(How) do underdeveloped writing systems distort grammar?:
written Burushaski 'words' in transcription

Noboru YOSHIOKA / 吉岡 乾

(National Museum of Ethnology, Japan)

noboru7y@minpaku.ac.jp

National Institutes for the Humanities
National Museum of Ethnology

Contents

- Burushaski
- Ever proposed Burushaski Abjads
- How to investigate
- Discussion
  - Transcribing of words: homographic clashes
  - Treating & judging of loanwords
  - Word segmentation
- Do underdeveloped writing systems distort grammar?
- How do underdeveloped writing systems distort grammar?

Burushaski

- bsk (ISO639-3)
- Language isolate
- ~130 000 speakers [Burusho people]
  - Muslim (Isma'ili > Shi'ite > Sunni)
- Gilgit-Baltistan, Pakistan
  - Eastern: Hunza & Nager valleys
    - Some in Srinagar, Jammu-Kashmir, India
  - Western: Yasin valley
- No writing tradition
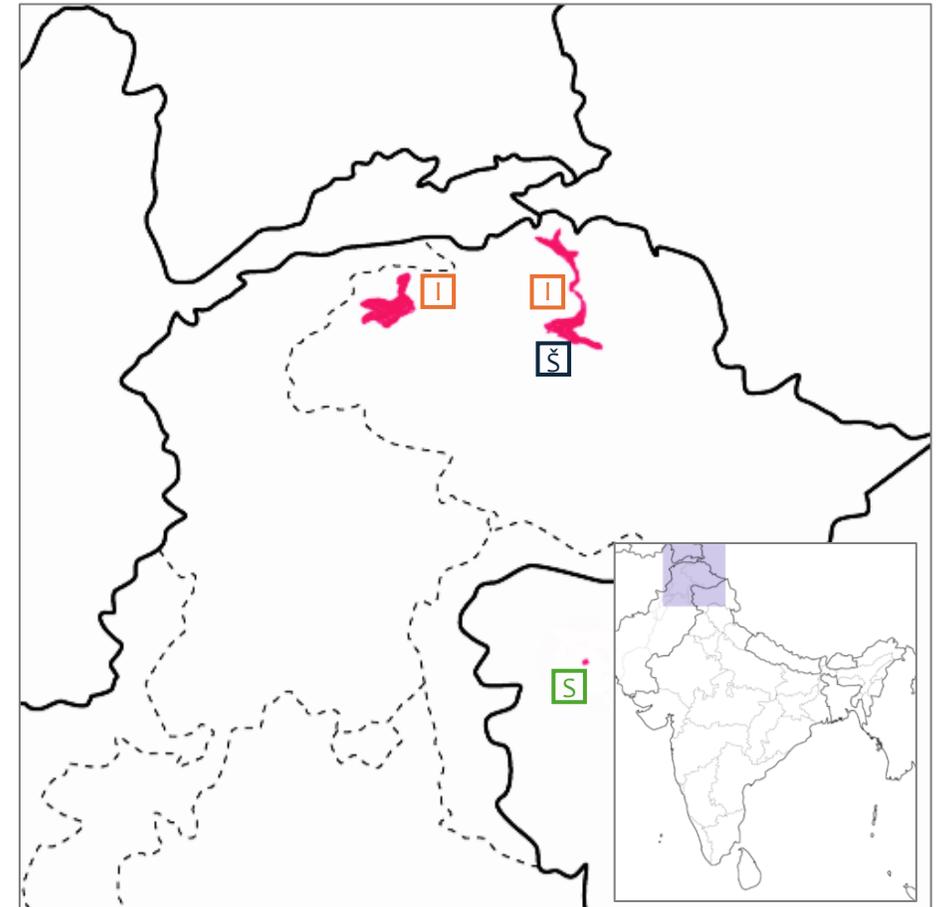
Burushaski

- 36 consonants
- 5 vowels
  - 2 lengths (short vs. long):
    V – VV [Vː]
  - 2 pitches (high vs. non-high):
    V́ [V˥] – V [V˦];
    V́V [Vː˥] – VV́ [Vː˦] – VV [Vː˦]

	front	central	back
close	i ii		u uu
mid	e ee		o oo
open		a aa	

	lab.	den.	ret.	pal.	vel.	uvu.	glo.
plo.	p b pʰ	t̪ d̪ t̪ʰ	ʈ ɖ ʈʰ		k g kʰ	q qʰ	
nas.	m	n			ŋ		
fri.		s z	ʂ	ɕ	ɣ		ɦ
aff.		ts tsʰ	ʈʂ ɖʐ ʈʂʰ	tɕ dʑ tɕʰ			
lat.		l					
tap		ɾ					
app.	w			j	ɰ		

Ever proposed Burushaski Abjads

- Hunzāī, Nasīr ud-Dīn Nasīr (1998): Sporadic word samples
  - Burushaski Research Academy (2006–s.a.): Lexical items & some sentences
- Hunzāī, Ghulām ud-Dīn Ghulām (2007): Massive text (Qur'an) BQ BRA
- Burusho Maraka (2023): Some word samples BM

- Newly proposed writing system <u>can become systematic</u>
  - Due to no historical noises
  - Systematicity requires sufficient knowledge of the language
- They all, native amatures, made their abjads based on the Urdu one
  - Urdu: the national language of Pakistan; Indo-Aryan < Indo-European

Ever proposed Burusha

- How to write the vowels

- Abjad is not good
  for representing of vowels
  - Diacritic marks are not
    obligatorily used
    - i/a/u
    - ii/ee [except word-final position], and
    - uu/oo distinctions sometimes
      become unclear

	BRA			BQ			BM		
i, í	يِـ	ـىِ / ـِـ		يِـ	ـىِ / ـِـ		يِـ	ـىِ / ـِـ	
ii, íi	ـى	يـ		ـى	يـ		ـى	يـ	
ií	ـٰى	يـٰ							
e, é	ـے	يـٰ		ـٔے			ـٔے		
ee, ée	ـے	يـ		ـے	يـ		ـے	يـ	
eé	ـے	يـٰ							
a, á	ٱہَ / ـَٰ			ہَ / ـَ			ہَ / ـَٰ		
aa, áa	ا			ا			ا		
aá	اٰ								
o, ó	و			ؤ			ؤ		
oo, óo	و			و			و		
oó	وّ								
u, ú	ـُو / ـُ			ـُؤ / ـُ			ـُؤ / ـُ		
uu, úu	ـُو			ـُو			ـُو		
uú	ـُوّ								

Ever proposed Burusha...

- How to write the vowels

- Abjad is not good
  for representing of vowels
  - Diacritic marks are not
    obligatorily used

< vowel length >

- BRA: ۲ (2); BQ: ء (hamza);
  BM: ‹ on short vowels /e, o/

< pitch accent >

- BRA: unmarked VV/V́V
  vs. marked VV́ with ۳ (3)

	BRA	BQ	BM
i, í	◌ِ/◌ِی	◌ِ/◌ِی	◌ِ/◌ِی
ii, íi	ی	ی	ی
íi	ی۲		
e, é	؎	ئ	ئ
ee, ée	ے	ے	ے
eé	؎		
a, á	◌َ/◌َا۲	◌َ/◌َه	◌َ/◌َه
aa, áa	ا	ا	ا
aá	ا۲		
o, ó	و۲	ؤ	ؤ
oo, óo	و	و	و
oó	و۲		
u, ú	◌ُ/◌ُو	◌ُ/◌ُؤ	◌ُ/◌ُؤ
uu, úu	◌ُو	◌ُو	◌ُو
uú	◌ُو۲		

Ever proposed Burushaski Abjads

مۃ = 4

- How to write the consonants

- There are 9 consonants
  not used (phonologically) in Urdu

	lab.	den.	ret.	pal.	vel.	uvu.	glo.
plo.	p b	t̪ d̪	ʈ ɖ		k g	q	
	pʰ	t̪ʰ	ʈʰ		kʰ	qʰ	
nas.	m	n				ŋ	
fri.		s z	ʂ	ɕ		ɣ		ɦ
aff.		ts	ʈʂ ɖʐ	tɕ dʑ			
		tsʰ	ʈʂʰ	tɕʰ			
lat.		l					
tap		ɾ					
app.	w			j	ɰ̈		

	BRA	BQ	BM
ts	ژ	ژ	ڎ
tsʰ	ح	ژ	ڎھ
ʈʂ	خ	خ	خ
ʈʂʰ	ض	څھ	څھ
ɖʐ	ژ	ژ	ژ
ʂ	سّ	ش	ش
ŋ	گّ	نّگ	نّ
ɰ̈	يِ	ىّ	ئ
qʰ	خ	خ	خ
pʰ	ف	ف	پھ

8

Ever proposed Burushaski Abjads

- How to write the consonants
- There are 9 consonants not used (phonologically) in Urdu

Ur. فائده **_fāiḍa_** > Bur. **_pʰaaiḍá_**

Ur. خدا **_xuḍā_** > Bur. **_qʰuḍáa_**

Ideal correspondence regarding aspiration

unaspirated	پ /p/	ت /t̪/	ٹ /ʈ/	ک /k/	چ /tɕ/	څ /ts/	څ /ʈʂ/	ق /q/
aspirated	پھ /pʰ/	تھ /t̪ʰ/	ٹھ /ʈʰ/	کھ /kʰ/	چھ /tɕʰ/	څھ /tsʰ/	څھ /ʈʂʰ/	قھ /qʰ/

What actually they proposed (fully or partially)

unaspirated	پ /p/	ت /t̪/	ٹ /ʈ/	ک /k/	چ /tɕ/	څ /ts/	څ /ʈʂ/	ق /q/
aspirated	ف /pʰ/	تھ /t̪ʰ/	ٹھ /ʈʰ/	کھ /kʰ/	چھ /tɕʰ/	ژ /tsʰ/	ض /ʈʂʰ/	خ /qʰ/

How to investigate

- Burushaski paginally translated Qur'an – Eastern Burushaski [Hunza]
  - Text body: 6 236 proses (total of 114 chapters) + annotations
- Digitising
- Transliterating: one-to-one romanization (to make light-to-right)
- Deciphering: what the author wanted to write
- Transcribing: actual pronunciation in Burushaski
- Analysing: morphologically
- Labelling
- Translating

- Digitising   ﴿٣٥﴾ ژَنیکه یه اُوے اُئ بَم۔ که خدا دُم بغیر خدایَن اَپائے، اوسُ مَر اُوے تکبُّر اهچَم

- Transliterating: one-to-one romanization (to make light-to-right)

ċanykh yh ʔuve ʔuy² bam. kh xdʔ cum bɣyr xdʔyan ʔapʔy²e, ʔvsu mar ʔuve tkbur ʔʰčam ((35))

- Deciphering: what the author wanted to write

tsʰáneke jé ué úi bám. ke xuḍātsum baɣɛ̄r xuḍājan apái, ósumar uée ṭakabbur étɕám

- Transcribing: actual pronunciation in Burushaski

| tsʰáneke | jé | ué | úi | bám. | ke | qʰuḍáatsum | baɣáir | qʰuḍáajan | apái, | ósumar | uée | ṭakabúr | étɕám |

- Analysing: morphologically

| tsʰán-e=ke | jé | ué | u-í | báan-m | =ke | xuḍā-ts-m | baɣɛ̄r | xuḍā-an | a-bái-Ø | u-s-m-ar | ué-e | ṭakabbur | i-ṭ-tɕ+báan-m |

- Labelling

straight-GEN=CONJN INTERJ DIST:PL.H 3PL.H:I-self COP:3PL.H-NPRS =CONJN god[Ur]-ADE-ABL without[Ur] god[Ur]-INDF.SG NEG-COP:3SG.HM-PRS 3PL.H:II-tell-ADJVLZ-DAT DIST:PL.H-ERG arrogance[Ur]
3SG.Y:II-do-IPFV+COP:3PL.H-NPRS

- Translating

Indeed, they themselves were such. When they were told that there is no god but the God, they always became arrogant. [Q37:35]

11

Transcribing of words: homographic clashes

- او: **ú** 'they (H)' / **uú** 'their father' / **óo** 'their (H)'
- اوے: **úe** 'of them (H)' / **ué** 'those (H)' / **uée** 'of those (H)' / **uúe** 'of their father'
- اوسس: **ósas** 'to tell them (H)' / **óosas** 'to put'
- اوتس: **óṭas** 'to make them (H/X)' / **óoṭas** 'to do for them (H/X)'
- فاش: **pʰáaɕ** 'notorious' / **pʰaáɕ** 'flowed out in small amounts'
- ایچم: **étɕam** 'I will do it (Y)' / **étɕám** 'you (SG/PL)/he/they was doing it (Y)'
- ایچئین: **étɕien** 'they (X) will do it (Y)' / **étɕién** 'they (X) do it (Y)'

Treating and judging of loanwords

- Ur. تكليف ***ṭaklīf*** 'trouble' > Bur. ***ṭaklíipʰ***

 BQ: 33

ṭaklíipʰtɕiŋ   ***d̪íibitsúm.***
ṭaklīf-tɕiŋ   d̪íi+bitsan-m
trouble[Uʀ]-PL   come:PFV.3PL.Y+COP:3PL.Y-NPRS
'... troubles had come.' [Q2:214]

 BQ: 37   Not تكليپ ?

úaɾ   ***ṭaklíip***   ***étɕaɾ.***
ú-aɾ   ṭaklīf   i-ṭ-tɕ-aɾ
DIST.PRN:PL.H-DAT   trouble[Uʀ]   3SG.Y:II-do-IPFV-DAT
'... to giving them the hardship.' [Q2:231]

Burushaski morphophonological rule:
$$C > C_{[-\text{aspirate}, -\text{voiced}]} / \_\#$$

13

Treating and judging of loanwords

- Pe./Ur. جماعت *ʤamāaṱ* 'gathering, congregation, company' >

BQ: 20

older loanword   newer loanword

Bur. *ʤamáaṱ* 'spouse' vs. *ʤamáaṱ* 'group'

~ *ʤamaáṱ*

kʰué	*ʤamaáṱ*	*níbáan.*
kʰué	ʤamāaṱ	ní+báan-Ø
PROX:PL.H	congregation[Uʀ]	go+COP:3PL.H-PRS

'These group members have gone.' [Q2:134]

BQ: 498

→ Not treated as a loanword!

(written in how it is pronounced)

ué	*ʤamáaṱiɕo*	*umáiman.*
ué	ʤamāaṱ-iɕo	u-man-tɕ-m-an
DIST:PL.H	congregation[Uʀ]-PL	3PL.H:I-become-IPFV-NPRS-3PL.H

'They will become wives.' [Q44:54]

Treating and judging of loanwords

- Ur. گناہ *gunāh* 'sin' > Bur. *gunáa*

BQ: 81

(nikáa	*éṭas)*	*mámaṭar*	*gunáa*	*apí.*
nikāh	i-ṯ-as	má~RDP-aṭ-ar	gunāh	a-bila-Ø
matrimony[UR]	3SG.Y:II-do-INF	you~OBL-INS-DAT	sin[UR]	NEG-COP:3SG.Y-PRS

'(getting marry) is not a sin on you.' [Q4:23]

BQ: 9

Not گناہر ~ گناہر ؟

Bare forms retain the silent letter, whereas declined forms drop it

dʑe	*máa*	*gunáar*	*máap*	*étɕam*
dʑé	má-e	gunāh-ar	muāf	i-ṯ-tɕ-a-m
I	you-GEN	sin[UR]-DAT	exempt[UR]	3SG.Y:II-do-IPFV-1SG-NPRS

'I will forgive your sin' [Q2:58]

15

Treating and judging of loanwords

- For judging loanwords what is reliable?

BQ: 8

ḏáa   **máimo**   **námsiɛo**   **ósqanin,**
ḏáa   ma-í-mo   náms̠-iɛo   uːs-ɣan-in
again   2PL:I-self-GEN.Z   mind-PL   3PL.X:II-CAUS-finish-IMP.PL
'and kill your minds' [Q2:54]

non-loanword?
نمس **nams** 'mind' ?< Ur. نفس **nafs** 'mind'

Berger (1998: 300)

Treating and judging of loanwords

- For judging loanwords what is reliable?

دَا خدایه بِئسن مار گِرمِنائے که اِته هُرُور اهِتِن۔

BQ: 29

ḍáa	*qʰudáaje*	*bésan*	*máar*	*girmínáike*	*iṭ̱é*	*ħuzúur*	*éṭin.*
ḍáa	xudā-e	bés-an	ma÷ar	girmín+bái-Ø=ke	iṭ̱é	ħuɽūr(?)	i-ṭ-in
again	god[Uʀ]-ERG	what-INDF.SG	2PL:II-DAT	write+COP:3SG.HM-PRS=CONJN	DIST:Y	reliance(?)[?]	3SG.Y:II-do-IMP.PL

'Keep what the God has written down for you.' [Q2:187]

Shina doesn't have the /z/ sound,
while Urdu doesn't use the /ʤ/ sound (here ذ)

Burushaski doesn't have

the /ɽ/ sound (Urdu ڑ)

non-loanword?

ħuzúur ?< Sh. **ħuʤúur** 'potentate'

هُذ ور: حضور ـ(ا ـ وا ـ ذ)ـ

Ziā (2010: 446)

?< Ur. حضور **ħuzūr** 'potentate, presence'

< Ar. حضور **ħudʕūr** 'presence, quietude'

اگر مَهٔ مئرَن که یا مَسقَنَن که خدایه حُضُورَلو گَٹی مَمائے مَن ۝۱۵۸۝

BQ: 71

'If you die or are killed, you shall be gathered into the presence of the God.' [Q3:158]

17

Word segmentation

- Are there consistent rules?

تکلیف چنگ دی بِدُم
BQ: 33

ṭaklíipʰtɕiŋ	**dʲiibitsúm.**
ṭaklīf-tɕiŋ	dʲii+bitsan-m
trouble[Ur]-PL	come:PFV.3PL.Y+COP:3PL.Y-NPRS

'... troubles had come.' [Q2:214]

Plural suffix: separated

Auxiliary Copula: **separated**

كُهؤے جماعت نِبان
BQ: 20

kʰué	**dʑamaáṭ**	**níbáan.**
kʰué	dʑamāaṭ	ní+báan-Ø
PROX:PL.H	congregation[Ur]	go+COP:3PL.H-PRS

'These group members have gone.' [Q2:134]

Auxiliary Copula: **not separated**

أوے جمات شو أمائے مَن
BQ: 498

ué	**dʑamáaṭiɕo**	**umáiman.**
ué	dʑamāaṭ-iɕo	u-man-tɕ-m-an
DIST:PL.H	congregation[Ur]-PL	3PL.H:I-become-IPFV-NPRS-3PL.H

'They will become wives.' [Q44:54]

Plural suffix: **separated**

Conjugation suffix: **separated**

دَا مَیٖمؤ نَمٔسٖٮشُو اؤشقَنٖن
BQ: 8

dʲáa	**máimo**	**námsiɕo**	**ósqanin,**
dʲáa	ma-í-mo	náms-iɕo	uːs-ɣan-in
again	2PL:I-self-GEN.Z	mind-PL	3PL.X:II-CAUS-finish-IMP.PL

'and kill your minds' [Q2:54]

Plural suffix: **not separated**

Conjugation suffix: **not separated**

Word segmentation

- Background: Treatment in Urdu
  - Case markers, Auxiliaries, Conjunctions and so on are written separately
- Attached
  - Case suffixes
  - Conjugation suffixes
  - Plural suffixes
- Detached
  - Case markers (clitics)
  - Verbal auxiliaries (compounding ~ clitics)
  - General conjunction **=ke** (کہ)

مجھ **mudʑʰ** 'I:OBL'
مجھے **mudʑʰ-ē** 'I:DAT'
مجھ سے **mudʑʰ=sē** 'I:ABL'

بول **bōl** 'speak! (IMP.SG)'
بولوں **bōl-ū̃** 'I may speak'
بولوں گا **bōl-ū̃=gā** 'I will speak'

کتابیں **kiṭāb-ẽ** 'books'
مردو **mard-ō** 'hey men! (VOC)'
مدروں نے **mard-õ̃=nē** 'men (ERG)'

Word segmentation

- General conjunction clitic: **=ke** (کہ)

béʝal	***momíndarue***	***káa***	***umáibáanke,***	***séibáan,***	***ke***
béʝal	mōmin-daro-e	káaʈ	u-man-tɕ+báan-Ø=ke	sén-tɕ+báan-Ø	=ke
when	pious[Uʀ]-PL-GEN	together	3PL.H:I-become-IPFV+	say-IPFV+COP:3PL.H-PRS	=CONJN
			COP:3PL.H-PRS=CONJN		

'when they meet with pious people, then they say that …' [Q2:14]

ágaɽ	***ma***	***momíṇḍaro***	***báanke,***
agaɽ	má	mōmin-ḍaro	báan-Ø=ke
if[Uʀ]	you	pious[Uʀ]-PL	COP:2PL-PRS=CONJN

'If you are pious, then …' [Q3:175]

Word segmentation

- Burushaski case markers are suffixes: e.g. LOC **-ulo** //-ul-e//

BQ: 71

Not دوزخلو

íne   ***ḓiɕ***   ***ḓuuzáqʰulo***   ***bilá.***

ín-e   ḓiɕ   ḓōzax-ul-e   bila-Ø

DIST.PRN:H-GEN   place   hell[UR]-LOC-GEN   COP:3SG.Y-PRS

'There is his place in the hell.' [Q3:162]

BQ: 35

ḓáa   ***aaqʰiráṭe***   ***(tɕáɣamiŋ)***   ***ulo***   ***(sambá***   ***éṭin)***

ḓáa   āxiraṭ-e   tɕáɣa-miŋ   -ul-e   sambá   i-ṭ-in

again   afterlife[UR]-GEN   story-PL   -LOC-GEN   thought   3SG.Y:II-do-IMP.PL

'and think inside the stories of the next life' [Q2:220]

Word segmentation

- Burushaski case markers are suffixes: e.g. LOC *-ulo* //-ul-e//

بِیۡسَن چِیزَن اَمَائَت قَبۡضُو (مَرین) لو نوسِن (اُش گَنِن) BQ: 49

bésan	***tɕíizan***		***amaanáṭ***	***qabzáu***		***(márian)***
bés-an	tɕíīz-an		amānaṭ	qabzā-u(l)		máaɾi-an
what-INDF.SG	thing[UR]-INDF.SG	trust[UR]		occupation[UR]-LOC		tribute[SH]-INDF.SG

lo	***nóosin***	***(uṣ***	***gánin)***
-(u)l-e	n-óos-n	úṣ	gán-in
-LOC-GEN	CP-put-CP	guilt	take-IMP.PL

'borrow some item having put them as collateral for the trust' [Q2:283]

Word segmentation

- Some Burushaski case markers have been derived from free positional nouns etymologically

- **-ul-o** locative 'in' < **úl-o** 'inside:GEN' (WB **úl-e**)
- **-aʈ-e** instrumental 'by, around' < **@-jáʈ-e** 'upwards:GEN'
- **-ts-e** adessive 'on' < **@-tsí-(e)** 'against:GEN'
- **-ar** dative 'to' < **@-jár** 'forwards'

†@- indicates the personal prefix slot

- Attachment/detachment tells us whether the spelled part is a case suffix or a positional noun in their scripts occasionally

23

Do underdeveloped writing systems distort grammar?

- Yes, possibly

@-ríiŋ 'hand'

		SG	PL
1		aríiŋ	miríiŋ
2		guríiŋ	maríiŋ
3	HM	iríiŋ	uríiŋ
	HF	muríiŋ	
	X	iríiŋ	uríiŋ
	Y	iríiŋ	iríiŋ

BRA, 2: 170

-ríiṅ -ćiṅ, ng. -ćaṅ, Has. -ríin Hand
1.8; 4.16; 45.5; 46.3; 48.2 u.a.; L 304
(ys. -rén, pl. -réiṅ, -réiṅćiṅ)   Berger (1998: 364)

- ***kʰóle ríiŋ bilá*** 'There is a (someone's) hand'
- ***kʰóle iríiŋ bilá*** = the default person-class is 3SG.HM
- Some native speakers (wannabe elites) intend to change even their speaking rules for making themselves "language teachers"

How do underdeveloped writing systems distort grammar possibly? (in the case of Burushaski)

<u>< Letter choices ></u>

- Able to mislead (/ accelerate) phonological shift: /p^h/ > [f]; /q^h/ > [x]

<u>< Negligence of the distinctive accent ></u>

- Neutralisation of different TAM forms

<u>< Judgement of loanwords ></u>

- Folk etymology > phonological/morphological revisionism

<u>< Word segmentation in writing ></u>

- Grammatical Urduisation
  - Reanalysis of inflectional affixes as adpositions (i.e., clitics or free words)
  - [consequently] Confusion on case suffixes with postpositional nominals

How do underdeveloped writing systems distort grammar possibly? (in the case of Burushaski)

<u>< Abjadhood ></u>

- Doesn't affect the grammar directly

<u>< Assymmetry/unsystematicity of letters ></u>

- Doesn't affect the grammar directly

<u>< Borrowing the spelling of loanwords ></u>

- Doesn't affect the grammar directly

<u>< Unstable spelling of inherent words ></u>

- Doesn't affect the grammar directly by itself

References

- **Berger**, Hermann. **1998**. *Die Burushaski-Sprache von Hunza und Nager, Teil III: Wörterbuch*. Wiesbaden: Harrassowitz Verlag.

- **Burushaski Research Academy. 2006–s.a.**. *Burūshaskī-Urdū Lughat* [Burushaski-Urdu Dictionary]. 3vols. Karachi: Bureau of Composition, Compilation & Translation, University of Karachi.

- **Burusho Maraka. 2023**. *Burūshaskī alif bā* [The Burushaski Alphabet]. s.l.: Burusho Maraka.

- **Hunzāī**, Ghulām ud-Dīn Ghulām (transl.). **2007**. *Al-Qur'ān al-Karīm: Burūshaskī tarjumā* [The Noble Qur'an: Burushaski translation]. Gilgit: Oxford Gilgit Printers & Gilgit Kashmir Printers.

- **Hunzāī**, Nasīr ud-Dīn Nasīr. **1998**. *Dékaran* [Small Garden]. Karachi: Burushaski Research Akademy.

- **Ziā**, Muhammad Amīn. **2010**. *Shinā-Urdū Lughat* [Shina-Urdu Dictionary]. Rawalpindi: ST Printers.