

**Realistic paradigms for
derivational morphology:
Ethnics and Animals in French**

Gauvain Schalchli & Gilles Boyé
Université Bordeaux-Montaigne & CLLE-ERSS

Outline

- Derivational Paradigms? Realistic?
 - Ethnic Morphological Derivation
 - Animal Lexical Families
- ➔ Realistic Derivational Paradigms

Derivational Paradigms?

- Paradigms are difficult to introduce in derivation because of their strong association with inflection
- But inflectional paradigms have particular properties
 - uniformity
 - productivity
 - popularity
- that might not be as fundamental to paradigmatic organisation as:
 - cell structures
 - semantic relations

Realistic Data

- Following Blevins (2006), we consider that morphological analysis should be based on data accessible to speakers
 - in nature
 - whole words
 - in quantity
 - ecological samples
 - in quality
 - emergent properties and structures
- Here we will focus on emergent structures

Realistic Derivational Paradigms

- Two systems with different structures:
 - ethnics:
 - systematic syncretism
 - uniform paradigm
 - regular semantics
 - animals:
 - occasional syncretism
 - no uniform paradigm
 - variable semantic stratification
- Different results
 - ethnics: inflection-like => classic paradigm
 - animals: not inflection-like => not a classic paradigm but yet

Ethnic Morphological Derivation

Ethnics : a 6-cell paradigm (Roché 2008)

- 6 Cells
 - 3 References
 - Ethnic, Area, Language
 - 2 Categories
 - Noun, Adjective

	Ethnic	Area	Language
N	<i>français</i>	<i>France</i>	<i>français</i>
Adj	<i>français</i>	<i>français</i>	<i>français</i>

	Ethnic	Area	Language
N	<i>russe</i>	<i>Russie</i>	<i>russe</i>
Adj	<i>russe</i>	<i>russe</i>	<i>russe</i>

Different derivation patterns

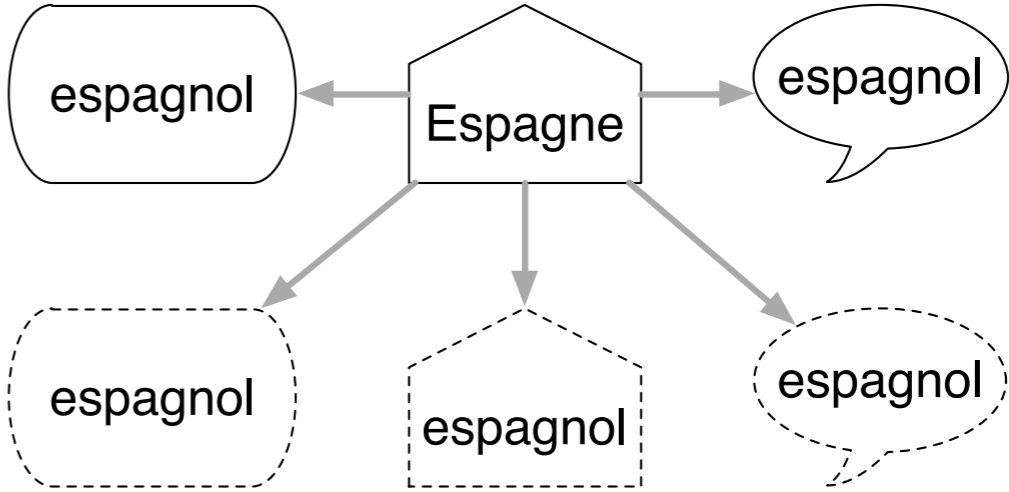
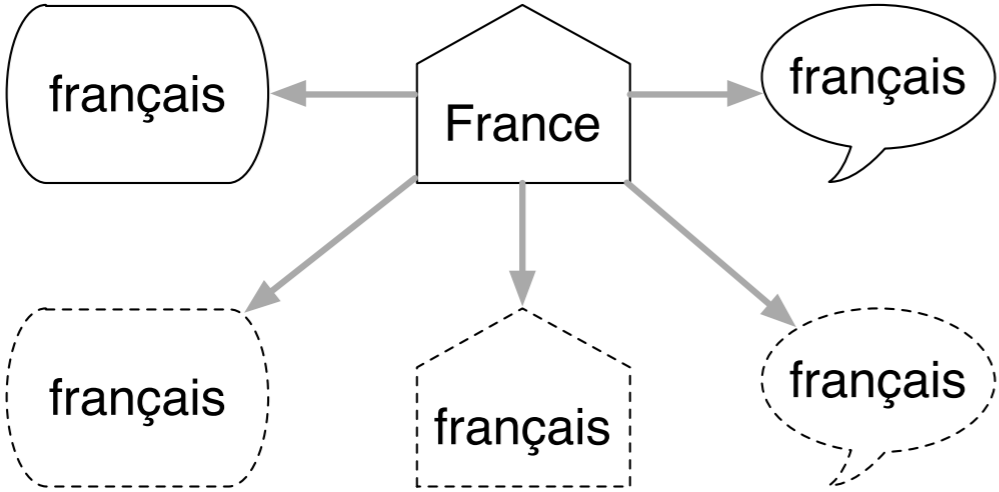
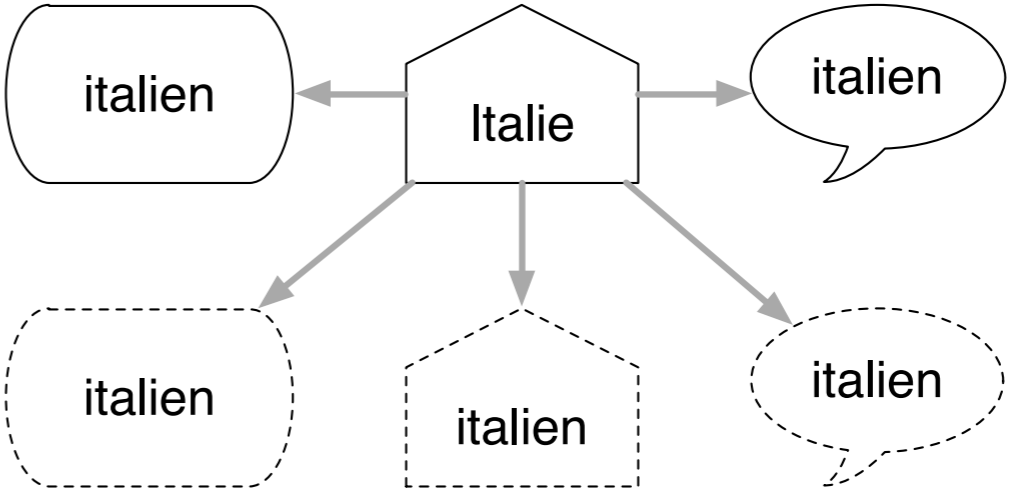
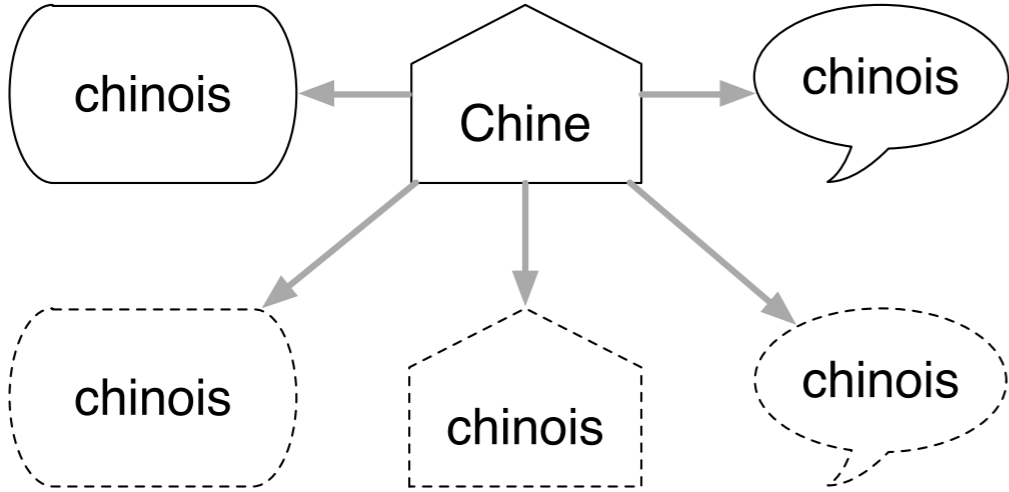
- Area+suffix => Ethnic
 - *France — Français, Italie — Italien, Espagne — Espagnol*
- Ethnic+suffix => Area
 - *Russie — Russe, Belgique — Belge, Tadjikistan — Tadjik*
- Conversion
 - *Suisse — Suisse, Argentine — Argentin*
- Complex relation
 - *Monaco — Monégasque, Nouvelle-Zélande — Néo-zélandais*

But only two forms

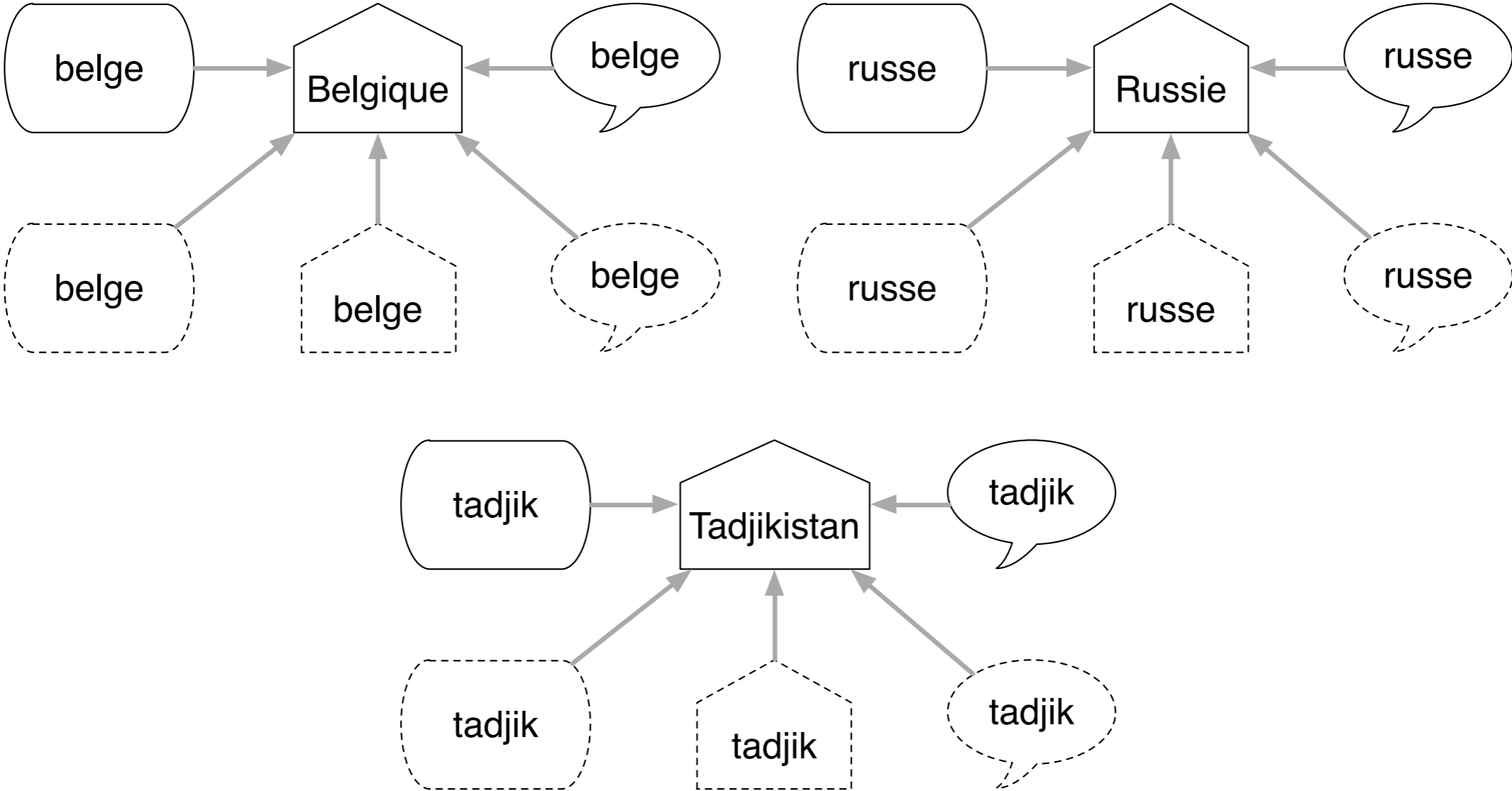
- 1 for the Area N
- 1 for the rest
 - for nouns: Language = Ethnic
 - for adjectives: Ethnic = Area = Language
 - for Ethnic: N = Adj
 - for Language: N = Adj

	Ethnic	Area	Language
N			
Adj			

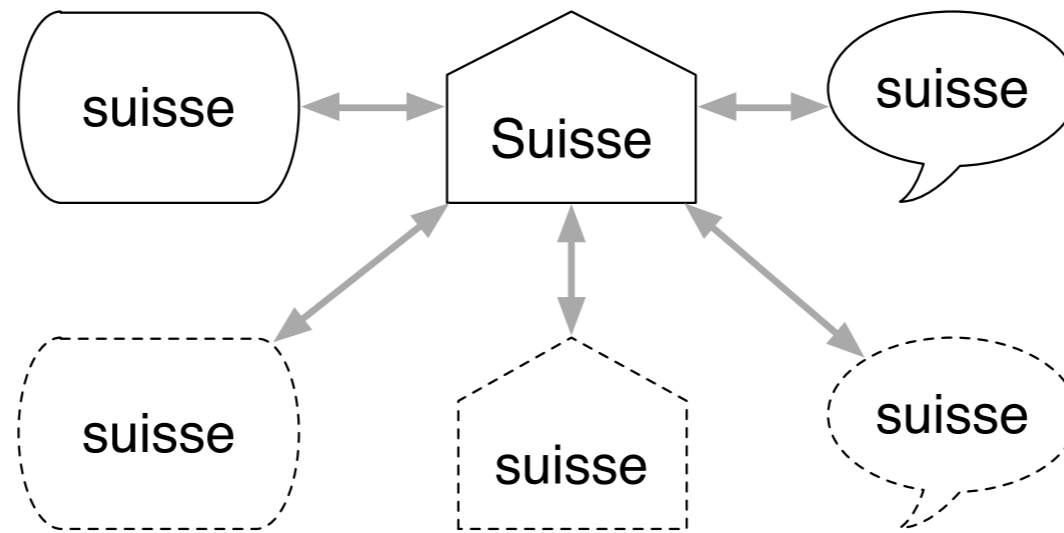
And two filling strategies



And two filling strategies



And two filling strategies



Morphosyntactic and Morphomic Paradigms

- Morphomic paradigms capture systematic syncretism
 - Latin indicative: **morphosyntactic** = morphomic
 - Mauritian: **morphomic**

	INDICATIVE					
	PRESENT	IMPERFECTIVE	FUTURE	PERFECT	PLU-PERFECT	FUTURE PAST
1SG	amo	amabam	amabo	amavi	amaveram	amavero
2SG	amas	amabas	amabis	amavisti	amaveras	amaveris
3SG	amat	amabat	amabit	amavit	amaverat	amaverit
1PL	amamus	amabamus	amabimus	amavimus	amaveramus	amaverimus
2PL	amatis	amabatis	amabitis	amavistis	amaveratis	amaveritis
3PL	amant	amabant	amabunt	amaverunt	amaverant	amaverint

Latin

LF	brize	brije	vãde	amãde	kõsiste	εgziste	fini	vini
SF	briz	brije	van	amãd	kõsiste	εgzis	fini	vin
TRANS.	'break'	'mix'	'sell'	'amend'	'consist'	'exist'	'finish'	'come'

Mauritian

Morphomic paradigms

- English conjugation can be represented with 8 morphomic cells

Sync ₁	[IND, PRS, 1, SG]	<i>set</i>	<i>talk</i>	<i>am</i>
Sync ₂	[IND, PRS, 2, SG], [IND, PRS, 1, PL], [IND, PRS, 2, PL], [IND, PRS, 3, PL]	<i>set</i>	<i>talk</i>	<i>are</i>
Sync ₃	[IND, PRS, 3, SG]	<i>sets</i>	<i>talks</i>	<i>is</i>
Sync ₄	[IND, PST, 1, SG], [IND, PST, 3, SG]	<i>set</i>	<i>talked</i>	<i>was</i>
Sync ₅	[IND, PST, 2, SG], [IND, PST, 1, PL], [IND, PST, 2, PL], [IND, PST, 3, PL], [SUBJ, PST, 1, SG], [SUBJ, PST, 2, SG], [SUBJ, PST, 3, SG], [SUBJ, PST, 1, PL], [IND, PST, 2, PL], [SUBJ, PST, 3, PL]	<i>set</i>	<i>talked</i>	<i>were</i>
Sync ₆	[SUBJ, PRS, 1, SG], [SUBJ, PRS, 2, SG], [SUBJ, PRS, 3, SG], [SUBJ, PRS, 1, PL], [IND, PST, 2, PL], [SUBJ, PRS, 3, PL], [INF]	<i>set</i>	<i>talk</i>	<i>be</i>
Sync ₇	[PART, PRS]	<i>setting</i>	<i>talking</i>	<i>being</i>
Sync ₈	[PART, PST]	<i>set</i>	<i>talked</i>	<i>been</i>

What about our derivations?

- Roché (2008) uses a 6-cell morpho-semantic paradigm
 - six lexemes and two categories
 - but only 2 forms
- We propose a 2-cell morphomic paradigm
 - all the elements in the paradigm do not belong to the same unit
 - but making one cell for Ethnic N/Adj, Area Adj & Language N/Adj makes sense of derivations
 - two cells => two orientations
 - no problem with the origin or circulation

Extending the Ethnic paradigm

- Four 4-cell lexical paradigms:
 - three area — ethnic pairs referring to different levels of conceptualisations
 - geographic area and its inhabitants (*Himalaya — Himalayen*)
 - political area and its citizen (*États-Unis — États-Unien*)
 - ethnic area and its ethnic group (*Tibet — Tibétain*)
 - one language — speaker pair (*Anglais — Anglophone*)
- Similar internal structures
 - same semantic relations between the pairs (Area — Member)
 - same semantic relations inside the pairs (Noun — Adjective)
- Different paradigmatic levels
 - *peuple français, langue française, état français, France métropolitaine, ...*

Paradigm structures

- The three Area — Ethnic paradigms share
 - the same structure
 - an area and a group associated with it
 - two filling strategies
 - the ethnic area derives from the ethnic group
les francs > la francie
 - the geographical/political area is the base for the group
la france > les français
- Language — Speaker paradigm is different
 - the speaker is a neo-classic compound based on the language
le français -> les francophones

Paradigms, paradigms, what paradigms?

- Every Area — Ethnic can combine with a Language — Speaker
 - one paradigm with all four?
 - three paradigms with the combinations?
 - The contrasts between the terms vary
 - depending on the point of view, the same term changes meaning
 - meaning stratification => paradigm stratification
- ➔ One paradigm with several levels of discrimination






Intermediate conclusion

- The ethnic paradigm behave like an inflectional one
 - there is a number of expected forms: cells
 - the semantic relations between cells are constant
- It falls under the scope of the PCFP, Paradigm Cell Filling Problem (Ackerman & al. 2009):
 - what licenses reliable inferences about the inflected and derived surface forms of a lexical item?
 - ➔ paradigmatic organisation
 - ➔ learnable regularities

Animal Lexical Families


Animals: a 3-cell paradigm (Roché 1996)

- 3 Cells
 - 3, 2 or 1 forms
 - syncretism:
 - Species = Male, Species = Female, Species = Male = Female

	Species	Male	Female
	<i>mouton</i>	<i>bélier</i>	<i>brebis</i>
	<i>chien</i>	<i>chien</i>	<i>chienne</i>
	<i>dinde</i>	<i>dindon</i>	<i>dinde</i>
	<i>chèvre</i>	<i>bouc</i>	<i>chèvre</i>
	<i>écureuil</i>	<i>écureuil</i>	<i>écureuil</i>

Extending the animal paradigm

- Roché (2009) includes the young animal names
 - as *pullisemblance* derivatives (Damourette & Pichon, 1930)
- Schalchli (2016) structures both young and adult families with the same structure
 - sex-neutral, male, female

	Sex-neutral	Male	Female
adult	<i>mouton</i>	<i>bélier</i>	<i>brebis</i>
young	<i>agneau</i>	<i>agneau</i>	<i>agnelle</i>

PCFP : some problems

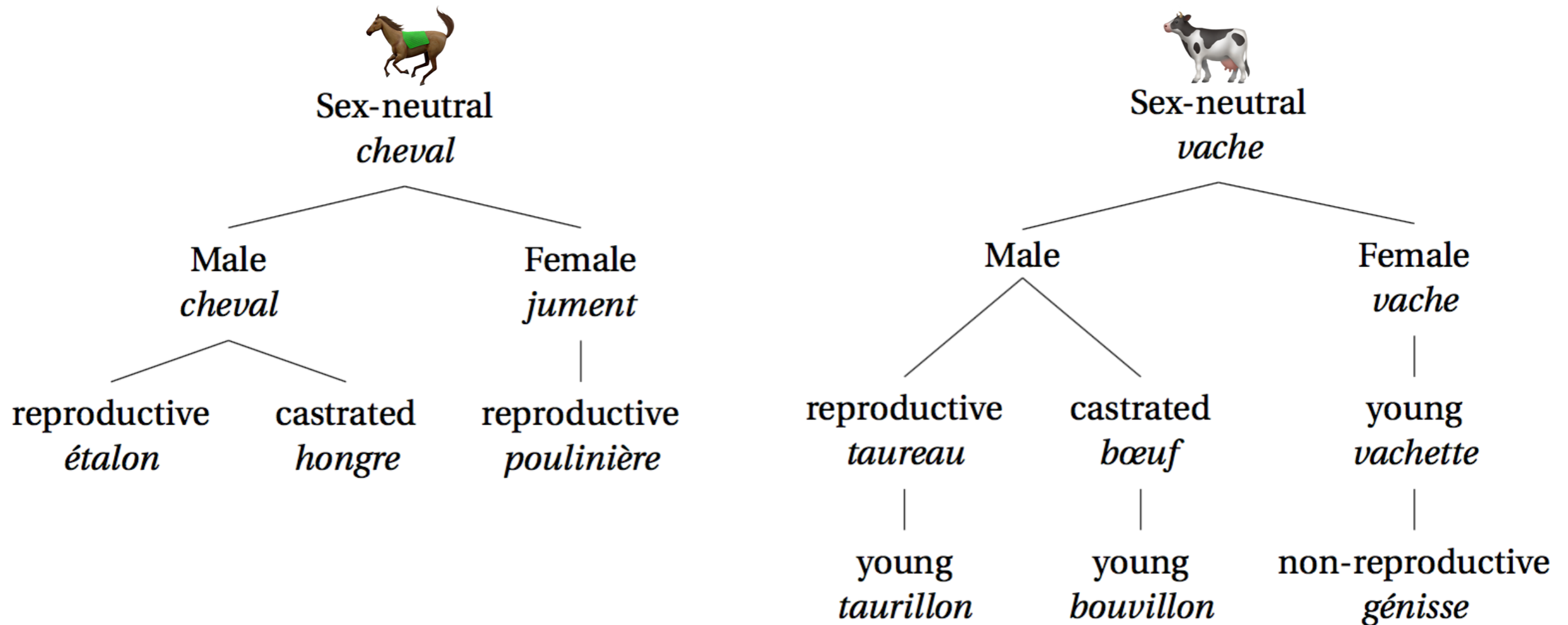
- Answering the PCFP seems more difficult than with the ethnics
 - lots of suppletion : vache/bœuf/taureau/génisse/veau
 - lots of derivations: suffixations/conversions
 - and ambiguities *chaton vs dindon*
 - massive syncretism but not systematic
 - species = male, species = female
- But the main problem lies with the Paradigm Cell Filling Problem itself:
 - what paradigm?

The Paradigm Cell Finding Problem

- In the extended animal family, the structure varies from one species to the next
- varying contrasts:
 - genders 2 or 1: *chien/chienne, mouche, escargot*
 - ages 2 or 3 or 1: *chien/chiot, poule/poulet/poussin, amibe*
 - reproduction 1 or 2: *chien, taureau/bœuf, étalon/hongre*
- varying hierarchy:
 - male reproduction-neutral : 🐎 *cheval*, 🐄 ?*taureau*, ?*bœuf*
- This is the Paradigm Cell Finding Problem
- what licenses reliable inferences about the paradigm of a lexical family?
 - ➔ learnable paradigmatic organisation

Paradigmatic organisation

- The organisation is semantic
 - it relies on an hierarchy of oppositions and morphological relations



Conclusion

- Paradigm organisation could emerge from realistic data
 - stratified morphomic paradigms
 - meanings and oppositions vary depending on the level of discrimination considered
 - inflection syncretism comes in two varieties
 - systematic syncretism => systematic polysemy (arbitrary)
 - ≠ natural syncretism/polysemy (= underspecification)
 - accidental syncretism => homophony, accidental polysemy
- Morphemes can be useful at a lexical level
 - a lexical morpheme can contain more than one classical lexeme