Pseudo-Pāṇinian Splits

Berthold CRYSMANN and Baptiste LOREAU UNGER

CNRS and ENS Paris-Saclay

IMM, WU Wien, 28–30 August 2024

Paradigm splits

- Patterns of syncretism
  - morphology
    - paradigm structure
    - representation of morphological features (e.g. underspecification)
  - morphology–syntax interface
    - correspondence between syntactic features (content paradigm) and morphological features (form paradigm)
    - morphomes support autonomous morphology
- Two extreme points (and a large area in between):
  - natural splits (systematic neutralisation): balanced, parallel zones
  - Pāṇinian (Bonami, 2015): completely nested zones
  - unmotivated splits (irreducibly morphomic)

	A	B
α	x	y
β	x	y
γ	x	y

(a) Natural

	A	B
α	x	x
β	x	x
γ	x	**z**

(b) Pāṇinian

	A	B
α	x	y
β	x	y
γ	**z**	**z**

(c) Pseudo-Pāṇinian

	A	B
α	x	y
β	x	x
γ	y	y

(d) Morphomic

Figure: Some types of splits

Paradigm splits

- Patterns of syncretism
  - morphology
    - paradigm structure
    - representation of morphological features (e.g. underspecification)
  - morphology–syntax interface
    - correspondence between syntactic features (content paradigm) and morphological features (form paradigm)
    - morphomes support autonomous morphology
- Two extreme points (and a large area in between):
  - natural splits (systematic neutralisation): balanced, parallel zones
  - Pāṇinian (Bonami, 2015): completely nested zones
  - unmotivated splits (irreducibly morphomic)

	SG	PL
NOM	Zeit	Zeiten
GEN	Zeit	Zeiten
DAT	Zeit	Zeiten
ACC	Zeit	Zeiten

(a) Natural: German

	SG	PL
NOM	Wagen	Wagen
GEN	**Wagens**	Wagen
DAT	Wagen	Wagen
ACC	Wagen	Wagen

(b) Pāṇinian: German

Pashto,
Limbu,
Koryak,
…

(c) Pseudo-Pāṇinian

Gender	SG	PL
I	v	b
II	y	d
III	y	y
V	d	d
VI	b	d

(d) Morphomic: Batsbi

Figure: Types of splits: examples

Pashto adjectival declension (Tegey & Robson, 1996)

- Pashto adjectives (and nouns) inflect for case, number and gender
- Inflectional features:
  - 2 core cases: direct and oblique
  - 2 numbers
  - 2 genders
- Exponence:
  - Oblique plural consistently marked with *-o*, irrespective of inflection class and gender
  - Unnatural systematic syncretism between OBL.SG and DIR.PL
  - Syncretism may extend to include DIR.SG cell

		MAS	FEM
SG	DIR	tit	tit-a
	OBL	tit	tit-e
PL	DIR	tit	tit-e
	OBL	**tit-o**	**tit-o**

		MAS	FEM
SG	DIR	təg-ay	təg-e
	OBL	təg-i	təg-e
PL	DIR	təg-i	təg-e
	OBL	**təg-o**	**təg-o**

Figure: Declension of *tit* 'scattered' (class I) and *təg-ay* 'thirsty' (class II) (Tegey & Robson, 1996, 75–76)

Pashto adjectival inflection: the challenge

		MAS	FEM
SG	DIR	tit	tit-a
	OBL	tit	tit-e
PL	DIR	tit	tit-e
	OBL	**tit-o**	**tit-o**

		MAS	FEM
SG	DIR	təg-ay	təg-e
	OBL	təg-i	təg-e
PL	DIR	təg-i	təg-e
	OBL	**təg-o**	**təg-o**

Figure: Declension of *tit* (class I) and *təg-ay* (class II) (Tegey & Robson, 1996, 75–76)

- Distribution of syncretic ∅, *-e*, *-i* does not correspond to natural class
- Class I masc and class II feminine look like Pāṇinian splits
  - General OBL.PL *-o* takes a chunk out of the defaults' distribution (∅/*-e*)
- Class I feminine and class II masculine may also be Pāṇinian splits
  - General OBL.PL *-o* takes a chunk out of the default's distribution
  - Class-specific overrides for the DIR.SG cell *-a* and *-ay* get another chunk, leaving "diagonal" pattern
- Problem: Override is too general to function as a Pāṇinian override
  - ✓ *-o* is *more specific* w.r.t. case and number
  - ✗ *-o* is **less specific** w.r.t. gender and inflection class

Pāṇinian vs. pseudo-Pāṇinian splits

- Pāṇinian split (Bonami, 2015)
  - Corresponds to widely assumed principle (dating back to 500 BCE)
    - Lexical Phonology (Elsewhere Condition; Kiparsky, 1985)
    - Paradigm Function Morphology (Stump, 2001)
    - Distributed Morphology (?)
    - Optimality Theory (?)
  - Pāṇini's Principle
    - Overriding cells properly included within default (proper nesting of cells)
    - Featural description of strictly more specific than the default (subsumption)
    - Override has a natural description, but its complement often does not
- Pseudo-Pāṇinian splits
  - combine
    - natural split
    - Pāṇinian split
  - But override is too general
    - typically across splits (yielding π–shape)

$$\begin{bmatrix} \textsc{mud} & \{\textit{obl}, \textit{pl}\} \\ \textsc{mph} & \langle[\textsc{ph}\ \text{o}]\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textsc{mud} & \{\textit{case}, \textit{num}\} \\ \textsc{ms} & \{[\textsc{cls}\ \textit{m3-cls}], \ldots\} \\ \textsc{mph} & \langle[\textsc{ph}\ \text{i}]\rangle \end{bmatrix}$$

- Desideratum: make pseudo-Pāṇinian overrides work like standard Pāṇini
  - make instances more specific to work as an override
  - keep generalisations

Pāṇinian vs. pseudo-Pāṇinian splits

- Pāṇinian split (Bonami, 2015)
  - Corresponds to widely assumed principle (dating back to 500 BCE)
    - Lexical Phonology (Elsewhere Condition; Kiparsky, 1985)
    - Paradigm Function Morphology (Stump, 2001)
    - Distributed Morphology (?)
    - Optimality Theory (?)
  - Pāṇini's Principle
    - Overriding cells properly included within default (proper nesting of cells)
    - Featural description of strictly more specific than the default (subsumption)
    - Override has a natural description, but its complement often does not
- Pseudo-Pāṇinian splits
  - combine
    - natural split
    - Pāṇinian split
  - But override is too general
    - typically across splits (yielding π–shape)

$$\begin{bmatrix} \textsc{mud} & \{\textit{obl}, \textit{pl}\} \\ \textsc{mph} & \langle[\textsc{ph}\ \text{o}]\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textsc{mud} & \{\textit{case}, \textit{num}\} \\ \textsc{ms} & \{[\textsc{cls}\ \textit{f1-cls}], \ldots\} \\ \textsc{mph} & \langle[\textsc{ph}\ \text{e}]\rangle \end{bmatrix}$$

- Desideratum: make pseudo-Pāṇinian overrides work like standard Pāṇini
  - make instances more specific to work as an override
  - keep generalisations

The framework: IbM

- Formal analysis carried out in Information-based Morphology (Crysmann & Bonami, 2016; Crysmann, 2023)
- Inferential-realisational theory of inflection couched in typed feature logic
  - Rule types are organised into multiple inheritance hierarchies
    ⇒ generalisation (about overrides' distributions)
  - Rule instances at the leaves
    ⇒ specificity (to account for competition)
- Many-to-many: rules are pairings of
  - *m* properties to express (MUD) with
  - *n* morphs (MPH) that realise them (phonology & position)
  - Rules can constrain on other morphosyntactic properties (MS) than those realised (MUD)
- Pāṇinian competition resolved globally at the leaves, based on rule informativity (subsumption)

Analysis of Pashto adjectival declension (broken)

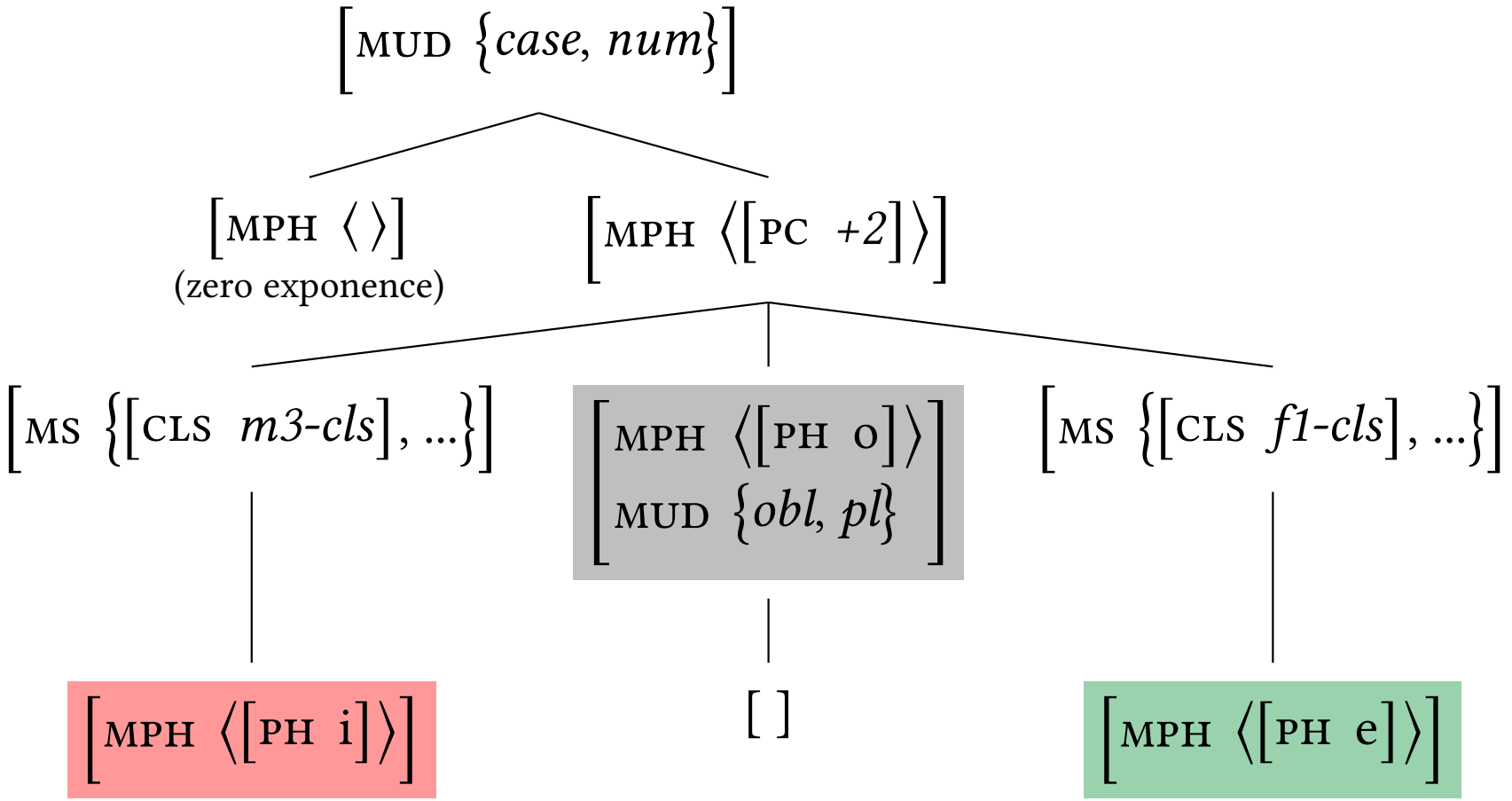

Figure: Rule hierarchy for affixes in Pashto data (broken Pāṇinian strategy)

Analysis of Pashto adjectival declension (correct)

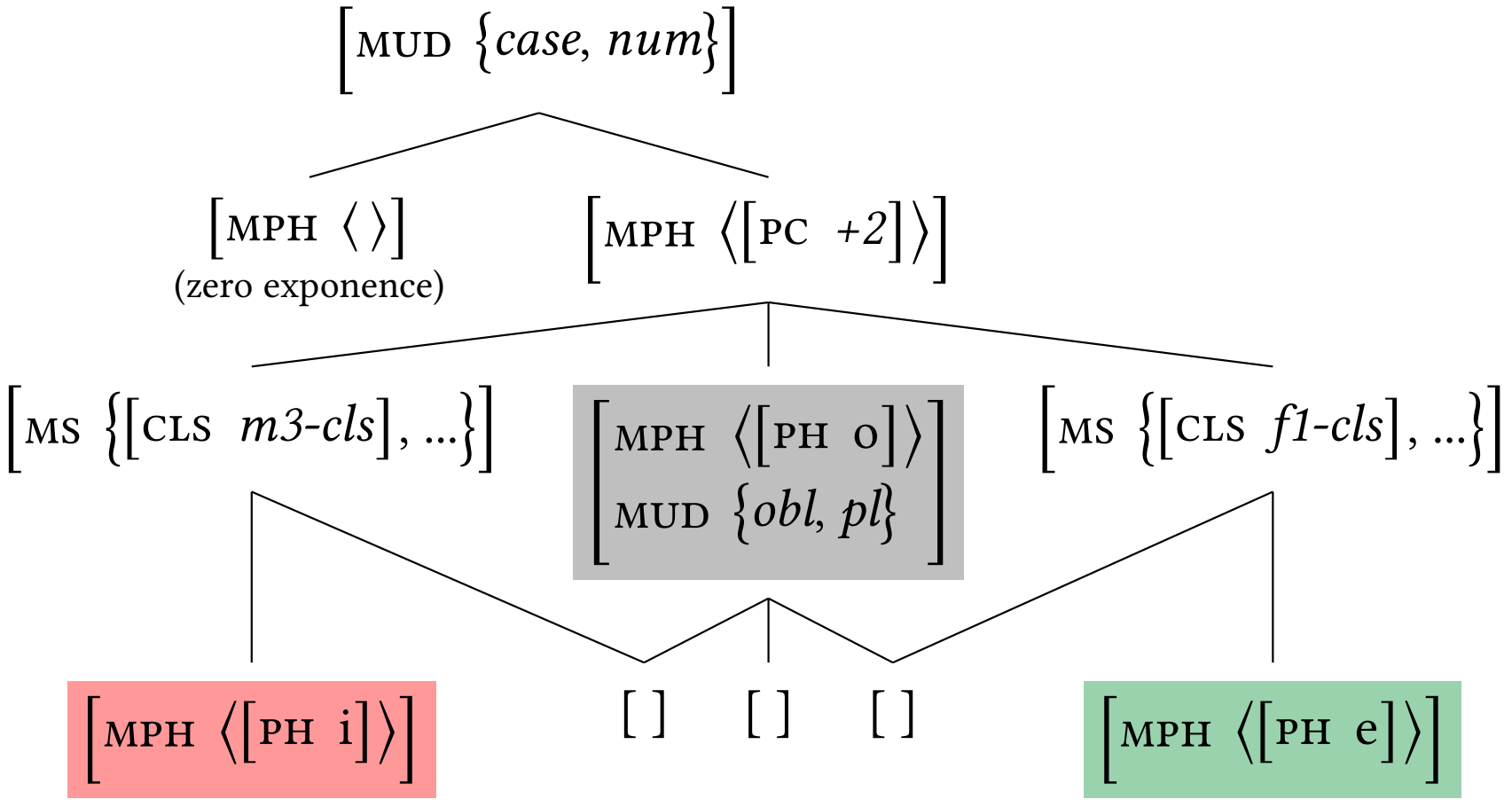

Figure: Rule hierarchy for affixes in Pashto data (pseudo-Pāṇinian strategy)

More pseudo-Pāṇini: Koryak agreement suffixes

- Koryak (Žukova, 1972) uses verbal affixes for S, A and O participants
- 3rd person O cells:
  - inflect for O number: *-n* SG vs. *-net* DU
  - Two exceptions (expressing A person/number instead): 2DU>3 and 3SG>3
- Problem identical to Pashto: *-tkə* and *-nin*
  - ✓ *more specific* w.r.t. A person/number
  - ✗ **less specific** w.r.t. O number

		object						
		∅	1		2		3	
			SG	DU	SG	DU	SG	DU
1	SG	-k	—	—	-gi	-tək	-n	-net
	DU		—	—	-gi	-tək	-n	-net
2	SG	-j	-j	-mək	—	—	-n	-net
	DU	-tək	-tək	-mək	—	—	**-tkə**	**-tkə**
3	SG	-j	-j	-mək	-gi	-tək	**-nin**	**-nin**
	DU	-gəhi	-gəm	-mək	-gi	-tək	-n	-net

Table: Koryak suffixes (Baerman et al., 2005)

Analysis of Koryak third-person object suffixes

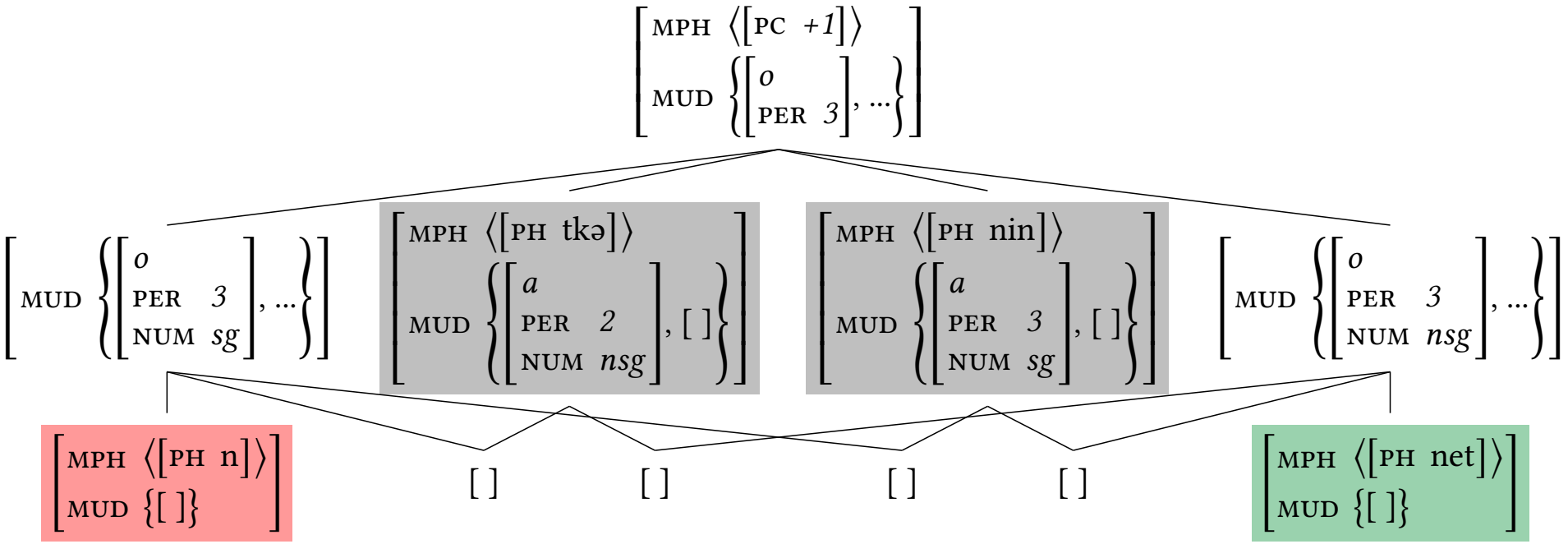

Figure: Koryak third person object suffixes

- Solution analogous to Pashto:
  - cross-classification of each override with featural description of its competitors
  - natural description of each override (purely conjunctive)

More pseudo-Pāṇini: Limbu portmanteau *-m?na*

- Limbu (van Driem, 1987) verbs inflect for S, A, O core participants
- Suffix *-m?na* is a portmanteau that jointly expresses
  - **preterite** tense
  - 1st person **exclusive plural**: S or A>3 (nominative)
- Competitors:
  - **Plural** suffix *-i* (S or O; absolutive)
  - **Plural** suffix *-m* (A)
  - Others: 1st person exclusive *-ge/-be*, 3rd-person O *-u*, preterite *-ɛ*
- Competition: *-m?na*
  - ✓ *more specific* w.r.t. almost all information, but
  - ✗ but **less specific** or **incommensurate** w.r.t. role

Role	A>2	A>3SG	S	3SG>O
1PE.PRET	-n-ɛ-tchi-ge	**-m?na**	**-m?na**	-i-ge
1PE.NPRET	-nɛ-tchi-ge	-u-m-be	-i-ge	-i-ge
2PL.PRET	—	kɛ- -u-m	kɛ- -i	kɛ- -i
2PL.NPRET	—	kɛ- -u-m	kɛ- -i	kɛ- -i
1PI.PRET	—	a- -u-m	a- -ɛ	a- -ɛ
1PI.NPRET	—	a- -u-m	a-	a-

Table: Pseudo-Pāṇinian split in the plural, involving *-m?na*

Analysis of Limbu preterite person marking

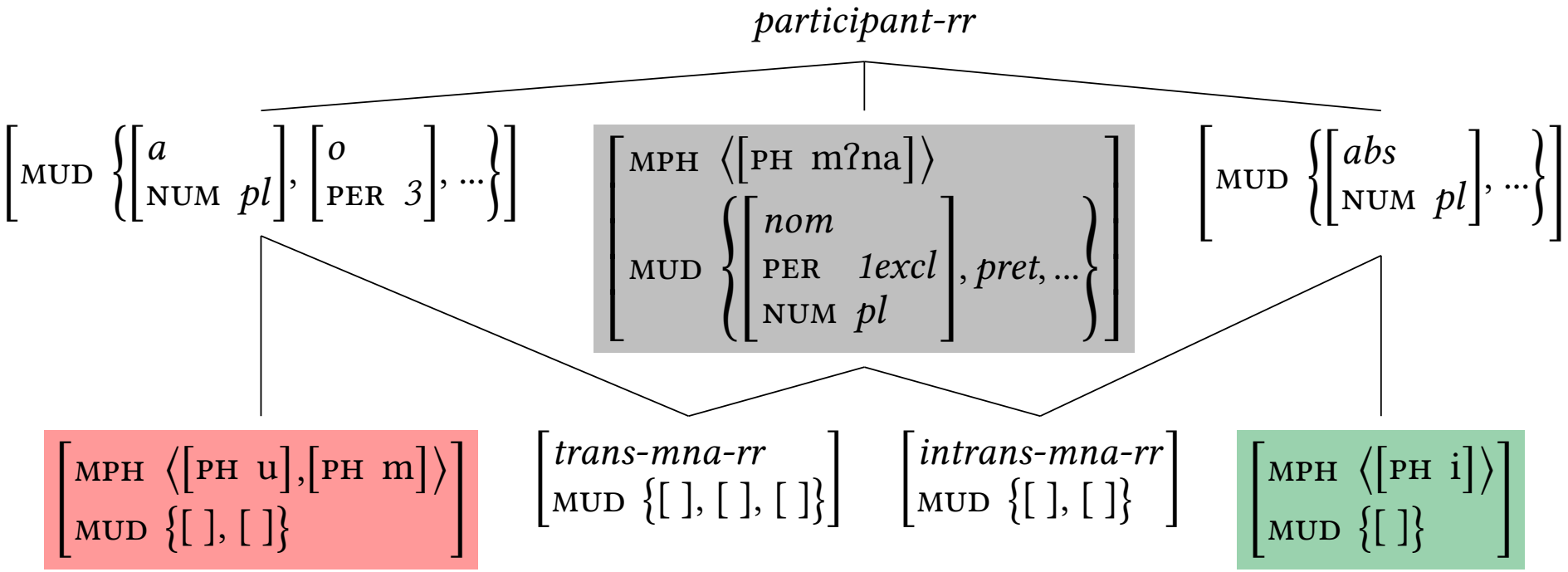

Figure: Rule hierarchy for Limbu person affixes

- Cross-classification of *-mʔna* with its competitors' supertypes
  - adds specificity to *-mʔna*'s role specification (make Pāṇini great again!)
  - fine-tunes the distribution (conjunctive supertype and disjunctive subtypes)

Almost morphomic: Koryak agreement prefixes

- Prefixal *ne-/ine-* functions as an inverse marker
- Distribution of *ine-*:
  - ✓ properly nested within *ne-*: only 1SG.O cells (like Pāṇini)
  - ✗ needs to invoke disjunctive constraints as well (unlike Pāṇini)

			object					
		∅	1		2		3	
			SG	DU	SG	DU	SG	DU
1	SG	tə-	—	—	tə-	tə-	tə-	tə-
	DU	mət-	—	—	mət-	mət-	mət-	mət-
2	SG		**ine-**	ne-	—	—		
	DU		**ine-**	ne-	—	—		
3	SG		**ine-**	ne-	ne-	ne-		
	DU		ne-	ne-	ne-	ne-	ne-	ne-

Table: Koryak prefixes (Baerman et al., 2005)

- Disjunction needed regardless of the atomic/composite analysis of *ine-*
  - *i-* still competes with suffixes/zero

Analysis of Koryak inverse prefixes

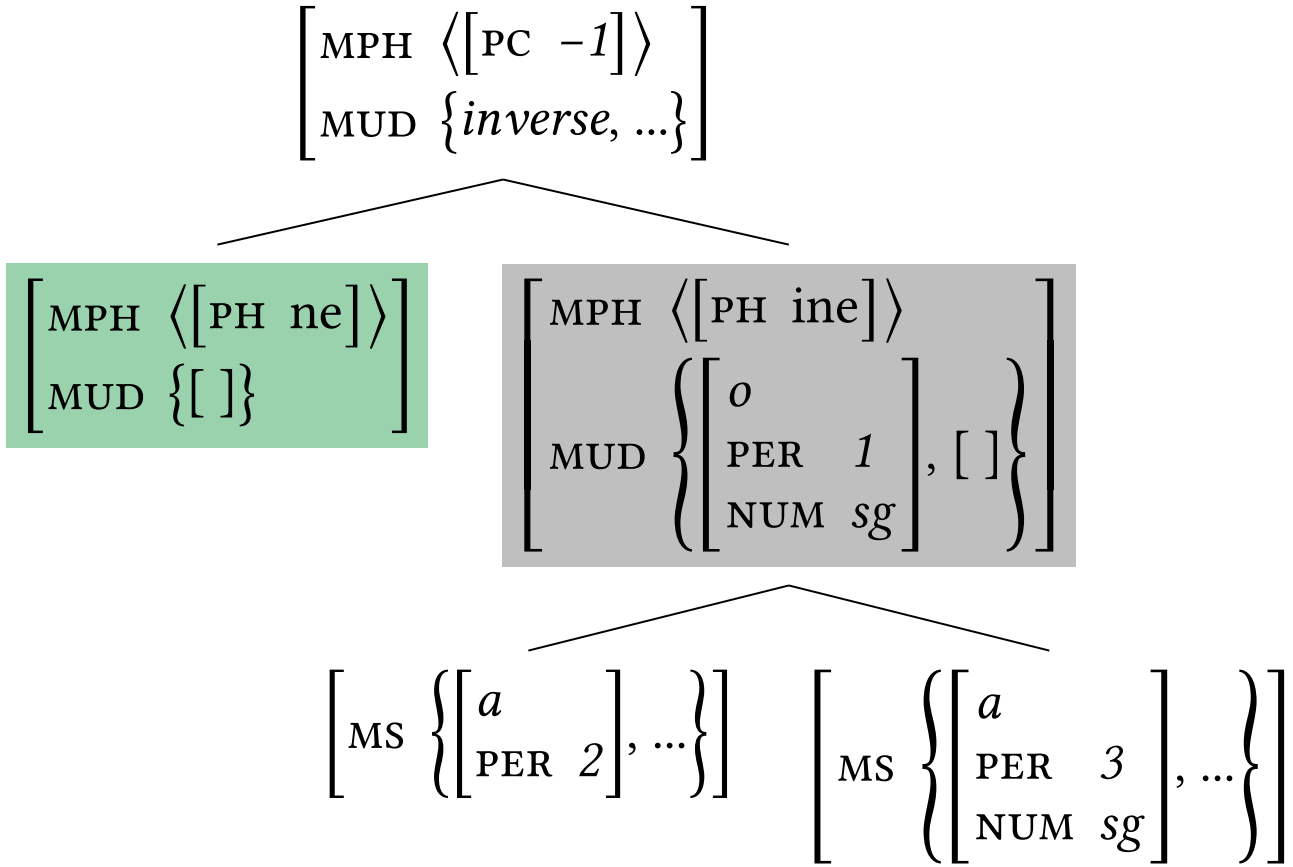

Figure: Koryak inverse prefixes

- Description of override
  - non-trivial generalisation (essentially a superset approximation)
  - disjunctive restriction of distribution (cf. Limbu)

Truly morphomic

- Disjunctive characterisation of override may also be found with classical morphomic split
- However:
  - Koryak *ine-* has a conjunctive generalisation that is properly nested within the default
  - Generalisation of disjunctive L-morphome is co-extensive with the entire split
- Truly morphomic splits lack a non-trivial generalisation of the override

'say'	PRS.IND	PRS.SBJV
1SG	dig-o	dig-a
2SG	diz-es	dig-as
3SG	diz	dig-a
1PL	diz-emos	dig-amos
2PL	diz-eis	dig-ais
3PL	diz-em	dig-an

Table: The L-pattern morphome (Portuguese; Maiden, 2005, 147)

Towards a typology of Pseudo-Pāṇini

- Pseudo-Pāṇinian splits in Pashto, Koryak and Limbu
  - display family resemblance with
    - natural splits
    - true Pāṇinian splits
  - can be reduced to these well-established splits by means of inheritance hierarchies

    "Have your cake and eat it!"
- Extended notion of Pāṇinian splits
  - gives rise to a 2×2 mini-typology, with two dimensions:
    - whether override applies across a split (±split)
    - fully natural vs. partially natural description of override (±natural)
  - can be formally distinguished from truly morphomic splits (which have entirely unnatural descriptions and trivial generalisations)

	−split	+split
+natural	standard Pāṇini	Pashto/Koryak suffixes
−natural	Koryak prefixes	Limbu

Table: Typology of (pseudo-)Pāṇinian splits

  - complements existing typologies of morphomes (?)

References I

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism.* Cambridge: Cambridge University Press.

Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25. 63–110.

Crysmann, Berthold. 2023. Morphotactic competition in Murrinh-Patha: Rule composition and rule interaction in Information-based Morphology. In Stefan Müller & Elodie Winckel (eds.), *Proceedings of the 30th International Conference on Head-Driven Phrase Structure Grammar, University of Massachusetts Amherst*, 27–44. Frankfurt/Main: University Library. doi:10.21248/hpsg.2023.2.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52(2). 311–374. doi:10.1017/S0022226715000018.

van Driem, George. 1987. *Limbu language.* Berlin: Mouton de Guyter.

Harris, Alice C. 2009. Exuberant exponence in Batsbi. *Natural Language and Linguistic Theory* 27. 267–303.

References II

Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2. 83–136.

Maiden, Martin. 2005. Morphological autonomy and diachrony. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2004*, 137–175. Dordrecht: Kluwer.

Stump, Gregory T. 2001. *Inferential–realizational morphology* 1–30. Cambridge Studies in Linguistics Cambridge University Press. doi:10.1017/CBO9780511486333.001.

Tegey, Habibullah & Barbara Robson. 1996. *A reference grammar of Pashto.* Washington, D.C.: Center for Applied Linguistics.

Žukova, Alevtina. 1972. *Grammatika korjakskogo jazyka: fonetika, morfologija.* Leningrad: Nauka.

Uninflectedness and Overabundance in Pashto

		MAS	FEM
SG	DIR	xāystá	xāystá
	OBL	xāystá	xāystá
PL	DIR	xāystá	xāystá
	OBL	xāystá, xāystá- **wo**	xāystá, xāystá- **wo**

Table: Declension of *xāystá* ‘pretty’ (class IV: *a*-final inflectables)

Koryak verbal agreement (with plural cells)

A \ O	3SG		3DU		3PL	
1SG	t-	-n	t-	-net	t-	-new
1DU	mət-	-n	mət-	-net	mət-	-new
1PL	mət-	-la-n	mət-	-la-n	mət-	-la-n
2SG		-n		-net		-new
2NSG		**-tkə**		**-tkə**		-la-**tkə**
3SG		**-nin**		**-nin**		**-nin**
3NSG	ne-	-n	ne-	-net	ne-	-new

Table: Koryak transitive person marking (past). Adapted from Žukova (1972, 252–254).

- Split between *-n* 3SGO, *-net* 3DUO and *-new* 3PLO:
  - extends all over the 3O subparadigm
  - disrupted in scenarios involving a 2NSG, 3SG or 1PL subject:
    - *-tkə* 2NSGA (can combine with (object) plural marker *-la*)
    - *-nin* 3SGA (cannot combine with object *-la*)
    - (*-la* PLA can combine with *-n* 3O, disrupting the object number split)

Turkana verbal marking

		1		2	3
		SG	PL		
1	SG	—	—	**ka-**	a-
	PL	—	—	kɪ-	kɪ-
2		kɪ-	kɪ-	—	ɪ-
3		**ka-**	kɪ-	kɪ-	ɛ-

Table: Turkana (Baerman et al., 2005)

Athpare verbal number marking

O \ A	3SG		3PL	
1SG		-ŋa	u-	-ŋa
1DE		-ci-ŋa	u-	-ci-ŋa
1PE		-i-ŋa	u-	-i-ŋa
1DI	**a-**	-ci	**a-**	-ci
1PI	**a-**		**a-**	
2SG	**ma-**		**ma-**	
2DU	**ma-**	-ci	**ma-**	-ci
2PL	**ma-**	-i	**ma-**	-i
3SG		-u	u-	-u
3NSG		-uci	u-	-uci
∅ (intr.)		∅	u-	

Table: Athpare verbal agreement. Adapted from **?**, 31.

- Split between *u-* 3PLA/S and ∅ 3SGA/S:
  - extends all over the 3A/S subparadigm
  - disrupted in scenarios with a 2 or 1INCL object (collapse of A number):
    - *ma-* 3>2
    - *a-* 1INCL.S/O