Opaque paradigms, transparent forms in Nepali conjugation

Olivier Bonami
Université Paris-Sorbonne
Institut Universitaire de France
LLF (UMR 7110)

&

Gilles Boyé
Université de Bordeaux
CLLE-ERSSàB (UMR 5263)

WOTM5, June 2010

Introduction

- Widespread conception : inflection realizes **morphosyntactic** features — i.e. syntactic or semantic properties.
  - Notable exceptions: inflectional classes are purely **morphological** features (**morphomic** in the sense of Aronoff (1994)).
- Corbett & Baerman (2006): Morphomic features are non-canonical, they should be avoided if possible.
  - For Nepali, we defend the use of morphomic features at the interface between morphosyntactic features and morphological exponents.
    - Exponent paradigms are nicely regular, but related in an opaque fashion to morphosyntactic features.
- This allows for a conservative analysis of apparently difficult data within Paradigm Function Morphology (Stump 2001).

Opaque paradigms in Nepali conjugation

Synthetic conjugation

- 8 synthetic TAM sub-paradigms: present, perfective, imperfective, future, imperative, injunctive, narrative present, narrative imperfective
- polarity is expressed synthetically
  - positive

PRESENT birsãtʃʰa	FUTURE birselā	NARRATIVE PRESENT birsãdatʃʰa
IMPERFECTIVE birsãtʃʰjo	INJONCTIVE birsos	NARRATIVE IMPERFECTIVE birsãdatʃʰjo
PERFECTIVE birsjo	IMPERATIVE birsa	

  - negative

PRESENT birsãdajna	FUTURE birsojna	NARRATIVE PRESENT birsãdajna
IMPERFECTIVE birsãdajnatʃʰjo	INJONCTIVE nabirsos	NARRATIVE IMPERFECTIVE birsãdajnatʃʰjo
PERFECTIVE birsena	IMPERATIVE nabirsa	

☞ and also many periphrastic sub-paradigms

Inflectional classes

- 4 regular inflectional classes

	PRESENT	PERFECTIVE	FUTURE	INFINITIVE	
1	suttʃʰa	sutjo	sutlā	sutnu	‘sleep’
2	birsãtʃʰa	birsjo	birselā	birsanu	‘forget’
3	ubhĩtʃʰa	ubhijo	ubhielā	ubhinu	‘stand’
4	gāutʃʰa	gājo	gāulā	gāunu	‘sing’

- The ending of the perfective stem determines the class

1	sut-jo	sut	VC
2	birs-jo	birs	CC
3	ubhi-jo	ubhi	i
4	gā-jo	gā	ā

- These inflectional classes can be reduced to phonological conditions on realization rules (Boyé 1999).

A sample sub-paradigm

- In a sub-paradigm, verbs inflect for:

  gender, number, person, honorific grade (LOW, MID, HIGH)

- Neutralizations:
  - no honorific grade for 1st person
  - no gender for plurals
  - high grade honorific neutralizes gender, number, and persons 2 & 3

BIRSANU ‘to forget’

	M.SG	F.SG	PL
1	birsẽ	birsẽ	birsjaũ
2.LOW	birsis	birsis	birsjau
2.MID	birsjau	birsjau	birsjau
3.LOW	birsjo	birsi	birse
3.MID	birse	birsin	birse
HIGH	birsanub^{h}ajo		

Syncretism

- Some sub-paradigms exhibit more syncretism than others
  - Short Negative Present is the most syncretic
    ☞ only 6 distinct forms

BIRSANU

	M.SG	F.SG	PL
1	birsanna	birsanna	birsannaũ
2.LOW	birsannas	birsannas	birsannau
2.MID	birsannau	birsannau	birsannau
3.LOW	birsanna	birsanna	birsannan
3.MID	birsannan	birsannan	birsannan
HIGH		birsanuhunna	

Systematic syncretism

- Some syncretisms hold for all 18 sub-paradigms of all verbs
  - including irregular and suppletive verbs
  - including concurrent realizations

BIRSANU			
	M.SG	F.SG	PL
1	birsãdinã	birsãdinã	birsãdajnaũ
2.LOW	birsãdajnas	birsãdinas	birsãdajnau
2.MID	birsãdajnau	birsãdinau	birsãdajnau
3.LOW	birsãdajna	birsãdina	birsãdajnan
3.MID	birsãdajnan	birsãdinan	birsãdajnan
HIGH	birsanuhũdajna		

The usual solutions fail

Syncretism in inferential-realizational morphology

- Feature neutralization (Stump's “unstipulated syncretism”): syncretism occurs because some feature bundle has no specific exponent.
  - everybody's favorite when applicable.
- Rule of referral (Zwicky, 1985; Stump, 2001; Baerman et al., 2005): the realization of some feature bundle is referred to that of another feature bundle.
  - good for directional, nonsystematic syncretism
- Metarules (Stump 2001, 222-223): if a rule realizes some feature bundle σ using f for exponence, then another rule must realize the feature bundle τ using the same exponence function f.
  - formally and conceptually ill-understood
- Disjunctive feature specifications (Karttunen 1986, Zwicky 2000, Baerman et al. 2005): in marked instances a rule may apply to a disjunctively or negatively specified feature bundles.
  - good for symmetric, nonsystematic syncretism

Feature neutralization for Nepali?

	M.SG	F.SG	PL
1	birsãd-**i**-**na**-**ã**	birsãd-**i**-**na**-**ã**	birsãd-**aj**-**na**-**aũ**
2.LOW	birsãd-**aj**-**na**-**s**	birsãd-**i**-**na**-**s**	birsãd-**aj**-**na**-**au**
2.MID	birsãd-**aj**-**na**-**au**	birsãd-**i**-**na**-**au**	birsãd-**aj**-**na**-**au**
3.LOW	birsãd-**aj**-**na**	birsãd-**i**-**na**	birsãd-**aj**-**na**-**n**
3.MID	birsãd-**aj**-**na**-**n**	birsãd-**i**-**na**-**n**	birsãd-**aj**-**na**-**n**

- Works fine for the **green**, **blue** and **grey** exponents
- Can't account for the distribution of **aj** and **i**: from this table, it looks like **aj** could be the default. In fact **aj** occurs only in imperfect paradigms whereas **i** is more general.
  - Not an isolated problem: see **e** in the simple past
- More generally, does not account for the systematicity of the relevant syncretism, which cuts accross the diversity of exponents.

Rules of referral for Nepali?

	M.SG	F.SG	PL
1	birsãd-i-na-ã	birsãd-i-na-ã	birsãd-aj-na-aũ
2.LOW	birsãd-aj-na-s	birsãd-i-na-s	birsãd-aj-na-au
2.MID	birsãd-aj-na-au	birsãd-i-na-au	birsãd-aj-na-au
3.LOW	birsãd-aj-na	birsãd-i-na	birsãd-aj-na-n
3.MID	birsãd-aj-na-n	birsãd-i-na-n	birsãd-aj-na-n

- We could postulate **three** portmanteau **rules** of **referral**. But :
  - Still no account of the distribution of **aj**
  - No argument to the effect that this syncretism is directional
  - As realization rules, referrals are expected to have exceptions. Here even the most irregular lexemes respect syncretism
  - Rules of referrals are usually last resort strategies for dealing with local oddities. Here their use would be very systematic.

Disjunctive feature specifications for Nepali?

	M.SG	F.SG	PL
1	birsãd-**i**-**na**-**ã**	birsãd-**i**-**na**-**ã**	birsãd-**aj**-**na**-**aũ**
2.LOW	birsãd-**aj**-**na**-**s**	birsãd-**i**-**na**-**s**	birsãd-**aj**-**na**-**au**
2.MID	birsãd-**aj**-**na**-**au**	birsãd-**i**-**na**-**au**	birsãd-**aj**-**na**-**au**
3.LOW	birsãd-**aj**-**na**	birsãd-**i**-**na**	birsãd-**aj**-**na**-**n**
3.MID	birsãd-**aj**-**na**-**n**	birsãd-**i**-**na**-**n**	birsãd-**aj**-**na**-**n**

aj realizes (POL *neg* ∧ FORM *long* ∧ ASP *ipfv* ∧ (NB *pl* ∨ (GEN *mas* ∧ ¬PER *1*)))

i realizes (POL *neg* ∧ FORM *long* ∧ NB *sg* ∧ (GEN *fem* ∨ PER *1*))

- Efficient but ugly
- Does not account for the systematicity of syncretism: that the same cells are syncretic sub-paradigm after sub-paradigm is an accident
- Why are the same disjunctions used realization rule after realization rule?

Transparent forms

A better view

- We take a different perspective to get a better view of the paradigm
  - by exchanging M.SG and F.SG

	M.SG	F.SG	PL
1	birsãd-**i**-**na**-**ã**		birsãd-**aj**-**na**-**aũ**
2.LOW	birsãd-**aj**-**na**-**s**	birsãd-**i**-**na**-**s**	
2.MID	birsãd-**aj**-**na**-**au**		birsãd-**aj**-**na**-**au**
3.LOW	birsãd-**aj**-**na**	birsãd-**i**-**na**	
3.MID	birsãd-**aj**-**na**-**n**		birsãd-**aj**-**na**-**n**

A simpler view

- From this perspective, two facts appear clearly:
  - There are only 2 forms on each line

  The M.SG and PL forms are similar

☞ The table can be collapsed to 2 columns

	A	B	PL
1	birsãd-i-na-ã		birsãd-aj-na-aũ
2.LOW	birsãd-i-na-s	birsãd-aj-na-s	
2.MID	birsãd-i-na-au		birsãd-aj-na-au
3.LOW	birsãd-i-na	birsãd-aj-na	
3.MID	birsãd-i-na-n		birsãd-aj-na-n

A transparent table

- Apart from the 1st person, suffixes are identical on each line.
- Adding a row makes the table completely transparent:
  - one exponent per column
  - one exponent per row

	A	B
1.α	birsãd-**i**-**na**-**ã**	birsãd-**aj**-**na**-**aũ**
1.β	birsãd-**i**-**na**-**s**	birsãd-**aj**-**na**-**s**
2.α	birsãd-**i**-**na**-**au**	birsãd-**aj**-**na**-**au**
2.β	birsãd-**i**-**na**	birsãd-**aj**-**na**
3.α	birsãd-**i**-**na**-**n**	birsãd-**aj**-**na**-**n**
3.β		

Morphomic features

	A	B
1.α	1, SG	
1.β		1, PL
2.α	2.H1, FEM, SG	2.H1, MAS, SG
2.β	2.H2, FEM, SG	2.H1, PL ; 2.H2
3.α	3.H1, FEM, SG	3.H1, MAS, SG
3.β	3.H2, FEM, SG	3.H1, PL ; 3.H2

Transparent forms through morphomic features

The idea: morphomic features

- In realizational morphology, the rules realize **morphosyntactic** features, i.e. features corresponding to syntactic or semantic properties.
- We propose to relax this constraint: rules can realize auxiliary features only indirectly related to morphosyntactic properties.
  - The Nepali sub-paradigms are not structured by the morphosyntactic features GENDER, NUMBER et HONORIFICATION but by the morphomic features COLUMN and ROW.
- This conceptual change does not change the formal properties of realizational morphology.
  - We illustrate this point by giving a detailled account of Nepali synthetic conjugation in PFM.

Relating morphosyntactic and morphomic features

- We relate the morphomic features ROW et COL to the morphosyntactic features by feature cooccurrence constraints

(1) a. {NB *pl*} ⊃ {COL *b*}

b. {GEN *fem*, NB *sg*} ⊃ {COL *a*}

c. {GEN *mas*, NB *sg*} ⊃ ({PER *1*} ≡ {COL *a*})

(2) a. {NB *pl*} ⊃ {ROW *β*}

b. {HON *mid*} ⊃ {ROW *β*}

c. {HON *low*, NB *sg*} ⊃ {ROW *α*}

d. {PER *1*, NB *sg*} ⊃ {ROW *α*}

	F.SG	M.SG	PL
1			
2.LOW			
2.MID			
3.LOW			
3.MID			

The canonical case: the long negative present

		COL *a*	COL *b*
PER *1*	ROW *α*	birsãd-**i**-**na**-**ã**	—
	ROW *β*	—	birsãd-**aj**-**na**-**aũ**
PER *2*	ROW *α*	birsãd-**i**-**na**-**s**	birsãd-**aj**-**na**-**s**
	ROW *β*	birsãd-**i**-**na**-**au**	birsãd-**aj**-**na**-**au**
PER *3*	ROW *α*	birsãd-**i**-**na**	birsãd-**aj**-**na**
	ROW *β*	birsãd-**i**-**na**-**n**	birsãd-**aj**-**na**-**n**

X,σ:{POL *neg*, FORM *long*, COL *a*} → X⊕**i**

X,σ:{ASP *ipfv*, POL *neg*, FORM *long*, COL *b*} → X⊕**aj**

X,σ:{POL *neg*} → X⊕**na**

X,σ:{MOOD *ind*, PER *1*, ROW *α*} → X⊕**ã**

X,σ:{PER *1*, ROW *β*, COL *b*} → X⊕**aũ**

X,σ:{MOOD *ind*, PER *2*, ROW *β*} → X⊕ **au**

X,σ:{PER *1*, ROW *α*} → X⊕**s**

X,σ:{PER *3*, ROW *β*} → X⊕**n**

Further syncretism: the short negative present

- Further syncretisms can be captured by feature neutralizations.

		COL *a*	COL *b*
PER *1*	ROW *α*	birsan-**na**	—
	ROW *β*	—	birsan-**na**-**aũ**
PER *2*	ROW *α*	birsan-**na**-**s**	birsan-**na**-**s**
	ROW *β*	birsan-**na**-**au**	birsan-**na**-**au**
PER *3*	ROW *α*	birsan-**na**	birsan-**na**
	ROW *β*	birsan-**na**-**n**	birsan-**na**-**n**

X,σ:{MOOD *ind*, PER *1*, ROW *α*} → X⊕**ã**

X,σ:{PER *1*, ROW *β*, COL *b*} → X⊕**aũ**

X,σ:{MOOD *ind*, PER *2*, ROW *β*} → X⊕ **au**

X,σ:{POL *neg*} → X⊕**na**

X,σ:{PER *1*, ROW *α*} → X⊕**s**

X,σ:{PER *3*, ROW *β*} → X⊕**n**

Covert syncretism: the positive future

- The content of the cell {TNS *fut*, COL *a*, ROW *α*} looks like a {COL *b*}.

		COL *a*	COL *b*
PER *1*	ROW *α*	birse-**ũ**-**lā**	—
	ROW *β*	—	birse-**aũ**-**lā**
PER *2*	ROW *α*	birse-**li**-**s**	birse-**lā**-**s**
	ROW *β*	birse-**au**-**li**	birse-**au**-**lā**
PER *3*	ROW *α*	birse-**li**	birse-**lā**
	ROW *β*	birse-**li**-**n**	birse-**lā**-**n**

- Analysis : this is a referral to an otherwise unused cell

X, σ:{POL *pos*, TNS *fut*, COL *a*} → X⊕**li**

X, σ:{POL *pos*, TNS *fut*, COL *b*} → X⊕ **lā**

X, σ:{POL *pos*, TNS *fut*, PER *1*, COL *a*, ROW *α*} → ⟨X, *σ*/{ROW *β*}⟩

Morphomic features...

- The introduction of the morphomic features COL and ROW allows a straightforward description of the verbal paradigm.
  - the COL and ROW are added to the description, the original features remain unmodified
  - COL and ROW capture a type of generalization which is not attainable via neutralization, rules of referral or disjunctive feature specifications
- Formal implementation
  - The solution presented here has the advantage of being formally conservative
  - Alternatively, one might want to segregate morphosyntactic (e.g. NUM, GEN, HON) and morphomic (e.g. COL, ROW) features to two different subsystems (à la Sadler & Spencer 2001).

More morphomic features

Concurrent forms

- Some feature bundles have multiple realizations:
  - Negative present: 2 realizations
    - “I do not forget”
      - long → birsãdinã
      - short → birsanna
  - Negative imperfective: 3 realizations
    - “I was not forgetting”
      - long → birsãdajnat^{h}ẽ
      - short → birsannat^{h}ẽ
      - *thi*-form → birsãt^{h}inã
  - Negative future: 2 realizations
    - “I will not forget”
      - prefixed → nabirsũlā
      - suffixed → birsojna

Concurrent forms

- On top of its concurrent negative forms, Nepali possesses narrative vs ordinary present and imperfective forms and the content associated with the contrast is not clear.
- This is in contradiction with the hypothesis underlying realizational frameworks that inflection realization is a function (Zwicky 1986, Anderson 1992, Aronoff 1994, Stump 2001).
- Nepali is not an isolated case: numerous examples in various languages
- Solution
  - relax again the requirement that only morphosyntactic features be realized by inflectional morphology : morphomic features
  - use a FORM feature with values (*long* vs. *short* vs. *thi*)

Inflectional classes

- In PFM, inflectional classes are separate from features.
  - Realization rules specify two types of input: a feature bundle to realize and an inflectional class restriction.
- This architecture does not allow for constraints relating inflectional classes and features.
  - Verbs of the VC class do not have negative short forms
  - Diachronically explainable, synchronically puzzling
- Solution:
  - Inflectional classes are morphomic features
  - Use feature co-occurrence restrictions :

$$\{\text{CLASS } \textit{vc}, \text{ POL } \textit{neg}\} \supset \{\text{FORM } \textit{long}\}$$

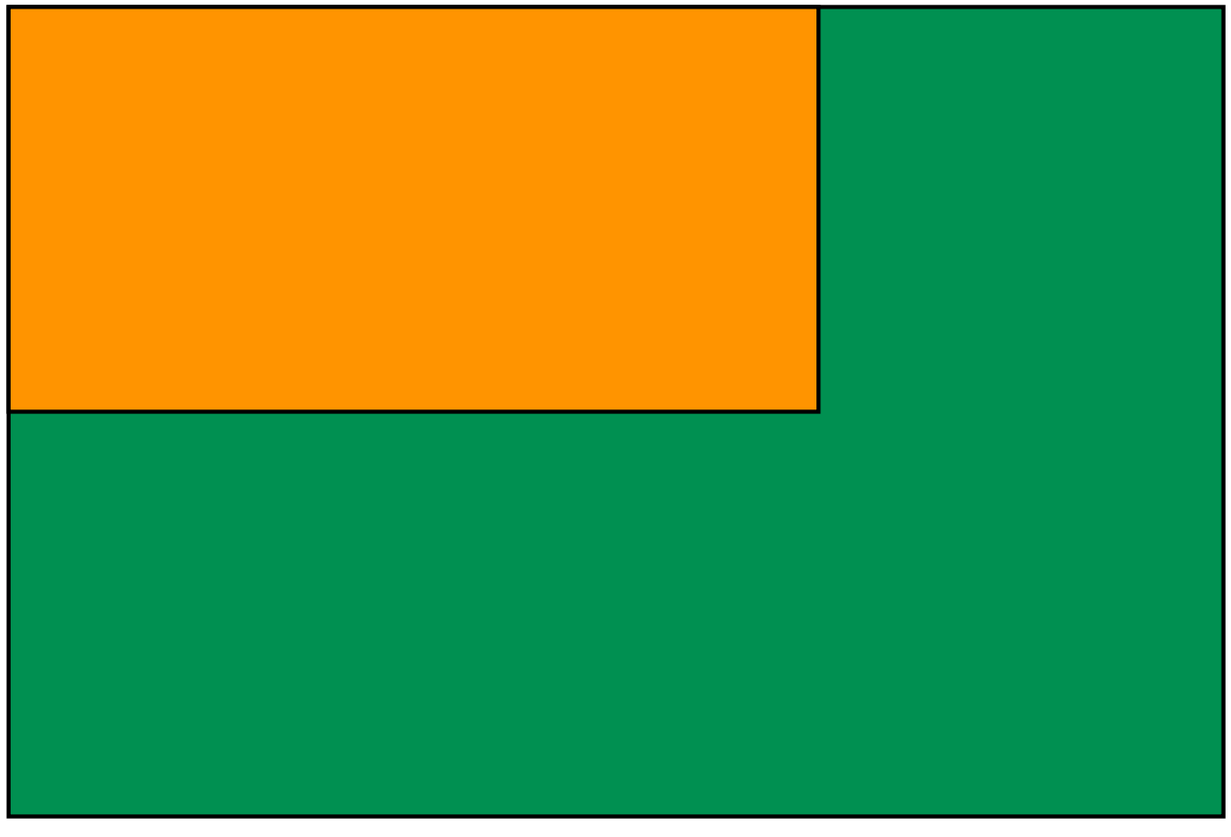
Conclusion

- We provide a full, formally conservative analysis of Nepali conjugation based on morphomic features.
  - Validated by a DATR implementation
- Conceptually, the description of inflectional morphology is modified:
  - We do not prejudge the characterization of the features realized by inflectional morphology.
  - The interface between syntax/semantics and morphology can be more or less transparent.
  - Different morphomic features serve different purposes
    - COL and ROW reduce the dimensionality of every TAM(P) sub-paradigm and thus clarify the morphology
    - CLASS and FORM provide the necessary dimensions needed to describe the whole array of TAM(P) sub-paradigms

Appendix

Syncretism as feature neutralization

- Among two applicable rules, the most specific rule wins.
  - **Syncretic natural class** + **default case**
  - It is possible to have **an apparently non-natural class** :
    - The **blue** rule applies to a **natural class**
    - In **this area**, **blue** competes with **green**
    - In **this area**, **blue** competes with **orange**



Syncretism using rules of referral

	I	II		III		IV		V
	M/F	M/F	NEU	M/F	NEU	M/F	NEU	M/F
NOM	aqua	dominus	donum	homo	nomen	gradus	cornu	res
ACC	aquam	dominum	donum	hominem	nomen	gradum	cornu	rem
GEN	aquae	domini	doni	hominis	nominis	gradus	cornus	rei
DAT	aquae	domino	dono	homini	nomini	gradui	cornui	rei
ABL	aqua	domino	dono	homine	nomine	gradu	cornu	re
	water	*master*	*gift*	*man*	*name*	*step*	*horn*	*thing*

Singular declension of Latin nouns

$X_{II/IV},\sigma$:{CASE *acc*} → X⊕um

$X_{II/IV},\sigma$:{CASE *nom*} → X⊕us

X, σ:{CASE *nom,* GEN *neu*} → ⟨X, σ/{CASE *acc*}⟩

X_{IV},σ:{CASE *acc,* GEN *neu*} → X⊕u

Syncretism using metarules

	I	II		III		IV		V
	M/F	M/F	NEU	M/F	NEU	M/F	NEU	M/F
NOM	aqua	dominus	donum	homo	nomen	gradus	cornu	res
ACC	aquam	dominum	donum	hominem	nomen	gradum	cornu	rem
GEN	aquae	domini	doni	hominis	nominis	gradus	cornus	rei
DAT	aquae	domino	dono	homini	nomini	gradui	cornui	rei
ABL	aqua	domino	dono	homine	nomine	gradu	cornu	re
	water	*master*	*gift*	*man*	*name*	*step*	*horn*	*thing*

For every rule realizing {CASE *acc*, GEN *neu*} by way of ***f***,
there is a rule realizing {CASE *nom*, GEN *neu*} by way of ***f***.

- Difference with referrals: does not take part in rule competition
- Problems:
  - Formally ill-defined
  - As far as we can tell, metarules are asymmetric, despite Stump's claims.

Syncretism using disjunctive feature specifications

	I	II		III		IV		V
	M/F	M/F	NEU	M/F	NEU	M/F	NEU	M/F
NOM	aqua	dominus	donum	homo	nomen	gradus	cornu	res
ACC	aquam	dominum	donum	hominem	nomen	gradum	cornu	rem
GEN	aquae	domini	doni	hominis	nominis	gradus	cornus	rei
DAT	aquae	domino	dono	homini	nomini	gradui	cornui	rei
ABL	aqua	domino	dono	homine	nomine	gradu	cornu	re
	water	*master*	*gift*	*man*	*name*	*step*	*horn*	*thing*

Singular declension of Latin nouns

$X_{II/IV},\sigma$:(CASE *nom* $\vee$ CASE *acc*) $\rightarrow$ X$\oplus$um

$X_{II/IV},\sigma$:(CASE *nom* $\wedge$ GEN *mas*) $\rightarrow$ X$\oplus$us

X_{IV},σ:((CASE *nom* $\vee$ CASE *acc*) $\wedge$ GEN *neu*) $\rightarrow$ X$\oplus$u

- Better alternative to metarules for dealing with symmetric syncretism
- **NB**: rule competition still plays a crucial role in the system.