8

Paradigm Uniformity and the French Gender System

Olivier Bonami and Gilles Boyé

8.1 Introduction

Corbett's view of canonical inflection (see, for example, Corbett 2009) holds that, in the canonical case, all lexemes belonging to the same part of speech should have the same paradigm structure. Obvious violations of the canon are defectivity (a lexeme missing a paradigm cell) and overdifferentiation (a lexeme with an extra paradigm cell). These are usually defined in terms of lexical exceptionalism: we have an expected paradigm shape, and a few irregular lexemes unexpectedly deviate from that paradigm shape. Another interesting family of deviations from paradigm uniformity involve situations where there is a systematic distinction of multiple paradigm shapes within a single part of speech. One obvious example comes from conjugation in languages exhibiting object agreement: clearly, in such languages, intransitive and transitive verbs have different paradigm shapes, and this is not a matter of lexical exceptionalism. In this paper we report on what we take to be another systematic case of paradigmatic non-uniformity resulting from the distribution of gender on personal nouns in contemporary French. Table 8.1 summarises the types of gender behaviour exhibited by French nouns, which we comment on in detail below.

In languages whose grammatical gender system opposes a masculine and a feminine, it is commonly the case that most nouns referring to males are masculine and most nouns referring to females are feminine. We refer to such nouns as gender-iconic nouns. Hence, French masculine *homme* 'man' and feminine *femme* 'woman' are gender-iconic nouns, but feminine *personne* '(male or female) person', *table* 'table' or *sentinelle* 'watchman' are not. It is also common, in such languages, for many gender-iconic nouns to come in pairs of morphologically related words; for instance, *instituteur* 'male schoolteacher' and *institutrice* 'female schoolteacher' belong to the same morphological family, as do *tigre* 'tiger' and *tigresse* 'tigress'. Such pairs of nouns we call gender-iconic pairs.

The main issue addressed in this paper is the morphological status of gender-iconic

Table 8.1 Attested gender assignment situations for French nouns

Type	Simplex	Derived
Personal nouns		
Single-gendered,		
non-iconic	personne$_F$ 'person'	mauvi-ette$_F$ 'wimp'
iconic	homme$_M$ 'man'	ménag-ère$_F$ 'housewife'
Common gender	enfant$_{M/F}$ 'child'	dent-iste$_{M/F}$ 'dentist'
Non-homophonous pairs	avocat$_M$/avocate$_F$ 'lawyer'	jou-eur$_M$/jou-euse$_F$ 'player'
Inanimate nouns	table$_F$ 'table'	lav-erie$_F$ 'laundry'
Single-gendered	clope$_{M/F}$	auto-route$_{M/F}$
Common gender	'cigarette' (informal)	'highway'
Non-homophonous pairs	ravin$_M$/ravine$_F$	photocopi-eur$_M$/photocopi-euse$_F$
	'ravine'	'copy machine'

pairs. A century-old line of argumentation takes gender-iconic pairs to be pairs of derivationally related lexemes (e.g. Nyrop 1936; Zwanenburg 1988; Matthews 1991). A main motivation for this position seems to be the very definition of grammatical gender as a classification of nouns (see Corbett 1991 for detailed discussion): if gender classifies nouns, then every noun should have one and only one gender. This tradition, however, is in striking contrast with the practice of lexicographers and traditional grammarians, who uniformly list gender-iconic pairs under a single entry. This suggests a conception where gender-iconic pairs correspond to a single lexeme, with each gender-iconic noun constituting a slab of that lexeme's paradigm.

Closely related to that issue is the status of common gender nouns, i.e. situations where the exact same form can be used in the masculine or feminine, with the use of grammatical gender matching social gender; compare *le dentiste* 'the male dentist' with *la dentiste* 'the female dentist' (Corbett 1991: 67, 181–2).[1] There are two possible views of such nouns, which are linked with the two possible views of gender-iconic pairs as derivationally or inflectionally related: (i) a common gender noun could be taken to have just two paradigm cells, and be underspecified for gender; or (ii) a common gender noun could be taken to be a gender-iconic pair, i.e. a pair of a masculine and a feminine noun, where the masculine and feminine forms happen to be homophonous. Note that both views are compatible with both approaches to gender-iconic pairs as derivationally or inflectionally related, although, as we will see below, there is some degree of congruence between the two issues.

Whether gender-iconic pairs are inflectionally or derivationally related should be decided, we argue, on a language-by-language basis, by examining which means the morphology deploys to relate gender-iconic nouns. In §8.2, we collect relevant empirical evidence on the French situation. We first evaluate the prevalence of common gender nouns and gender-iconic pairs in the lexicon, showing that both are

far too high to be considered lexically exceptional. We then examine how productive lexeme formation processes derive new personal nouns. We show that the formation of masculine and feminine personal nouns almost always goes in parallel, either through the formation of a common gender noun, or through parallel affixations.

In §8.3 we discuss the theoretical consequences of our findings, and argue that, for contemporary French, the traditional view of the lexicographers is the correct one: gender-iconic pairs correspond to a single inflectional paradigm. Our argument is twofold. First, we argue that the productive formation of common gender nouns cannot be reconciled with the view that normal nouns are gender specific: a vast, open and quickly growing family of French nouns are compatible with both genders. Second, we argue that parallel derivation of masculine and feminine forms for gender-iconic pairs can only be accommodated by postulating that gender-iconic pairs correspond to a single inflectional paradigm. We conclude that there cannot be paradigmatic uniformity of nouns in French: almost all inanimate nouns indisputably have only two paradigm cells, while thousands of personal nouns have four.

A study of the gender system of contemporary French cannot be undertaken without taking into account the rapid evolution of the system under social pressure, both in the form of language planning (see, for example, Bousquet and Abily 2015) and spontaneous evolution. As a striking piece of anecdotal evidence, the noun *médecin* 'physician' is traditionally masculine, and had no recognised feminine counterpart until the beginning of the twenty-first century, as evidenced by examination of both the Google Books and Frantext collections of texts. However, using *médecin* in the feminine when referring to a female, as in the following newspaper example, has become the de facto standard in recent years, despite much conservative prescriptive outrage. Note also the use of *gynécologue* as a common gender noun, and the explicitly feminine form of *obstétricienne*.[2]

(1)   Le tribunal dit que le décès de l'enfant est imputable à des fautes commises par la
        médecin gynécologue obstétricienne. (*Le Télégramme*, 31 July 2006, T. Charpentier)
        'The court states that the death of the child is due to mistakes made by the$_F$ **medical_
        doctor gynecologist obstetrician**$_F$.'

In this paper we do our best to document actual usage in a quickly evolving domain where conscious planning is frequent, while making abstraction both of political debate on the relationship between social and grammatical gender, and of the numerous fundamental sociolinguistic questions raised by the evolution of the system.

8.2 Empirical Evidence

In this section we assess empirically the status of common gender nouns and gender-iconic pairs in French. We first examine the prevalence of common gender nouns and gender-iconic pairs in the extant lexicon, as documented in dictionaries and other lexical resources. We then examine the organisation of lexeme formation processes producing nouns with respect to gender.

8.2.1 Gender in the Stable Lexicon

We begin by examining the distribution of common gender nouns and gender-iconic pairs in the French lexicon. The Morphalou lexicon (Romary et al. 2004) is a machine-readable French lexicon derived from information contained in the *Trésor de la langue française* dictionary. As such, it provides explicit information about pairs of morphologically related nouns differing only in gender, irrespective of whether the two forms are homophonous and of whether the noun refers to an animate entity. This gives us a quantitative basis to evaluate how prevalent these are. As Table 8.2 indicates, about 15 per cent of feminine nouns form a pair with a masculine, and, conversely, about 15 per cent of the masculines form a pair with a feminine.

These proportions are hard to interpret, however, because of the high prevalence of inanimate nouns. First, there exist some pairs of gender-differentiated synonymous inanimate nouns (e.g. *photocopieur*$_M$ vs *photocopieuse*$_F$ 'copy machine') or pairs of an animate and an inanimate noun (e.g. *perceur*$_M$ 'piercer' vs *perceuse*$_F$ 'drill'), so that it is not obvious what proportion of the 4,441 paired nouns actually are common gender nouns or gender-iconic pairs. Second, it would be more informative to know the relative type frequency of these types of nouns among nouns with human reference; however, Morphalou does not document any semantic information. To make up for that limitation, we rely on a handmade classification by a research assistant of all the nouns in the Flexique lexicon (Bonami et al. 2014),[3] indicating for each noun whether it has established uses referring to a human, animal, inanimate or abstract entity, as documented in lexicographic sources. There are 24,990 nouns in Morphalou that are also fully documented in Flexique; among these, 4,544 were validated as personal nouns. As Table 8.3 indicates, among the validated personal nouns, 78 per cent of feminine nouns and 51 per cent of masculine nouns are associated with a noun of contrasting gender.

We now turn to the distribution of common gender nouns. These are listed as paired nouns in Morphalou, and hence included in the counts in Table 8.3. To assess which pairs have homophonous masculine and feminine forms, we use transcriptions

Table 8.2 Types of M and F nouns in the overall Morphalou lexicon

	With associate	Without associate	Proportion
Feminine nouns	4,441	28,223	15%
Masculine nouns	4,441	28,276	15%

Table 8.3 Types of validated personal F and M nouns in the Morphalou lexicon

	With associate	Without associate	Proportion
Feminine nouns	2,021	575	78%
Masculine nouns	2,021	1,948	51%

Table 8.4 Homophonous and non-homophonous pairs of personal nouns in Morphalou

	M = F	M ≠ F	Prop. M = F
Validated personal nouns	846	1,175	42%

from Flexique (Bonami et al. 2014). As Table 8.4 indicates, it turns out to be the case for 42 per cent of our validated personal nouns.[4]

In the end then, we have established the existence of at least 846 common gender nouns and another 1,175 gender-iconic pairs with non-homophonous masculines and feminines in French. These correspond in turn to 33 per cent (homophonous) and 45 per cent (non-homophonous) of all validated feminine personal nouns, and respectively to 21 per cent (homophonous) and 30 per cent (non-homophonous) of all validated masculine personal nouns. These numbers should be taken to be low estimations, in terms of both absolute and relative frequency. In absolute terms, remember that Flexique is a relatively small lexicon; in particular, remember that Morphalou contained about two times more pairs of morphologically related nouns contrasting in gender. In relative terms, the proportion of common gender nouns and gender-iconic pairs is certain to be underestimated. As we said before, Morphalou derives from the *Trésor de la langue française*, a dictionary constructed between the late 1960s and the early 1990s, and intended to reflect usage from the late eighteenth century to 1960. Given social change in the last half century, and a strong push towards using gender-iconic nouns for professions and activities, we have a strong expectation that the prevalence of gender-iconic pairs in contemporary usage is significantly higher. A precise estimation will have to await future research,[5] but it is important for the arguments to follow to remember that, if anything, we are underestimating the prevalence of common gender nouns and gender-iconic pairs.

A final piece of evidence that can be derived from the present dataset is the distribution of morphophonological alternations within gender-iconic pairs. This is indicated in Table 8.5. It is striking that the alternation types are well-behaved, and

Table 8.5 Phonological alternation types among confirmed Morphalou gender-iconic pairs

Alternation	Count	Example		
		M	F	Translation
X ~ XC	445	avocat	avocate	'lawyer'
Xœʁ ~ Xøz	318	joueur	joueuse	'player'
XṼ ~ XVn	238	voisin	voisine	'neighbour'
Xœʁ ~ Xʁis	145	auditeur	auditrice	'listener'
Xf ~ Xv	11	veuf	veuve	'widow(er)'
X ~ Xɛs	10	traître	traîtresse	'traitor'
Xo ~ Xɛl	6	jumeau	jumelle	'twin'
XṼ ~ XVnœʁɛs	1	devin	devineresse	'soothsayer'
Xk ~ Xʃɛs	1	archiduc	archiduchesse	'archduke'

match almost exactly the types of alternations found between the masculine and feminine forms of adjectives (Bonami and Boyé 2005): loss of a final consonant in the masculine with possible nasalisation of the preceding vowel, devoicing of /v/ /o/ ~ /ɛl/ alternation, and the two suffix pairs *-eur/-euse* (attaching to the default stem) and *-eur/-rice* (attaching to a learned stem). The only exceptions are the last two types, each found in only one lexeme, and which exhibit a combination of the use of the feminine suffix *-esse* and some other morphological or phonological process.[6]

Another striking observation is that, overall, the language almost never uses an affix to derive a feminine personal noun from a masculine one or the other way around. Only *-esse* is used for that purpose, and this happens only 0.5 per cent of the time. This suggests that we should take a more thorough look at the place of gender in the lexeme formation system, which we turn to presently.

8.2.2 Gender in Lexeme Formation

In this section we examine how the French lexeme formation system constrains the gender of derived nouns.

8.2.2.1 Suffixation

Most processes forming nouns involve suffixes. French has plenty of independent gender-specific affixes forming nouns denoting inanimate entities, as illustrated in Table 8.6.

It is striking, however, that there barely is any suffixal process specifically devoted to forming feminine personal nouns. The only relevant suffix is *-esse$_2$*. However, this suffix is little used: we found only 46 relevant examples in the *Trésor de la langue française*, and 69 attested in the massive Google ngrams dataset (Michel et al. 2010). Moreover, the suffix has stopped being productive for a long time. The *Trésor de la langue française* documents only 8 coinings in the nineteenth century, and none after 1867. The Google ngrams dataset provides three relevant *-esse* nouns with a later first attestation date, the youngest being *emmerderesse* 'annoying woman (vulgar)' first seen in 1955, and clearly a play on words. In addition to not being productive, *-esse$_2$*

Table 8.6 French gender-specific suffixes

| Masculine	-age	marier 'to marry'	mari-age 'wedding'
	-ment	sentir 'to feel'	senti-ment 'feeling'
	-at	assassin 'murderer'	assassin-at 'murder'
	-isme	race 'race'	rac-isme 'racism'
Feminine	-ion	presser 'to press'	press-ion 'pressure'
	-ité	digne 'dignified'	dign-ité 'dignity'
	-ure	blesser 'to wound'	bless-ure 'wound'
	-ance	confier 'to confide'	confi-ance 'confidence'
	-esse$_1$	rude 'rough'	rudesse 'roughness'
	-esse$_2$	traître 'male traitor'	traîtresse 'female traitor'
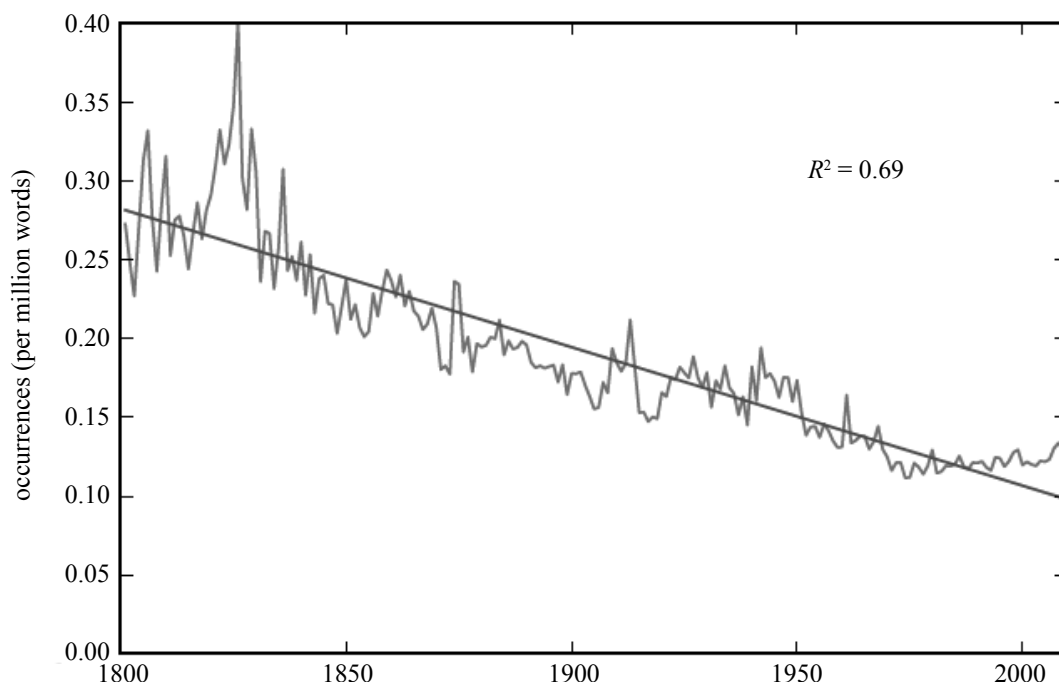
Figure 8.1 Cumulated relative frequency over time of feminine personal nouns in *-esse*. All nouns are taken from the *Trésor de la langue française* dictionary, frequency data from the Google ngrams dataset

nouns are progressively disfavoured by usage since the middle of the nineteenth century, as indicated by the decay of their use in the Google ngram dataset (Figure 8.1).

It is thus a very clear property of the contemporary French system that creation of new feminine personal nouns by suffixation to the corresponding masculine noun simply does not happen.

On the other hand, many affixes are compatible with both genders, as exemplified in Table 8.7. Most of these affixes can form personal nouns (the single exception is *-oir/-oire*), and hence affixes compatible with both genders are a major source of new common gender nouns.

Finally, many French affixes compatible with an animate denotation actually come in pairs of related affixes, combining a masculine and a feminine form in a clear

Table 8.7 French suffixes agnostic to gender

-able	contribuer 'contribute'	contribu-able_{M/F} 'taxpayer'
-aire	révolution 'revolution'	révolutionn-aire_{M/F} 'revolutionary'
-iste	journal 'newspaper'	journal-iste_{M/F} 'reporter'
-ite	Jésus 'Jesus'	jésu-ite_{M/F} 'Jesuit'
-oir(e)	raser 'to shave'	ras-oir_M 'razor'
	baigner 'to bathe'	baign-oire_F 'bathtub'

Table 8.8 Pairs of morphologically related suffixes with contrasting genders

-ain/aine	Tibet 'Tibet'	tibét-ain$_M$	'male Tibetan'
		tibét-aine$_F$	'female Tibetan'
-ien/ienne	Italie 'Italy'	ital-ien$_M$	'male Italian'
		ital-ienne$_F$	'female Italian'
-ais/aise	France 'France'	franç-ais$_M$	'Frenchman'
		franç-aise$_F$	'French woman'
-ant/ante	perdre 'to lose'	perd-ant$_M$	'male loser'
		perd-ante$_F$	'female loser'
-ier/ière	police 'police'	polic-ier$_M$	'policeman'
		polic-ière$_F$	'policewoman'
-eur/euse	chasser 'to hunt'	chass-eur$_M$	'hunter'
		chass-euse$_F$	'huntress'
-eur/rice	inspecter 'to inspect'	inspect-eur$_M$	'male inspector'
		inspect-rice$_F$	'female inspector'
-on/onne	sauvage 'savage'	sauvage-on$_M$	'wild boy'
		sauvage-onne$_F$	'wild girl'
-et/ette	poule 'hen'	poul-et$_M$	'cockerel'
		poul-ette$_F$	'young hen'
-in/ine	plaisanter 'to joke'	plaisant-in$_M$	'male joker'
		plaisant-ine$_F$	'female joker'

morphological relation. These pairs of affixes are a prolific source of gender-iconic pairs.[7] Table 8.8 provides relevant examples.

These observations on gender and French suffixes already provide motivation for most of the observations derivable from Table 8.5. There are few types of alternations in gender-iconic pairs because the grammar systematically proposes paired strategies for coining new personal nouns: either a single suffix used for both the masculine and the feminine, or two morphologically related suffixes for the masculine and feminine. It is no coincidence that the processes under examination can also be used to form adjectives, and hence need to provide parallel strategies for masculine and feminine forms.[8] Be that as it may, the end result is that gender-iconic pairs typically have the exact same structure as adjectival paradigms.[9]

8.2.2.2 Composition

Having considered suffixal lexeme formation, we turn to other processes, and establish a surprising generalisation: all non-suffixal processes that form new personal nouns form nouns of common gender. We start with different kinds of compounding and then move to non-concatenative processes.

French VN compounds (Villoing 2009) provide a striking example of the strong productivity of common gender personal nouns. It is a well-established generalisation that inanimate VN compounds are always masculine. This, however, is not true in contemporary French for those VN compounds that are personal nouns. Table 8.9 illustrates the general pattern, and (2) exhibits some attested examples.

Table 8.9 Examples of VN compounds

Inanimate	
ouvrir 'open', boîte$_F$ 'can'	ouvre-boîte$_M$ 'can-opener'
presser 'to press', papier$_M$ 'paper'	presse-papier$_M$ 'paper-weight'
tirer 'to pull', bouchon$_M$ 'cork'	tire-bouchon$_M$ 'cork-screw'
Personal nouns	
casser 'to break', pied$_M$ 'foot'	casse-pied$_{M/F}$ 'irritating person'
rabattre 'to lower', joie$_F$ 'joy'	rabat-joie$_{M/F}$ 'kill-joy'
porter 'to carry', parole$_F$ 'speech'	porte-parole$_{M/F}$ 'spokesperson'
lécher 'to lick', botte$_F$ 'boots'	lèche-botte$_{M/F}$ 'boot-licker'
piquer 'to steal', assiette$_F$ 'plate'	pique-assiette$_{M/F}$ 'freeloader'
pisser 'to piss', copie$_F$ 'copy'	pisse-copie$_{M/F}$ 'hack'
briser 'to break', coeur$_M$ 'heart'	brise-coeur$_{M/F}$ 'heart-breaker'

(2) a. Ce **casse-pied** de photographe nous aura fait gagner pas mal de temps. (*Terreur sur Saïgon*, 2014, Philippe Geluck, p. 29)
   'This$_M$ **ball-breaker** of a photographer finally saved us quite some time.'
   b. Ta mère est une belle **casse-pieds**, et ton mec réagit au quart de tour. (forum. aufeminin.com, 15 October 2012, adoravel)
   'Your mother is a$_F$ perfect$_F$ **ball-breaker**, and your guy responds instantly.'

We observe the same general situation with other types of compounds, as illustrated in Table 8.10: while inanimates have a fixed gender, personal nouns have common gender, or at least readily acquire it if the social conditions are such that both male and female referents are available. The examples in (3) provide empirical support for these claims.

(3) a. « Bonjour, je suis le **sage-femme** qui va vous prendre en charge. ». (jactiv.ouest-france.fr, 8 March 2015, D. Le Normand)
   'Hello, I am the$_M$ **midwife** who is going to take care of you.'

Table 8.10 Other examples of gender in compounds

Inanimate	
bloc$_M$ 'block', moteur$_M$ 'engine'	bloc-moteur$_M$ 'engine block'
chou$_M$ 'cabbage', fleur$_F$ 'flower'	chou-fleur$_M$ 'cauliflower'
pause$_F$ 'break', café$_M$ 'coffee'	pause-café$_F$ 'coffee break'
Personal nouns	
sage 'wise', femme$_F$ 'woman'	sage-femme$_{M/F}$ 'midwife'
sans 'without', papiers$_M$ 'papers'	sans-papiers$_{M/F}$ 'illegal immigrant'
sans 'without', abri$_M$ 'shelter'	sans-abri$_{M/F}$ 'homeless person'
faire 'to make', valoir 'to be worth'	faire-valoir$_{M/F}$ 'stooge'

Table 8.11 Examples of human-denoting neoclassical compounds

-logue	zoo- 'animal'	zoo-logue$_{M/F}$ 'zoologist'
-pathe	psycho- 'psyche'	psycho-pathe$_{M/F}$ 'psychopath'
-phile	haltéro- 'weight'	haltéro-phile$_{M/F}$ 'weight lifter'
-vore	herbi- 'grass'	herbi-vore$_{M/F}$ 'herbivore'
-mane	mytho- 'myth'	mytho-mane$_{M/F}$ 'mythomaniac'
-morphe	poly- 'several'	poly-morphe$_{M/F}$ 'polymorph'

    b. Avec son film, Martin Provost voulait rendre hommage à sa manière à la **sage femme** qui lui sauvé la vie à la naissance. (www.allocine.fr, 8 March 2015, D. Le Normand)
        'Through this movie, Martin Provost wished to pay a tribute of his own to the$_F$ **midwife** who saved his life at birth.'

We have again a very similar situation with neoclassical compounding. The gender of a neoclassical compound is determined by the head (second element), and it is obviously fixed for inanimates (e.g. *démocratie*$_F$ 'democracy', *théocratie*$_F$ 'theocracy', etc. vs *homicide*$_M$ 'homicide', *génocide*$_M$ 'genocide', etc.). However, animate neoclassicals are systematically of common gender, as illustrated in Table 8.11.

8.2.2.3 Truncation

We now turn to non-concatenative processes. French very commonly uses clipped nouns to form colloquial new nouns that may be more or less synonymous with their base (Kerleroux 2004).[10] Where the base form denotes an inanimate, the clipped form inherits the gender of its base (e.g. *manifestation*$_F$ 'demonstration' > *manif*$_F$, *vélocipède*$_M$ 'bicycle' > *vélo*$_M$). However, where the base form denotes a human, the clipped form normally has common gender. Table 8.12 illustrates the patterns, and (4) provides attestations.

(4)   a. Ce n'est que le lendemain, lorsque vous vous rendrez compte que vous êtes en présence d'un authentique **beauf**, que vous réaliserez votre erreur. (www.demotivateur.fr, 6 September 2016, N. Weber)
        'It is only on the following day, that you will realise that you are with an$_M$ authentic **dork**, and understand your mistake.'
    b. Qu'est-ce qu'il insinuait? Qu'elle avait l'air d'une **beauf** qui avait gagné au loto? (*Ma vie, mon ex et autres calamités*, 2014, M. Vareille, p. 124)
        'What was he hinting at? That she looked like a$_F$ **dork** who just won the lottery?'
(5)   a. Ou, plus vraisemblablement, comme un clando se glisse jusqu'à un zodiac râpeux alors que des mouettes couinent dans le noir. (*Jours tranquilles d'un prof de banlieue*, 2011, M. Quenehen, p. 51)
        'Or, more plausibly, like an$_M$ **illegal** floats towards an old dinghy while seagulls squeal in the night.'

Table 8.12 Examples of personal clipped nouns

Simple truncations	
beau-frère$_M$ 'brother-in-law'	beauf$_{M/F}$ 'dork'
chef-opérateur$_M$/-trice$_F$ 'chief cameraman'	chef-op$_{M/F}$ 'chief cameraperson'
documentaliste$_{M/F}$ 'school librarian'	doc$_{M/F}$ 'school librarian'
indicateur$_M$ 'personal informant'	indic$_{M/F}$ 'personal informant'
instituteur$_M$/-trice$_F$ 'teacher'	instit$_{M/F}$ 'teacher'
prématuré$_M$/-ée$_F$ 'premature baby'	préma$_{M/F}$ 'premature baby'
professeur$_M$ 'professor'	prof$_{M/F}$ 'professor'
sous-officier$_M$ 'non-commissioned officer'	sous-off$_{M/F}$ 'NCO'
quinquagénaire$_{M/F}$ 'fifty year old'	quinqua$_{M/F}$ 'fifty year old'

Affixed truncations	
clandestin$_M$/-ine$_F$ 'illegal immigrant'	clando$_{M/F}$ 'illegal immigrant'
propriétaire$_{M/F}$ 'owner'	proprio$_{M/F}$ 'owner'
anglais$_M$/-aise$_F$ 'English'	angliche$_{M/F}$ 'English'
bolchévique$_{M/F}$ 'bolchevik'	bolcho$_{M/F}$ 'bolchevik'

b. Elle, une bombasse mais elle parle comme une **clando**. Lui, il est dégueulasse mais il roule en lambo. (*On peut pas tout avoir*, 13 December 2010, Rohff)
'Her, a hottie but she speaks like a$_F$ **clandestine**. Him, he's disgusting but he drives a Lamborghini.'

Note that, interestingly, the gender status of clipped personal nouns does not have to match that of its base. *Instituteur/institutrice* or *indicateur/indicatrices* clearly are gender-iconic pairs with distinct masculine and feminine forms, but the clipped forms coincide, giving rise to a common gender noun. On the other hand, *beau-frère* is indisputably masculine because of its lexical meaning, but readily acquires common gender as a clipped noun with a shifted meaning. Of particular interest is the case of *professeur*, which, despite belonging to the family of agent nouns in *-eur* based on a Latinate stem, lacks a matching feminine, as do all similar nouns based on stems in *-s* (Bonami and Boyé 2006). Probably because of the prestige and conservatism associated with the professorial function, *professeur* resisted becoming a common gender noun for decades after female teachers had become frequent.[11] In the meantime, *prof* has been commonly used in both genders at least since the 1950s.[12]

8.2.2.4 Acronyms

We now turn to acronyms, which reveal a very similar picture. Nominal acronyms referring to inanimate entities have fixed gender, usually inherited from the head of the source phrase (e.g. *confédération$_F$ générale des travailleurs* 'general confederation of workers' > *CGT$_F$*, name of a trade union). However, when an acronym refers to a human, whether the source phrase already had human reference or the use of the acronym has somehow shifted from inanimate to human reference, the result is

Table 8.13 Gender of acronym personal nouns

Bon Chic Bon Genre 'posh' good chic, good style	BCBG$_{M/F}$ 'posh person'
X 'École Polytechnique'	X$_{M/F}$ 'an X graduate'
Sans Domicile Fixe 'homeless' without a fixed address	SDF$_{M/F}$ 'homeless person'
Compagnie Républicaine de Sécurité 'riot squad' national squad of security	CRS$_{M/F}$ 'member of the CRS'
Very Important Person 'V.I.P.'	VIP$_{M/F}$ 'V.I.P.'
Vice-Président 'vice-president'	VP$_{M/F}$ 'V.P.'
Directeur Général 'C.E.O.'	DG$_{M/F}$ 'C.E.O.'

readily of common gender. Table 8.13 provides a few examples, and some attestations are given in (6).

(6) a. Le comédien est hilarant dans ce teen-movie potache, sorte de parcours initiatique à l'envers, où un **BCBG** coincé va progresser . . . en régressant. (*Côté Ciné*, 6 December 2013, T. Séguéla, p. 14)
'The actor is hilarious in this farcical teen-movie, kind of a reverse spiritual journey, where a$_M$ rigidᴍ **posh_guy** will evolve . . . by regressing.'

   b. Imposant régulièrement l'image d'une **BCBG** un peu coincée, sa popularité grandit à l'orée des années 1980 avec « Les hommes préfèrent les grosses » et « Y a-t-il un Français dans la salle ? ». (fr.wikipedia.org, 21 May 2008, Alexdarkchild)
'Promoting on a regular basis the image of a$_F$ rather stiff$_F$ **posh_woman**, her fame started to grow in the 80s with « Les hommes préfèrent les grosses » and « Y a-t-il un Français dans la salle ? ».'

8.2.2.5 Borrowing

We end this discussion by considering gender assignment in borrowed nouns. While the situation is complex and many such nouns initially had a single gender, current usage readily uses them in both genders. Whether a gender distinction exists in the source language is mostly immaterial. Table 8.14 provides some examples, the use of which is documented in (7) and (8).

(7) a. Lorsque j'ai empoigné les ciseaux et commencé à raccourcir les cheveux de mon **cobaye**, un attroupement s'est formé. (*Champion!*, 2015, R. Poulidor, p. 14)
'When I grabbed the scissors and started to shorten my$_M$ **guinea_pig**'s hair, a crowd gathered.'

   b. Tu as de la chance d'avoir trouvé une **cobaye** qui adore les maths! (*La vie commence demain*, 2017, C. Sébillon, p. 49)
'You are lucky to have found a$_F$ **guinea_pig** with a taste for maths!'

(8) a. Tu vois, finalement, je l'ai trouvé mon **baby-sitter** confident, sauf qu'on l'appelle un psy. (*L'irrésistible confident*, 2015, E. Peille, p. 90)

Table 8.14 Borrowed personal nouns

Example	Translation	Source language
cobaye$_{M/F}$	'guinea pig'	Portuguese
cosaque$_{M/F}$	'Cossack'	Russian
minus$_{M/F}$	'moron'	Latin
yankee$_{M/F}$	'American'	English
baby-sitter$_{M/F}$	'babysitter'	English
nabab$_{M/F}$	'nabob'	Urdu
clebs$_{M/F}$	'dog'	Arabic
toubib$_{M/F}$	'physician'	Arabic
soprano$_{M/F}$	'soprano'	Italian

'You see, in the end, I found him, my$_M$ **babysitter** and confidant, only he's called a shrink.'

b Je me rappelle qu'on m'avait confié quelques jours aux Fulconis, les parents de ma **baby-sitter**. (*Comment tu parles de ton père*, 2016, J. Sfar)

'I remember being left for a few days with the Fulconis, my$_F$ **babysitter**'s parents.'

8.2.3 Interim Conclusion

In this section we examined the place of common gender and gender-iconic pairs in the French nominal system, both from the point of view of their prevalence and classification in the extant lexicon, and from the point of view of the resources of the lexeme formation system. Three overall conclusions emerge from that exploration.

First, the prevalence of common gender nouns and gender-iconic pairs in the French system is high and on the rise. From dictionaries documenting conservative usage about 50 years ago, it can be ascertained beyond doubt that more than half of personal nouns either are common gender nouns or belong to a gender-iconic pair. All relevant evidence points to the conclusion that this proportion has risen starkly in the ensuing years.

Second, the common strategies for generating new personal nouns are either common gender (identical forms in the masculine and feminine) or paired gender-specific derivational affixes. The strategy of deriving a feminine personal noun from a masculine personal noun, common as it may be in other languages, basically died out in the middle of the nineteenth century.

Third and most importantly, common gender has the status of a default strategy. Wherever the lexeme formation system does not provide a systematic way of deriving parallel masculine and feminine forms through paired affixes (gender-ambiguous suffixes, compounding, clipping, acronyms, borrowing), newly coined personal nouns acquire common gender, and previously gender-specific nouns tend to shift in that same direction.

Meaning: 'male kid' Gender: M Paradigm:		Meaning: 'female kid' Gender: F Paradigm:		Meaning: 'kid' Gender: — Paradigm:		Meaning: 'kid' Paradigm	
						M	F
SG	gɔs	SG	gɔs	SG	gɔs	gɔs	gɔs
PL	gɔs(z)	PL	gɔs(z)	PL	gɔs(z)	gɔs(z)	gɔs(z)
View (9a)				View (9b)		View (9c)	

Figure 8.2 Schematic lexical entries for *gosse* 'kid' under three views of common gender nouns

In the next section we draw theoretical consequences of these observations for a proper analysis of the French gender system.

8.3 The Shape of French Nominal Paradigms

Let us now reflect on the consequences of our findings for the structure of French nominal paradigms. We start by examining the status of common gender nouns, and then move on to gender-iconic pairs.

8.3.1 Common Gender and Gender Specificity

There are three possible ways of conceiving of the status of common gender nouns in the lexicon. Figure 8.2 illustrates the status of the common gender noun *gosse* 'kid' under these three views.

(9)   a. A maximally conservative view holds that all nouns have to be gender specific by definition, and hence that the intuition that a common gender noun is a single noun with two genders is illusory: we are really dealing with two homophonous nouns differing in grammatical gender.
   b. A first alternative maintains paradigm uniformity but abandons universal gender specificity. Under this view, a French common gender noun has only two paradigm cells, but is underspecified for gender.
   c. A second alternative assumes that common gender nouns have separate paradigm cells for masculine and feminine, but that the morphology happens to syncretise the realisation of masculine and feminine forms. This view abandons both paradigm uniformity and universal gender specificity.

The evidence presented in §8.2 provides a strong argument against view (9a). If homophonous masculine and feminine nouns correspond to different lexemes, given their high prevalence and the productive introduction of new such pairs in the lexicon, they have to be morphologically related. Thus, under this view one would have to posit a highly productive process of gender-changing conversion for personal nouns. The prediction is then that this conversion should be able to apply to any

personal noun matching its input: in other words, there should be derivations such as *laveur*$_M$ 'male washer' > *laveur*$_F$ 'female washer', or, conversely, *laveuse*$_F$ 'female washer' > *laveuse*$_M$ 'male washer'. But such derivations are clearly unattested: wherever the grammar provides an alternate-gender strategy, the use of that strategy is mandatory – common gender noun formation only occurs when that is not possible. There is no way to capture this generalisation under view (9a), short of enumerating as exceptions to the application of gender conversion bases formed with exactly these affixes which happen to be paired with an alternate-gender affix.

It is clear then that universal gender specificity cannot be maintained in the context of productive common gender noun formation: one must admit that, while some nouns, including almost all inanimates, are gender specific, many nouns are compatible with both genders.[13] It is worth emphasising that, at least in French, these nouns are too numerous to be treated as individually listed lexical exceptions: while productive common gender nouns may be a sign of non-canonical gender (Corbett and Fedden 2016), they have to be taken at face value as a systematic pattern.

We are thus left with alternatives (9b) and (9c): either common gender nouns are simply underspecified for gender, or they have separate paradigm cells for masculine and feminine. Which of these two possibilities is preferable depends on the status of gender-iconic pairs in the language, to which we turn now.

8.3.2 Gender-Iconic Pairs and Paradigm Uniformity

As stated in the introduction, there are two ways one may see the relation between the masculine and feminine words in a gender-iconic pair. The status of the gender-iconic pair *avocat*, *avocate* under each view is outlined in Figure 8.3.

(10) a. One may see the two nouns as being derivationally related: we are dealing with two lexemes belonging to the same derivational family. Under such a view, paradigm uniformity is maintained: each noun has a single gender and all nouns may have the same number of paradigm cells, whether or not they belong to a gender-iconic pair.

Meaning: 'male lawyer'		Meaning: 'female lawyer'	
Gender: M		Gender: F	
Paradigm:		Paradigm:	
SG	avoka	SG	avokat
PL	avoka(z)	PL	avokat(z)

Meaning: 'lawyer'
Paradigm:

	M	F
SG	avoka	avokat
PL	avoka(z)	avokat(z)

View (10a)   View (10b)

Figure 8.3 Schematic lexical entries for *avocat/avocate* 'lawyer' under two views of gender-iconic pairs

    b. One may see the two nouns as inflectionally related. Gender-iconic pairs constitute a single lexeme, with separate paradigm cells for masculine and feminine. This view abandons paradigm uniformity.

Which view is correct clearly depends on the language.[14] In languages (or language states) where feminine personal nouns are productively derived from masculine personal nouns, (10a) is clearly warranted. As 8.2 showed in detail, this is not so at all in French.

8.3.2.1 Gender-iconic pairs as derivationally related

Let us consider the consequences of view (10a) for an analysis of the French system. As in the case of common gender nouns, if gender-iconic pairs are separate lexemes, then they must somehow be related by the lexeme formation system. There are two possibilities here: either the two gender-iconic nouns derive from one another, or they are both derived from other sources.

We start with the first possibility, and consider the consequences of taking the two purported lexemes in a gender-iconic pair to be derived from one another. A cursory look at Table 8.5 reminds us that the processes under consideration would include final consonant truncation, substitutions of suffixes, and various marginal operations. View (10a) then leads to three problems. First, one is forced to postulate highly non-canonical processes in what seems to be an otherwise very simple system. Second, it is a mystery why each process is restricted in the way it is; for instance, why can the process of final consonant truncation not apply to feminine nouns in *-rice*? Third and finally, one is left with no account for the observation that the set of derivational alternations found with gender-iconic pairs coincides with the set of inflectional alternations found with adjectives.

Having refuted the idea that gender-iconic pairs could be related directly by a lexeme formation process, we turn to the alternative possibility that they be pairs of lexemes both deriving from a third lexeme. The simplest instantiation of this idea assumes that they are derived in parallel from their immediate base: thus *joueur*$_M$ 'male player' and *joueuse*$_F$ 'female player' would both be derived from the verb *jouer* 'play' by separate processes outputting masculine and feminine agent nouns respectively. Likewise, *épicier*$_M$ 'male grocer' and *épicière*$_F$ 'female grocer' would be derived separately from *épice* 'spice'. While this certainly initially seems to be a reasonable view, it is hard to reconcile with the existence of synonymous processes. To see this, consider again agent nouns. There are four sequences that may be added to a verb's basic stem to form an agent noun: two in the masculine, *-eur* and *-ateur*, and two in the feminine, *-euse* and *-atrice*.[15] All four strategies can be applied to the same bases, as is attested by the existence of quadruplets such as *exporteur*$_M$, *exportateur*$_M$, *exporteuse*$_F$, *exportatrice*$_F$ 'exporter'. The strategies seem to be essentially in free variation. If masculine and feminine agent nouns were derived independently of each other, we would then expect to find situations where the lexicon contains an *-eur* noun and an *-atrice* noun but no *-ateur* or *-euse*, or, equivalently, where it contains an

-*ateur* noun and an -*euse* noun but no -*eur* or -*atrice*; i.e. any combination of a masculine and a feminine agent noun formed on the same base should be able to constitute a gender-iconic pair. But, strikingly, such a situation does not arise. In fact, the French version of Wiktionary, by far the largest and most permissive lexicographic source for French, documents no such case.[16]

It seems clear then that the idea that gender-iconic pairs are derived by parallel suffixal derivation processes is not defensible. However, this is not the only way that the masculine and feminine nouns could be related derivationally. In the context of an analysis of French adjectives in -*eur*, Bonami and Boyé (2005) present an alternative. Building on previous work by Corbin and Corbin (1991), we proposed that many French personal nouns are derived from adjectives by conversion: from the adjective *directeur* 'guiding' with feminine form *directrice*, we hypothesised two separate conversion operations leading to the two personal nouns *directeur*$_M$ 'male director' and *directrice*$_F$ 'female director'. Whatever the merits of such an analysis, it does not scale up to the full set of gender-iconic pairs, as there are whole classes of gender-iconic pairs whose morphological family does not contain a corresponding adjective. This is the case, for instance, for nouns denoting sport practitioners. These are productively formed by suffixing -*eur*$_M$ and -*euse*$_F$ to the name of the sport, e.g. *basket* 'basketball' > *basketteur*$_M$ 'male basketball player', *basketteuse*$_F$ 'female basketball player'.

8.3.2.2 Gender-iconic pairs as inflectionally related

To sum up, we have seen that there is no clear way of maintaining the idea that gender-iconic pairs consist of two lexemes related by lexeme formation, directly or indirectly. It thus becomes clear that view (10a) maintains paradigm uniformity at the expense of completely failing to capture the shape of the French morphological system. We thus propose to adopt the alternative view (10b), and take gender-iconic pairs in French to be inflectionally rather than derivationally related. Under such a view, lexeme formation processes outputting gender-iconic pairs provide strategies for the derivation of both a masculine and a feminine stem, used concurrently for the formation of masculine and feminine forms. Hence, we have one rule forming agent nouns with stems in -*œʁ* and -*œz* from a verb's main stem, and a separate rule forming agent nouns with stems in -*œʁ* and -*ʁis* from a verb's Latinate stem; a rule forming nouns with stems in -*je* and -*jɛʁ* from a noun; etc.

Positing lexeme formation processes outputting multiple stems for the same lexeme is unusual, but, we argue, independently motivated. Bonami and Boyé (2005, 2006) show that such processes are necessary to account for productive stem allomorphy in the inflection of French adjectives – a situation where positing distinct lexemes is not an option. Hence, we can account directly both for the fact that masculine and feminine affixes come in pairs, and for the parallelism between the paradigms of gender-iconic pairs and those of adjectives.

Finally, let us return to the relationship between common gender nouns and gender-iconic pairs. At the end of §8.3.1 we left open whether it was more adequate

for common gender nouns to have a smaller paradigm with gender underspecification or a larger paradigm with syncretism between masculine and feminine forms. Now that we have concluded that gender-iconic pairs are inflectionally related, and hence that paradigm uniformity across all French nouns is not attainable, we have good reasons to opt for option (9c), which amounts to saying that common gender nouns are a special case of gender-iconic pairs with heavy syncretism. There are two immediate motivations for this. First, this makes good sense of the fact that new personal nouns default to common gender. If, by hypothesis, new personal nouns are normally gender-iconic pairs, and non-syncretic gender-iconic pairs involve two stems standing in an allomorphic relation, we expect that the default situation for a new gender-iconic pair is to have no stem allomorphy – hence to use the same form in the masculine and the feminine. Thus, processes that do not provide explicit separate strategies to form two stem alternants will output a single stem and give rise to common gender nouns. Second, this allows for a single locus for iconic gender assignment, which applies to both common gender nouns and gender-iconic pairs: whereas single-gendered nouns may be gender-iconic or not, it is categorically true of both common gender nouns and gender-iconic pairs that their grammatical gender (as manifest in agreement) has to match their social gender.[17] If common gender nouns are a special case of gender-iconic pairs and both are characterised by an expanded paradigm, this paradigm can be stated to be the structure of which the constraint holds.

We thus conclude that common gender nouns are a special case of gender-iconic pairs, and that both instantiate a situation where masculine and feminine nouns form together a single inflectional paradigm. There are thus two types of nouns in French: some have a smaller paradigm and only one gender, while others have a larger paradigm providing distinct cells for both genders.

8.3.3 Variable Content Paradigms, Uniform Form Paradigms

Much work remains to be done before we fully understand the interplay of grammatical gender and semantic gender within the French system. One important concern that needs to be explored in future research is the relationship between derived adjectives and animate and inanimate nouns. It is a general observation that all three types of lexemes tend to rely on the same derivational suffixes, and that these suffixes come in gendered pairs. For instance, *-eur* and *-rice* form in parallel masculine and feminine forms of adjectives (e.g. *moteur/motrice* 'driving', *directeur/directrice* 'guiding'), gender-iconic pairs (e.g. *directeur*$_M$ 'male director', *directrice*$_F$ 'female director') and gender-specific inanimate nouns (e.g. *moteur*$_F$ 'motor', *motrice*$_F$ 'power car'). A promising strategy to capture this parallelism is to rely on the distinction between content and form paradigms (Stump 2006, 2015). The idea is that all French nouns and adjectives share the same basic bidimensional paradigm structure distinguishing two genders and two numbers; lexeme formation processes that output nouns or adjectives need to provide strategies to fill that enlarged paradigm, i.e. to generate paired masculine and feminine forms. The difference between gender-iconic

pairs and other nouns would then stand at the interface between form and content paradigms: by virtue of being specified for gender lexically, non-gender-iconic pairs have a content paradigm that is smaller than their form paradigm by using only forms corresponding to one gender.

One attraction of this analytic scheme is that it is congruent with the current fluidity between gender-specific personal nouns and gender-iconic pairs: many nouns that were historically masculine are becoming gender-iconic pairs under the influence of social change. The system already predicts a feminine form for most of these nouns, by virtue of their morphological make-up. This is exactly what the notion of uniformly gender-variable form paradigms provides: even when a noun has a single gender by lexical stipulation, the forms corresponding to the other gender are readily available if that stipulation is dropped.

8.4 Conclusion

In this paper we have argued against paradigm uniformity across lexemes for French nouns. We started by observing that common gender nouns are too numerous to be treated as lexical exceptions; indeed, they form an open class, and common gender is the default situation for newly coined personal nouns. We then saw that pairs of distinct but morphologically related personal nouns were also very common, and were formed by parallel suffixation, rather than by derivation of one noun from the other. We accounted for both observations by proposing that, in addition to an unsurprising class of single-gendered nouns, French possesses a class of nouns with variable and semantically potent gender. These nouns have a larger paradigm by virtue of accommodating separate masculine and feminine cells.[18]

Although this proposal may go against analytic muscle memory, we submit that it solves more problems than it raises. Non-uniformity of paradigm shape is seldom discussed in the context of nouns, but is an unescapable reality in the conjugation of languages with object agreement; hence, there is no reason it would not also sometimes be found in the nominal domain. In the view we defend, gender is still inherent inflection on nouns, although, for some nouns, it is semantically potent, in the same way that number is. Be that as it may, it can still be defined as a morphosyntactic property of nouns manifest in agreement. Finally, single nominal lexical items with more than one gender value are generally recognised as an existing phenomenon in the case of common gender nouns (we purposefully paraphrase Corbett and Fedden 2016: 507); our contention is only that these are much more frequent than is usually recognised, and that the consequences of that fact for models of morphology should be taken at face value.

Acknowledgements

Agency (ANR) as part of the 'Investissements d'Avenir' programme (reference: ANR-10-LABX-0083).

Notes

1. Although there are also a few inanimate nouns of common gender, as exemplified in Table 8.1, these are not the focus of this paper. In the interest of readability, we take the liberty of referring to only personal nouns with the label 'common gender nouns'.
2. We decided not to provide full glosses of illustrative examples, as only gender marking in one NP is relevant. Rather, we indicate with a subscripted M or F in the translation which relevant words carry explicit gender marking. We also highlight in boldface nouns under consideration.
3. We thank Aurélie Chlebowki for her work on this project.
4. The proportion falls to 27 per cent if one looks at homography rather than homophony. This is due to the fact that French orthography quite often marks an overt gender difference by suffixing an <e> to the feminine that has no phonological reality. All classifications in this paper rely on phonology rather than orthography. An account based on written forms leads to different analyses in individual cases but does not alter the overall phenomenology.
5. As a first step in that direction, we examined whether nouns that are documented as single-gendered in Morphalou were found in co-occurrence with a determiner of the other gender in at least 1 per cent of their occurrences in the FrWac web corpus (Baroni et al. 2009). This was true for 435 purported masculines and 357 purported feminines. We refrain from drawing any strong conclusion from this observation in the absence of human validation.
6. There are other candidate pairs instantiating minor patterns that happen not to be part of the present dataset. One prominent example is the association between *serviteur* 'male servant' and *servante* 'female servant'. The main point holds that such unsystematic associations are strikingly rare.
7. Note that it is not possible, in Modern French, to segment sequences such as *-aine* into a derivational affix *-ain* and a gender suffix *-e*. Such a possibility is only suggested by the misleading orthographic conventions of French: the final orthographic -e does not indicate the realisation of a vowel, but the fact that the preceding orthographic consonant is realised phonetically. The pairs of words in Table 8.8 hence enter alternations that are morphologically principled but not phonologically predictable. See Bonami and Boyé (2005) for a relevant discussion of the parallel facts in adjectival paradigms.
8. At this point it is unclear whether homophonous nouns and adjectives such as *défenseur* 'defender/defending' should be said to be derived in parallel or in a conversion relation, or if that question even makes sense, *pace* Bonami and Boyé (2005).
9. Even the suffix *-esse* is marginally used in adjectival inflection: witness *le bras vengeur* 'the avenging arm' vs *la main vengeresse* 'the avenging hand'.
10. The length of the clipped form is variable, but most commonly two syllables. Sometimes clipping is accompanied by some sort of suffixation, mostly of -o: *propriétaire* 'owner' > *proprio*, *bolchévique* 'bolshevik' > *bolcho*, etc.
11. Such conservative usage is now quickly decaying, and *professeur* has common gender in the spontaneous speech of contemporary university students. In writing both *la professeur* and the homophonous *la professeure* are both commonly used, although the ultra-conservative *Madame le professeur* is still in usage in some circles. Conscious planning is too frequent in this area for one to be able to establish what spontaneous usage is.
12. The Frantext database provides attestations in correspondence between Jean-Paul Sartre and Simone de Beauvoir in 1937.

13. There are a few inanimate nouns found with both genders in the speech of the same speaker with no meaning difference, such as *autoroute* 'highway' and *après-midi* 'afternoon'. These, however, are not numerous enough to be of much consequence on the shape of the system.

14. It is actually quite conceivable that different gender-iconic pairs in the same language receive contrasting analyses. That was probably the case in French at some point, where productive *-esse* coexisted with paired suffixes. The existence of doublets such as *demandeuse* and *demanderesse* as the feminine counterpart of *demandeur* 'requester' testifies to such a state of the system.

15. In the terms used above, this is a consequence of *-eur*/*-rice* being suffixed to a Latinate stem, and the augment *-at* being used by default in the formation of Latinate stems (Bonami et al. 2009).

16. We established this by searching through the GLÀFF lexicon (Hathout et al. 2014), which compiles in tabular form information from the French Wiktionary.

17. We leave aside the proper treatment of hypernymic use of one of the two genders, as when *professeur* is used to refer to either male or female referents. Note that such uses are not limited to gender-iconic pairs (e.g. *homme* 'men' is sometimes used to refer to either men or women, despite being morphologically unrelated to *femme* 'woman'), and are clearly conventionalised (e.g. *garçon* can only refer to boys). Whatever the correct account of such uses, it remains that both common gender nouns and gender-iconic pairs exhibit systematic parallelism between grammatical gender and semantic properties of the referent, which may not exactly coincide with being male or female.

18. We make no claim that such an analysis generalises to other languages, or even to older varieties of French. See, for instance, Rainer (2012) for convincing evidence that German feminine agent nouns are derived from their masculine counterparts.

References

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi and Eros Zanchetta (2009), 'The wacky wide web: A collection of very large linguistically processed web-crawled corpora', *Language Resources and Evaluation*, 43, pp. 209–26.

Bonami, Olivier and Gilles Boyé (2005), 'Construire le paradigme d'un adjective', *Recherches Linguistiques de Vincennes*, 34, pp. 77–98.

Bonami, Olivier and Gilles Boyé (2006), 'Deriving inflectional irregularity', in *Proceedings of the 13th International Conference on HPSG*, Stanford, CA: CSLI Publications, pp. 39–59.

Bonami, Olivier, Gilles Boyé and Françoise Kerleroux (2009), 'L'allomorphie radicale et la relation flexion-construction', in Bernard Fradin, Françoise Kerleroux and Marc Plénat (eds), *Aperçus de morphologie du français*, Saint-Denis: Presses de l'Université de Vincennes, pp. 103–25.

Bonami, Olivier, Gauthier Caron and Clément Plancq (2014), 'Construction d'un lexique flexionnel phonétisé libre du français', in Franck Neveu, Peter Blumenthal, Linda Hriba, Annette Gerstenberg, Judith Meinschaefer and Sophie Prévost (eds), *Actes du quatrième congrès mondial de linguistique française*, pp. 2583–96.

Bousquet, Danielle and Gaëlle Abily (2015), *Guide pratique pour une communication publique sans stéréotype de sexe*, Tech. rep. Haut Conseil à l'Égalité entre les femmes et les hommes.

Corbett, Greville G. (1991), *Gender*, Cambridge: Cambridge University Press.

Corbett, Greville G. (2009), 'Canonical inflectional classes', in Fabio Montermini, Gilles Boyé and Jesse Tseng (eds), *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–11.

Corbett, Greville G. and Sebastian Fedden (2016), 'Canonical gender', *Journal of Linguistics*, 52, pp. 495–531.

Corbin, Danièle and Pierre Corbin (1991), 'Un traitement unifié du suffixe -ier(e)', *Lexique*, 10, pp. 61–145.

Hathout, Nabil, Franck Sajous and Basilio Calderone (2014), 'GLÀFF, a large versatile French lexicon', in Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds), *Proceedings of LREC 2014, Ninth International Conference on Language Resources and Evaluation*, European Language Resources Association, <http://www.lrec-conf.org/proceedings/lrec2014/index.html> (last accessed 26 September 2018).

Kerleroux, Françoise (2004), 'Sur quels objets portent les opérations morphologiques de construction?', *Lexique*, 16, pp. 85–123.

Matthews, P. H. (1991), *Morphology*, 2nd edn, Cambridge: Cambridge University Press.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak and Erez Lieberman Aiden (2010), 'Quantitative analysis of culture using millions of digitized books', *Science*, 14, pp. 176–82.

Nyrop, Kristoffer R. (1936), *Grammaire historique de la langue française, tome troisième formation des mots*, Copenhagen: Nordisk Forlag.

Rainer, Franz (2012), 'Morphological metaphysics: Virtual, potential, and actual words', *Word Structure*, 5 (2), pp. 165–82.

Romary, Laurent, Suzanne Salmont-Alt and Gil Francopoulo (2004), 'Standards going concrete: From LMF to Morphalou', in Michael Zock and Patrick Saint Dizier (eds), *ElectricDict '04 Proceedings of the Workshop on Enhancing and Using Electronic Dictionaries*, Stroudsburg, PA: Association for Computational Linguistics, pp. 22–8.

Stump, Gregory T. (2006), 'Heteroclisis and paradigm linkage', *Language*, 82, pp. 279–322.

Stump, Gregory T. (2015), *Inflectional Paradigms*, Cambridge: Cambridge University Press.

Villoing, Florence (2009), 'Les mots composés VN', in Bernard Fradin, Françoise Kerleroux and Marc Plénat (eds), *Aperçus de morphologie du français*, Saint-Denis: Presses Universitaires de Vincennes, pp. 175–98.

Zwanenburg, Wiecher (1988), 'Flexion et dérivation: le féminin en français', in Ronald Landheer (ed.), *Aspects de linguistique française, hommage à Q. I. M. Mok*, Amsterdam: Rodopi, pp. 191–208.