« Nuit debout » n'est plus un mouvement* **franco-français**.
      "'Nuit debout" is no longer a Franco-French movement.'

3.1.2. Secretion

MoNoPoli contains 15 deanthroponyms that instantiate the pattern *Xistan* as in (25). All of them are proper names of places where PPN referents hold power.

(25)   a. *Pourquoi ne pas coller le nom de chacun des 12 départements pendant qu'ils y sont ?* **Wauquiezistan** *ça sonnait mieux non ?* (*Laurent Wauquiez*)
            'Why not stick the name of each of the 12 departments while they are at it? Wauquiez-istan sounded better?'
        b. *Le petit train-train de la honte dans la catégorie humeurs et gueule de bois en* **sarkozistan**. (*Nicolas Sarkozy*)
            'The little train of shame in the category moods and hangovers in sarkozistan.'

We have not found any analysis of *Xistan* in morphology works. The final *-istan* is also not listed in dictionaries of exponents like Cottez (1982). Some Internet users, however, consider *-istan* to be a suffix and have dedicated an entry to it in the *Wiktionary*,[6] which provides two types of information.

- According to the collaborative platform, the form comes from the suffix *-stan*, which designates a 'place' in Persian (*dari*). It is found in Asian toponyms like *Turkmenistan*$_{Npr}$ or *Kyrgyzstan*$_{Npr}$.
- The form *-istan*, with the vowel /i/, would be a meta-analysis of *Afghanistan* which the speakers interpret as 'the country of the Afghans' and thus split into /afgã/ + /istã/.

PPN derivatives in *-istan* are imaginary countries where despotism or corruption embodied by the PPN reigns. *Wauquiezistan*$_{Npr}$ (25a) is the 'place ruled by Laurent Wauquiez in an authoritarian and dishonest way' (i.e., the French region *Auvergne-Rhône-Alpes*). *Sarkozistan*$_{Npr}$ (25b) is the 'place ruled by Nicolas Sarkozy in an authoritarian and dishonest way' (i.e., France when it was led by Nicolas Sarkozy). If we examine the toponyms in *-istan* (*Afghanistan*$_{Npr}$, *Kurdistan*$_{Npr}$, *Pakistan*$_{Npr}$), we notice that the associated countries are often at war, non-democratic or considered by the Western press as corrupt. We can therefore establish a semantic correlation between the properties of the toponym referents and the (imaginary) referents *Xistan* of our corpus. The context of the neologism *Zemmouristan*$_{Npr}$ based on Éric Zemmour,[7] cre-

---

[6]https://fr.wiktionary.org/wiki/-istan

[7]Éric Zemmour is an editorialist and politician of the extreme right.

ated by Jean-Luc Mélenchon[8] during a televised debate, illustrates our point. It is transcribed in (26). It contains the properties that Jean-Luc Mélenchon attaches to the country called *Zemmouristan*$_{\text{Npr}}$: "A country where women are demeaned, where there is the death penalty that you like, a country where homosexuals are punished and a country where we do not adhere to international conventions on refugees".

(26) *Pour le candidat LFI à la présidentielle, « le* **Zemmouristan***, ça existe » : « Un pays où les femmes sont rabaissées, où il y a la peine de mort qui vous plaît, un pays où les homosexuels sont punis et un pays où on n'adhère pas aux conventions internationales sur les réfugiés, ça s'appelle l'Arabie Saoudite! ».* (*Éric Zemmour*)
   'For the LFI presidential candidate, "the **Zemmouristan**, it exists": "A country where women are demeaned, where there is the death penalty that you like, a country where homosexuals are punished and a country where we do not adhere to international conventions on refugees, it is called Saudi Arabia!"'

Our hypothesis is that this is a new pattern of secretive suffixation. As a reminder (Fradin, 2000; Mattiello, 2007; Fradin, 2015; Mattiello, 2018), secretive suffixation is a process of segmentation change where a non-affixed part of a *model word* is reanalyzed as an affix: in this case -*istan*. The meaning related to the construction of complex words with the secreted affix contains only part of the meaning of the model word: in this case, the stereotype 'place ruled in an authoritarian and dishonest way'.

The originality of the X*istan* pattern is that it is impossible to determine a single model word. It can be hypothesized that the series of institutionalized toponyms is pressing to generate the form /istã/. -*istan* is thus a secreted element, but, unlike the secreted forms analyzed in the literature,[9] it is generated from a morphological series.

Finally, this secretion pattern is not specific to anthroponymic bases, nor to French. It applies to any participant that speakers/writers associate with one or more decisions that they consider hegemonic or totalitarian. It can be the person or company (27) that makes these decisions or the purpose to which it relates (28). For example, the documentary *Facebookistan*$_{\text{Npr}}$ (27) deals with the limits to freedom of expression imposed by the social network *Facebook*. More recently, the sanitary measures taken during the Coronavirus epidemic gave birth to *Hygiénistan*$_{\text{Npr}}$ ('hygien-istan') (28a) or *Vaccinistan*$_{\text{Npr}}$ (28b), particularly used in conspiracy discourses.

(27) **Facebookistan** *is a new documentary that takes a close look at Facebook, its laws, power and its influence on privacy and freedom of expression.*[10]

---

[8]Jean-Luc Mélenchon is a far left politician.

[9]For instance, Mattiello (2018) gives a list of related model words for each secreted form she identifies in English.

[10]https://facebookistan.org/

(28)   a. *Yes, un nouveau néologisme de Jérôme Blanchet-Gravel. « Nous ne sommes plus au Québec mais en* **Hygiénistan**. *».*[11]
       'Yes, a new neologism by Jérôme Blanchet-Gravel. "We are no longer in Quebec but in hygien-istan.".'
   b. **Vaccinistan**. *Since the start of the coronavirus pandemic, anti-vax conspiracy theories have become dominant in the disinformation landscape.*[12]

3.2. Subverted grammatical processes

We have just seen that MoNoPoli contains original complex words and analyzed two extragrammatical patterns not described in French morphology. However, the originality of our data also lies in the subversion of processes that are traditionally considered as grammatical, or more regular, in French, such as suffixation and composition. Among these, we propose to examine three types of constructions exemplified in (29). In Section 3.2.1, we analyze what we call *foreign suffixation* (29a); in Section 3.2.2 we focus on disease nouns such as *macronite*$_N$ 'macron-itis' contextualized in (29b) and in Section 3.2.3 we discuss compounds such as *macrono-vallso-hollando-montebouresque*$_{Adj}$ contextualized in (29c).

(29)   a. *N'oubliez pas* **Macronix**, *sûrement plus dangereux pour notre pays que les 3 précédemment cités réunis.* (*Emmanuel Macron*)
       'Don't forget Macron-ix, surely more dangerous for our country than the 3 previously mentioned together.'
   b. *La* **macronite** *aigüe médiatique est lourde. Ce type est nul, un jeune déjà vieux, sans carrure, arrogant, méprisant le peuple !* (*Emmanuel Macron*)
       'The media suffer of a severe case of acute macron-itis. This guy is garbage, a young man already old, without stature, arrogant, despises the people!'
   c. *Les insoumis, eux, ont finalisé leur programme, alors que les autres, à droite, font semblant d'être en désaccord quand ils sont tous pareils, et à « gauche », fausse gauche* **macrono-vallso-hollando-montebouresque**, *c'est le pataquès.* (*Emmanuel Macron*, *Manuel Valls*, *François Hollande*, *Arnaud Montebourg*)
       'The insoumis [far left party] have finalized their program, while the others, on the right, pretend to disagree when they are all the same, and on the "left", false left macrono-vallso-hollando-montebour-esque, it's a jumble.'

---

[11] https://www.facebook.com/9936886137/posts/10158771088016138/

[12] https://eooh.eu/articles/plandemic

3.2.1. Foreign suffixation

MoNoPoli contains 339 deanthroponyms constructed from a process we call *foreign suffixation*. The deanthroponyms in (30) and (31) exhibit one or more inflectional or derivational exponent(s) inherited or borrowed from another language.

- The forms of (30) are derivatives in *-ix*, *-(i)us*, *-(i)um*. These are Latin exponents. According to Blancher (2015), they are used to evoke in the mind of the speakers/writers fantasy languages, by analogy with the lexicon of *The adventures of Asterix*.
- Deanthroponyms of (31) result from the addition of a suffix on both the *first name* and the *last name* of the same PPN. The suffixes used on the two sub-names may be identical as in *Jérômo Cahuzaco*$_{\text{Npr}}$ (31a) or not, as in *Nicolaï Sarkozine*$_{\text{Npr}}$ (31b). The anthroponym endings in *-o* recall anthroponyms of Italian or Spanish origin (*Livio*, *Diego*, *Pedro*). *-aï* and *-ine* are French transpositions of the final sequences of Russian anthroponyms (*Lénine* 'Lenin', *Staline* 'Stalin').

(30) a. **Alain juppix** *dénonce la « nullité du débat ». Celui de « nos ancêtres les gaulois »,*
lancé par ce **nicolas sarkozix**. (*Alain Juppé, Nicolas Sarkozy*)
'Alain jupp-ix denounces the "pointlessness of the debate". That of "our ancestors the Gauls", launched by this nicolas sarkoz-ix.'

b. **Jospinus** *est en exil,* **hollandus** *est chassé de chez lui.* (*Lionel Jospin, François Hollande*)
'Jospin-us is in exile, holland-us is driven from his home.'

c. *Olivier Desbordes a truffé le livret d'allusions superfines à notre politique actuelle*
*"**Balladurium**, **Mitterrandium**, **Chiracium**, mysterium".* (*Édouard Balladur, François Mitterrand, Jacques Chirac*)
'Olivier Desbordes has filled the booklet with super-subtle allusions to our current politics "Balladur-ium, Mitterrand-ium, Chirac-ium, mysterium".'

(31) a. **Jérômo Cahuzaco** *mis en examen.* (*Jérôme Cahuzac*)
'Jérôm-o Cahuzac-o is indicted.'

b. *Je ne parlerai pas de monsieur* **Nicolaï Sarkozine** [...]. (*Nicolas Sarkozy*)
'I will not speak about Mr. Nicol-aï Sarkoz-ine [...].'

The derivatives of (30) are created by analogy with names of Gauls and Romans, as in the French comic strip *The adventures of Asterix* by A. Uderzo and R. Goscinny. Moreover, formally, these derivatives often mobilize long sub-names, i.e., *full names* as in (30a), following the example of what is done in the comic strip where the forms used as bases are often syntagmatic, thus also long (32).

(32) a. *ordre alphabétique* ('alphabetical order') > *Ordralphabetix*$_{\text{Npr}}$ (named *Unhygienix*$_{\text{Npr}}$ in English, according to Delesse, 2006)

b. *âge canonique* ('canonical age') > *Agecanonix*$_{\text{Npr}}$ (named *Geriatrix*$_{\text{Npr}}$ in English)

Blancher (2015), who analyzes the ludic mechanisms of *The adventures of Asterix*, calls these wordplays on anthroponyms *patronimi drôlati sistematici* (see also Kabatek, 2015). According to the author, R. Goscinny developed a model of anthroponymic game from the Gaulish *Vercingétorix*$_{\text{Npr}}$ from which he extracted the suffix form *-ix* /iks/. The writer then proceeded to systematize the use of the suffix for characters of Gaulish origin. He operated in the same way, but on other models, to generate the forms in *-us*, specific to Roman referents, in *-en* /ɛn/ (e.g., *Zøøvinsen*$_{\text{Npr}}$) for Vikings or *-is* /is/ (e.g., *Amonbofis*$_{\text{Npr}}$) for Egyptians. The choice of a suffix sequence thus evokes the ethnocultural belonging of the referent.

In our corpus, all PPN referents are French and contemporary. The suffix differences are therefore not related to the ethnicity of the referent. Rather, it is a matter of attributing stereotypes, or ethnic typing of the PPN referents, attributing to them properties (often assumed) emblematic of a people. For example, deanthroponyms in *-(i)us* and *-(i)um*, formed on the Latin model, sometimes have a warlike or conquering connotation as in (33a). The forms in *-ix* can accentuate the French side of the referent. This is the effect of meaning obtained in *montebourgix*$_{\text{Npr}}$ (33b) where the context refers to the fact that Arnaud Montebourg, former Minister of Economy, is known to have defended the relocalizing of French industry.

(33)   a. *Mais très rapidement les coriaces de l'opposition abattent les champions de* **sarkosius***, en dénichant des malversations qui virent au scandale et causent leur mort politique :* **blancus***,* **joyandetus***,* **estrosius** *sont ainsi rapidement mis hors combat.* (*Nicolas Sarkozy*, *Christian Blanc*, *Alain Joyandet*, *Christian Estrosi*)
            'But very quickly fierce of the opposition slaughter the champions of sarkosius, by unearthing malpractices which turn to the scandal and cause their political death : blanc-us, joyandet-us, estros-ius are thus quickly put out of combat.'

        b. *Il nous reste notre chef et président, le bien nommé Abraracourcix,* **airaultix** (*erotix*) *qui nous prépare des potions magiques aux effets aléatoires, et bien sûr* **montebourgix** *qui confond souvent les chefs d'entreprises gaulois avec des légionnaires romains.* (*Jean-Marc Ayrault*, *Arnaud Montebourg*)
            'We still have our chief and president, the well named Abraracourcix, airault-ix (erotix?) who prepares magic potions with random effects, and of course montebourg-ix who often confuses Gaulish company managers with Roman legionnaires.'

The method adopted by the authors of *The adventures of Asterix* to create anthroponymic wordplays is thus subverted by the speakers/writers to create nonce-formations based on PPN. These games indirectly allow speakers/writers to express their abjection towards the referents of the PPN or to ridicule them by attaching stereotypes to them. This process is not extragrammatical since it is indeed suffixation. However, the affixes are borrowings and the derivatives have an exclusively playful function.

This justifies the analysis of these derivatives as the result of subverted grammatical process.

3.2.2. Disease nouns

Another way for speakers/writers to express their views about PPN referents is to create medical-like names for diseases. MoNoPoli contains a total of 193 disease nouns that instantiate 6 different patterns shown in Table 4.

Pattern	Type	Example	Gloss
Xine	1	*Nicolas Sarkozy > sarkozine*	'sarkoz-ina'
Xpathie	4	*Claude Guéant > guéantopathie*	'guéanto-pathy'
Xide	5	*Nicolas Sarkozy > sarkozide*	'sarkoz-id'
Xose	17	*Dominique De Villepin > villepinose*	'villepin-osis'
Xite	81	*Rama Yade > ramanite*	'raman-itis'
Xmanie/mania	85	*Dominique Strauss-Kahn > dskomanie*	'dsko-mania'

Table 4. Disease nouns of MoNoPoli

These nouns question the sanity of the PPN referent or its supporters: they are 'disease caused by the PPN referent' as in (34), or 'diseased passion for the PPN referent' as in (35).

(34) a. *C'est Nathalie Kosciusko-Morizet qui prendra donc le numérique où elle pourra soigner sa* **borlose** *allergique.* (*Jean-Louis Borloo*)
'Therefore Nathalie Kosciusko-Morizet will put in charge of the digital where she will be able to get her allergic borl-osis treated.'

b. *C'est pire que la grippe cette* **macronite** *aiguë, elle se répand dans toutes les rédactions.* (*Emmanuel Macron*)
'This acute macron-itis is worse than the flu, it's spreading to every newsroom.'

c. *Marrant cette nouvelle maladie qu'est la* **sarkozine**. (*Nicolas Sarkozy*)
'Funny this new disease that is the sarkoz-ina.'

(35) a. *L'ancien Premier ministre garde l'espoir : « Je ne peux pas remonter mon handicap en pleine* **ségolènomania**.». (*Ségolène Royal*)
'The former Prime Minister remains hopeful: "I can't raise my handicap in the midst of ségolèno-mania."'.

The patterns in Table 4 are attested in the medical lexicon. In other words, the processes that manipulate these patterns are not limited to anthroponymic bases: X can be a common noun.

- For instance, X*ite* nouns are derivatives where the base specifically designates the site of an inflammation: e.g., a *bronchite*$_N$ ('bronchitis') is an 'inflammation of the bronchi'.
- The X*ose* derivatives also designate names of pathological processes where the base is the diseased organ: e.g., *dermatose* ('dermatosis') (Chebouti, 2014), but the base sometimes also designates the manifestation of the phenomenon (*furoncle*$_N$ 'furuncle' > *furonculose*$_N$ 'furunculosis'), or the origin of the disease (*bacille*$_N$ 'bacillus' > *bacillose*$_N$ 'bacillosis').

In anthroponym based games, these distinctions are irrelevant. While these exponents are used to distinguish between pathology names in medical terminology, they behave uniformly on PPNs. The learned exponents (mostly from Latin or Greek) are thus systematically subverted from their initial terminologizing function. A disease noun allows, metaphorically, to express an aversion to the PPN. This phenomenon of terminological despecialization has already been reported for other patterns handling learned exponents. Namer and Villoing (2015) and Lasserre (2016) have shown that the neoclassical exponents *-logue* and *-logie* are now added to any base X to designate— often ironically—the one who talks about X and about his so-called specialty (X*logie*).

3.2.3. Subverted compounding

MoNoPoli contains 1,925 deanthroponyms instantiating the general pattern X*o*-(X'(*o*)-)*Y*suff* where the brackets indicate the optionality of an element and the asterisk notes that the number of components at that position is 0, 1 or more, cf. (36).

(36)   a.   *Ainsi, selon une méthode éprouvée, le « camp du bien », pensant pouvoir l'achever, se livre en vain à une exégèse sémantique de sa critique du totem* **aubryo -strausskhanien**. (*Martine Aubry, Dominique Strauss-Kahn*)
'Thus, according to a tried and fruitlessly tested method, the "camp of the good", thinking to be able to finish it, engages in vain in a semantic exegesis of its criticism of the aubryo-strausskhanian totem.'

   b.   *Voilà le fruit de quinze années de* **pasquaïo-sarkozo-bessonisme**. (*Charles Pasqua, Nicolas Sarkozy, Éric Besson*)
'Here is the fruit of fifteen years of pasquaïo-sarkozo-bessonism.'

   c.   *C'est triste que le seul pendant à votre pensée unique* **bobo-marxo-stalino-taubiro-hollando-demissiono-comploto-lgbt-communiste**, *soit juste un propos « anti-système » d'extrême droite.* (*Christiane Taubira, François Hollande*)
'It's sad that the only counterpart to your boho-marxo-stalino-taubiro-hollando-resigno-conspiratoro-lgbt-communist unique thought, is just an extreme right-wing "anti-establishment" statement.'

The compound *aubryo-strausskhanien*$_{Adj}$ from (36a) has the minimal format of the pattern: X*o*-Y*suff*. It includes the adjective *strausskahnien*$_{Adj}$ ('strausskahn-ian') and the sub-names *Aubry* suffixed by /o/. It is the same /o/ that we have already ob-

served in Section 3.1. It is the typical vowel of learned compounding, subverted from its usual function since, here again, the stem is not inherited. This compound adjective is interpreted as a coordination: 'aubry-ist and strausskahn-ian'. We can see that the first component is a suffixed truncated adjectival form. The meaning of the compound guides our analysis: since the compound is coordinative and the coordinated elements are, by definition, of the same morpho-semantic type, we conclude that if Y*suff* is a denominative adjective, X is a denominative adjective like Y*suff*. So X is probably the truncated form of the relational adjective *aubryiste*$_{Adj}$ ('aubry-ist') of the sub-name *Aubry* (*aubryien*$_{Adj}$ 'aubry-ian' is attested with a lower frequency).

In (36b), we see that the minimal pattern can be extended to X*o*-X'*o*-Y*suff*. We can thus accumulate constituents in /o/. The compound is coordinative as in (36a). So, X*o* constituents are truncated forms of deanthroponyms of the same nature as Y*suff* *bessonisme*$_N$ ('besson-ism') which refers to the ideology of Eric Besson. They are truncated forms of the common nouns *pasquaïsme*$_N$ ('pasqua-ism') and *sarkozysme*$_N$ ('sarkoz-ism').

The examination of (36c), finally shows that what counts for the writer is above all the rhyme in /o/, since *bobo* 'boho' is not suffixed. Moreover, in *bobo-marxo-stalino-taubiro-hollando-demissiono-comploto-lgbt-communiste*$_{Adj}$, the accumulation of components to the left of the final suffixed component Y*suff* (*communiste*$_{Adj}$ 'communist') is not limited to a suffixed form in /o/. Indeed, one of the constituents is the acronym LGBT. In any case, all these forms always refer to adjectival properties, as does the last constituent *communiste*$_{Adj}$. The writer's goal is to accumulate as many constituents as possible, like an outburst, to distinguish himself.

These compounds are exclusively coordinative. Moreover, the more constituents the writer adds, the more the effect of meaning obtained is that of a cacophony. The longer the deanthroponym, the more original and remarkable it is. In conclusion, even if the compounding process is grammatical, compound deanthroponyms are not always grammatical (especially when they involve more than two bases). In our corpus, compounding is sometimes subverted in the benefit of the writer's argumentation or humor.

4. Conclusion

The method used to create the MoNoPoli database is replicable and adaptable to other languages or other inputs (bases or affixes). The database created is accessible online. It provides a large corpus of contextualized deanthroponyms, which to our knowledge does not exist in French.

MoNoPoli also displays a diverse and large panel of morphological processes. Its contents are relevant for further research on words based on anthroponyms or French nonce-formations. Its analysis reveals that deanthroponyms are often nonce-formations sometimes constructed by extragrammatical processes. We have also shown that grammatical processes can be subverted to satisfy the enunciative needs of the

writer. These needs demonstrate, at the same time, the writer's epilinguistic capacity to play with language.

To our knowledge, intensifying reduplication (*sarkozo-sarkoziste*), X*istan* pattern secretion, foreign suffixation (*sarkozix*), disease names subversion (*macronite*), and iterative composition (*bolcho-bayrouo-trotsko-villepiniste*) are all processes that have never been described in French. We have shown that these processes were not systematically specific to anthroponymic bases, nor to French (e.g., X*istan*). However, these new patterns seem specific to nonce-formations.

In the future, the content of MoNoPoli will be used as a basis for answering questions that we have not discussed in this paper.

- Do anthroponyms form a homogeneous category? For instance, do the names of authors, the names of journalists or collective anthroponyms have the same morphological functioning as proper names of politicians (e.g., are they the basis of the same processes)?
- Does the collection of deanthroponyms from other languages lead to the same results? The French are known to be dissenting, and are often portrayed as such in the international press. They express via morphological creation an appreciative dimension towards the referents of PPNs. Will another political culture translate into the language of its citizens by less expressive or less varied results?
- Do proper names form a homogeneous category? We could extend our study to other categories of proper names (e.g, toponyms, ergonyms). One of the objectives of this analysis would be to clarify whether the notion of sub-name is a specific feature of anthroponyms or whether other units possess similar dimensions (i.e., group together a set of syntactically autonomous and co-referential lexical units). This last question will feed the discussions about the definition of the nominal category and the units manipulated as bases in morphology.

Acknowledgements

Bibliography

Bat-El, Outi. Selecting the best of the worst : The grammar of Hebrew blends. *Phonology*, 13: 283–328, December 1996. doi: 10.1017/S0952675700002657.

Bauer, Laurie. *English Word-Formation*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge, 1983. ISBN 978-0-521-28492-9. doi: 10.1017/CBO9781139165846.

Bauer, Laurie. System vs. norm: coinage and institutionalization. In Booij, Geert, Christian Lehmann, and Joachim Mugdan, editors, *Morphology: an International Handbook of*

Inflection and Word Formation, pages 832–840. De Gruyter, Berlin, New York, 2000. doi: 10.1515/9783110111286.1.11.832.

Blancher, Marc. « Ça est un bon mot ! » ou l'humour (icono-)textuel à la Goscinny. In Winter-Wroemel, Esme and Angelika Zirker, editors, *Enjeux du jeu de mots : Perspectives linguistiques et littéraires*, pages 273–290. De Gruyter, Berlin, München, Boston, October 2015. ISBN 978-3-11-040834-8.

Bonami, Olivier, Gilles Boyé, and Françoise Kerleroux. L'allomorphie radicale et la relation flexion-construction. In Fradin, Bernard, Françoise Kerleroux, and Marc Plénat, editors, *Aperçus de morphologie*, pages 103–126. Presses Universitaires de Vincennes, Saint-Denis, 2009.

Chebouti, Karim. *Le vocabulaire médical du point de vue des trois fonctions primaires.* Phd thesis, Université Paris 13, 2014.

Cottez, Henri. *Dictionnaire des structures du vocabulaire savant : éléments et modèles de formation*. Les Usuels du Robert. Le Robert, Paris, 1982. ISBN 978-2-85036-090-9.

Dal, Georgette and Dany Amiot. La composition néoclassique en français et l'ordre des constituants. In Amiot, Dany, editor, *La composition dans une perspective typologique*, pages 89–113. Artois Presses Université, Arras, 2008.

Dal, Georgette and Fiammetta Namer. Internet. In Müller, Peter O., Ingeborg Ohnheiser, Susan Olsen, and Franz Rainer, editors, *Word-Formation:An International Handbook of the Languages of Europe*, pages 2372–2386. De Gruyter Mouton, Berlin, New York, 2015.

Dal, Georgette and Fiammetta Namer. Productivity. In Stump, Gregory and Andrew Hippisley, editors, *The Cambridge Handbook of Morphology*. Cambridge University Press, Cambridge, 2016. doi: 10.1017/9781139814720.

Dal, Georgette and Fiammetta Namer. Playful nonce-formations in French: Creativity and productivity. In Arndt-Lappe, Sabine, Angelika Braun, Claudine Moulin, and Esme Winter-Froemel, editors, *Expanding the Lexicon Linguistic Innovation, Morphological Productivity, and Ludicity*, number 5 in The Dynamics of Wordplay, pages 203–228. De Gruyter, Berlin, Boston, 2018. doi: 10.1515/9783110501933-205.

Delesse, Catherine. Les noms propres dans la série Astérix et leur traduction anglaise. *Palimpsestes. Revue de traduction*, Hors série:297–315, September 2006. ISSN 1148-8158. doi: 10.4000/palimpsestes.1067.

Dressler, Wolfgang U. Extragrammatical vs marginal morphology. In Doleschal, Ursula and Anna Thornton, editors, *Marginal and Extragrammatical Morphology*, pages 1–10. Lincom Europa, Müchen, 2000.

Dressler, Wolfgang U. and Lavinia M. Barbaresi. *Morphopragmatics*. Trends in Linguistics. Studies and Monographs, 76. De Gruyter Mouton, reprint 2011 ed. edition, 1994. ISBN 978-3-11-014041-5. doi: 10.1075/hop.m2.mor1.

Fradin, Bernard. Combining forms, blends and related phenomena. In Doleschal, Ursula and Anna M. Thornton, editors, *Extragrammatical and Marginal Morphology*, pages 11–59. Lincom Europa, München, 2000.

Fradin, Bernard. *Nouvelles approches en morphologie*. Linguistique Nouvelle. Presses Universitaires de France, Paris, 2003. doi: 10.3917/puf.fradi.2003.01.

Fradin, Bernard. Blending. In Müller, Peter O., Ingeborg Ohnheiser, Susan Olsen, and Franz Rainer, editors, *Word-Formation – An International Handbook of the Languages of Europe*, volume Volume 1, page 386. De Gruyter Mouton, Berlin, New York, 2015.

Fradin, Bernard, Georgette Dal, Natalia Grabar, Stéphanie Lignon, Fiammetta Namer, Delphine Tribout, and Pierre Zweigenbaum. Remarques sur l'usage des corpus en morphologie. *Langages*, 171(3):34–59, 2008. ISSN 0458-726X. doi: 10.3917/lang.171.0034.

Fradin, Bernard, Fabio Montermini, and Marc Plénat. Morphologie grammaticale et extragrammaticale. In *Aperçus de morphologie du français*, Sciences du langage, pages 22–45. Presses Universitaires de Vincennes, Saint-Denis, 2009.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen, and Kevin Russell. Contrastive Focus Reduplication in English (The Salad-Salad Paper). *Natural Language & Linguistic Theory*, 22(2): 307–357, 2004. ISSN 0167-806X. doi: 10.1023/B:NALA.0000015789.98638.f9.

Goldsmith, John. *Autosegmental phonology*. Phd thesis, MIT, 1976.

Gérard, Christophe. Variabilité du langage et productivité lexicale: Problèmes et propositions méthodologiques. *Neologica*, 12:23–45, 2018. doi: 10.15122/ISBN.978-2-406-08196-8.P.0023.

Hohenhaus, Peter. Identical Constituent Compounding – a Corpus-based Study. *Folia Linguistica*, 38(3-4):297–332, 2004. ISSN 1614-7308. doi: 10.1515/flin.2004.38.3-4.297.

Hohenhaus, Peter. Lexicalization and institutionalization. In Štekauer, Pavol and Rochelle Lieber, editors, *Handbook of Word-Formation*, Studies in Natural Language and Linguistic Theory, pages 353–370. Springer, Dordrecht, 2005. doi: 10.1007/1-4020-3596-9_15.

Hohenhaus, Peter. How to do (even more) things with nonce words (other than naming). In Munat, Judith, editor, *Lexical Creativity, Texts and Contexts*, Studies in Functional and Structural Linguistics 58, pages 15–38. John Benjamins Publishing Company, Amsterdam, Philadelphia, 2007. ISBN 978-90-272-1567-3. doi: 10.1075/sfsl.58.08hoh.

Hohenhaus, Peter. Anti-naming through non-word-formation. *SKASE Journal of Theorical Linguistics*, 12(3):272–291, 2015.

Huguin, Mathilde. Anthroponyms and paradigmatic derivation in French. *Lingue e Linguaggio. Defining paradigms in word formation*, XVII(2):217–232, 2018. doi: 10.1418/91866.

Huguin, Mathilde. *Analyse morphologique des mots construits sur base de noms de personnalités politiques*. Phd thesis, Université de Lorraine, December 2021.

Kabatek, Johannes. Wordplay and Discourse Traditions. In Zirker, Angelika and Esme Winter-Wroemel, editors, *Wordplay and Metalinguistic / Metadiscursive Reflection: Authors, Contexts, Techniques, and Meta-Reflection*, pages 213–228. De Gruyter, Berlin, München, Boston, October 2015. ISBN 978-3-11-040671-9. doi: 10.1515/9783110406719-010.

Kerremans, Daphné. *A Web of New Words: A Corpus-Based Study of the Conventionalization Process of English Neologisms*. English Corpus Linguistics. Peter Lang GmbH, Internationaler Verlag der Wissenschaften, 2015. ISBN 978-3-631-65578-8. doi: 10.1515/east-2016-0007.

Kilani-Schoch, Marianne and Wolfgang U. Dressler. *Morphologie naturelle et flexion du verbe français*. Gunter Narr Verlag, Tübingen, 2005.

Kleiber, Georges. *La Sémantique du prototype. Catégories et sens lexical*. Presses Universitaires de France, Paris, 1990.

Lasserre, Marine. *De l'intrusion d'un lexique allogène. L'exemple des éléments néoclassiques*. PhD thesis, Université Jean Jaurès, Toulouse, 2016.

Leroy, Sarah. Les noms propres et la dérivation suffixale. *Neuphilologische Mitteilungen*, 109: 55–71, 2008.

Lignon, Stéphanie, Fiammetta Namer, Nabil Hathout, and Mathilde Huguin. When sarkozysation leads to the hollandade, or the rejection of phonological well-formedness constraints by anthroponym-based derived words. In *International Symposium of Morphology (ISMo) 2019*, Paris, France, September 2019.

Lipka, Leonhard. Lexical creativity, textuality and problems of metalanguage. In Munat, Judith, editor, *Lexical Creativity, Texts and Contexts*, Studies in Functional and Structural Linguistics 58. John Benjamins Publishing Company, 1st edition, 2007. ISBN 978-90-272-1567-3. doi: 10.1075/sfsl.58.06lip.

Lüdeling, Anke, Stefan Evert, and Marco Baroni. Using web data for linguistic purposes. In Hundt, Marianne, Nadja Nesselhauf, and Carolin Biewer, editors, *Corpus Linguistics and the Web*, number 59 in Language and Computers: Studies in Practical Linguistics, pages 7–24. Rodopi, Amsterdam, New York, 2007.

Mattiello, Elisa. Combining forms and blends: The case of scape. In Jottini, Laura, Gabriella Del Lungo, and John Douthwaite, editors, *Cityscapes: Islands of the Self. Language Studies*, volume 2, pages 115–130, Cagliari, 2007. CUEC.

Mattiello, Elisa. Paradigmatic morphology splinters, combining forms, and secreted affixes. *SKASE Journal of Theoretical Linguistics*, 15(1):2–22, January 2018.

Moravcsik, Edith. Reduplicative Constructions. In Greenberg, Joseph H., editor, *Universals of Human Language*, number 3 in Word Structure, pages 297–334. Stanford University Press, Standford, 1978.

Namer, Fiammetta and Florence Villoing. Composition néoclassique en –logue et en –logiste : les noms en –logue sont ils encore des noms de spécialistes ? *Verbum*, 34(2):213–231, 2015.

Plénat, Marc. Les contraintes de taille. In Plenat, Marc, Bernard Fradin, and Françoise Kerleroux, editors, *Aperçus de morphologie du français*, Sciences du langage, pages 47–64. Presses Universitaires de Vincennes, Saint-Denis, 2009.

Plénat, Marc and Michel Roché. Prosodic constraints on suffixation in French. In Booij, Geert, Janet DeCesaris, Angela Ralli, and Sergio Scalise, editors, *Topics in Morphology. Selected Papers from the Third Mediterranean Morphology Meeting*, pages 285–299, Barcelona, September 2001. IULA-Universitat Pompeu Fabra (Barcelona).

Plénat, Marc, Stéphanie Lignon, Nicole Serna, and Ludovic Tanguy. La conjecture de Pichon. *Corpus et recherches linguistiques*, 1:105–150, 2002. doi: 10.4000/corpus.15.

Prince, Alan and Paul Smolensky. Optimality theory: constraint interaction in generative grammar. *Technical Report, Rutgers University Center for Cognitive Science and Computer Science Department*, 1993. doi: 10.1002/9780470756171.ch1.

Roché, Michel. Base, thème, radical. *Recherches linguistiques de Vincennes*, 39:95–134, December 2010. ISSN 0986-6124. doi: 10.4000/rlv.1850.

Roché, Michel and Marc Plénat. Le jeu des contraintes dans la sélection du thème présuffixal. In *SHS Web Conferences*, volume 8, pages 1863–1878. EDP Sciences, 2014. doi: 10.1051/shsconf/20140801143.

Scalise, Sergio, Antonietta Bisetto, and Emiliano Guevara. Selection in Compounding and Derivation. In Dressler, Wolfgang Ulrich, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer, editors, *Morphology and its demarcations: Selected papers from the 11th Morphology meeting, Vienna, February 2004*, number 264 in Current Issues in Linguistic Theory, pages 133–150. John Benjamins Publishing Company, Amsterdam, Philadelphia, 2005. doi: 10.1075/cilt.264.09sca.

Schweickard, Wolfgang. *"Deonomastik". Ableitungen auf der Basis von Eigennamen im Französischen*. De Gruyter, Tübingen, 1992. doi: 10.1515/9783110933901.

Tanguy, Ludovic. *Complexification des données et des techniques en linguistique : contribution du TAL aux solutions et aux problèmes*. Habilitation à diriger des recherches, Université Toulouse-Le Mirail, Toulouse, 2012.

Zwicky, Arnold M. and Geoffrey K. Pullum. Plain morphology and expressive morphology. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, pages 330–340, 1987. doi: 10.3765/bls.v13i0.1817.

Štekauer, Pavol. On the Theory of Neologisms and Nonce-formations. *Australian Journal of Linguistics*, 22(1):97–112, April 2002. ISSN 0726-8602. doi: 10.1080/07268600120122571.

Address for correspondence:
Mathilde Huguin
`mathilde.huguin@univ-lorraine.fr`
ATILF (CNRS & Université de Lorraine)
44, avenue de la Libération
B.P. 30687 54063 Nancy Cedex, France