

The uvularization of *g-/d-* in Tibetic languages¹

This study reviews four competing explanations for the origins of uvular preinitials in Tibetic lects, making a specific case study of modern uvular preinitial reflexes from Old Tibetan *g-/d-*. The first explanation is from Huang (2012), who claims that uvular preinitials were phonologically present in Pre-Tibetan, and thus at least some of the uvular preinitials in modern Tibetic lects descend from this Pre-Tibetan strata. Her argument is predicated on the hypothesis that Tibetic lects broke up into different languages before Old Tibetan was reduced to writing in the 7th century AD. The second explanation is from Hill (2010), who argues that all uvular preinitials are not inherited from Pre-Tibetan but are the result of language contact with Qiangic and/or Mongolic languages. Differing from Huang's explanation, Hill's explanation rests on the theory that all modern Tibetic lects descend from Old Tibetan. The third explanation assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from *g-/d-* to velar/postvelar/uvular fricatives (except before velar initials, where the change is to *r-*) in Amdo lects and to uvular fricatives in Gyalrongic lects. For WAT lects, *g-* regularly changes to velar/postvelar/uvular fricatives, but *d-* changes to velar/postvelar/uvular or *ʃ-/r-*. The fourth explanation is that in Old Tibetan *ɣ* and *g-/d-* were in velar and uvular free variation, and thus uvular preinitials do come from Old Tibetan, but originate from phones and not phonemes. The first three explanations are scientific hypotheses; i.e., they can be tested through evidence and are falsifiable. The final explanation (appealing to free variation) is not a falsifiable. After examining the evidence on the timing of the breakup of the Tibetic lects, Huang's hypothesis is eliminated, leaving only Hill's explanation and 'Explanation 3' standing. However, Explanation 3 is the only explanation that proposes a set of regular sound changes to summarize the uvularization of *g-/d-*.

Keywords: Sa-skya Paṇḍita's Law; Pre-Tibetan language; Old Tibetan language; Gyalrongic languages; uvular reflexes of *g-/d-* in Tibetan.

1. Introduction

This study reviews four competing explanations for the origins of uvular preinitials in Tibetic lects, making a specific case study of modern uvular preinitial reflexes of Old Tibetan *g-/d-*. Section 2 discusses Sa-skya Paṇḍita's Law and the resulting complementary distribution of *g-* and *d-*, and Section 3 presents words with *g-/d-* reflexes in modern Tibetic languages and loanwords with *g-/d-* reflexes in modern Gyalrongic languages.

Section 4 discusses an explanation from Huang (2012), who postulates that uvular preinitials were phonologically present in Pre-Tibetan (inherited from Tibeto-Burman) and thus uvular *g-/d-* in modern Tibetic lects are reflexes, descending from an unwritten Old Tibetan

¹ This paper has received funding from: China Postdoctoral Science Foundation (73rd batch), Funding number: 2023M731815: "Anticausative marking in West Gyalrongic languages" and National Social Science Fund Key Project (21AYY024), People's Republic of China: "Research on Tibetan poetry, rhythm, and related prosodic phonology."

Glossing in the examples follow the Leipzig Glossing Rules. Old Tibetan and Written Tibetan transcriptions follow the Wylie transcription system (Wylie 1959), except in the case of α where *ɣ* is used instead of an apostrophe, following Hill (2019b).

with phonological uvular preinitials. Her argument is predicated on the notion that there was a split in Old Tibetan that resulted in a non-uvular written Old Tibetan (based on an old U-Tsang lect) as one branch (with Written Tibetan and modern U-Tsang as descendants), and another branch of uvular Tibetic lects (e.g., Kham and Amdo)².

Section 5 examines Hill's 2010 explanation, including his explanation for why *g-/d-* has uvular fricative reflexes in some modern Tibetic lects. Hill argues that uvular preinitials are not inherited from Pre-Tibetan but are the result of language contact with Qiangic and/or Mongolic languages. Differing from Huang's explanation, Hill's explanation rests on the theory that all modern Tibetic lects descend from Old Tibetan.

Section 6 discusses a third explanation. Explanation 3 assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from *g-/d-* to velar/postvelar/uvular fricatives (except before velar initials, where the change is to *r-*) in Amdo lects and to uvular fricatives Gyalrongic lects. For WAT lects, *g-* regularly changes to velar/postvelar/uvular fricatives, but *d-* changes to velar/postvelar/uvular or *ʃ-/r-*. Hill's notion of contact induced uvularization is acknowledged as a possible origin story, but in light of the regularity of the sound change it is not explored further.

Section 7 discusses the fourth explanation, which is that in Old Tibetan *h* and *g-/d-* were in velar and uvular free variation, and thus uvular preinitials do come from Old Tibetan, but originate from phones and not phonemes.

Section 8 evaluates each of these four explanations. The first three explanations are scientific hypotheses; i.e., they can be tested through evidence and are falsifiable. The final explanation (appealing to free variation) is not a testable hypothesis, but merely one possibility. After examining the evidence on the timing of the breakup of the Tibetic lects, Huang's hypothesis is eliminated, leaving only Hill's explanation and Explanation 3 standing. Explanation 3 is the only explanation that proposes a set of regular sound changes to summarize the uvularization of *g-/d-*, and thus this explanation is considered as the best choice thus far.

2. Sa-skya Paṇḍita's Law

In the synchronic phonology of both Written Tibetan and Old Tibetan, *g-* and *d-* are in complementary distribution³. Sa-skya Paṇḍita's Law claims that since Old Tibetan *g-* and *d-* are in complementary distribution, *g-* only coming before acutes and *d-* only coming before graves, then the following sound changes must have taken place: **g- > d-* before graves (labials and velars) and **d- > g-* before acutes (dentals and palatals) (Hill 2011: 443-444, 2019b: 24)⁴. This conditioned neutralisation happened before Old Tibetan (i.e., Pre-Tibetan), or may have even occurred much earlier in the split off of Proto-Bodish from Proto-Trans-Himalayan (Proto-Sino-Tibetan).

² Personally, I don't like the term 'Written Tibetan' as simply 'Tibetan' would suffice for this *état de langue*. I use 'Written Tibetan' to identify with other scholars who use this term and in order to avoid confusion with modern spoken Tibetan (or Tibetic) lects. In Tibetan literary studies the synonym 'Classical Tibetan' is used.

³ Phonotactics play an important role in the allophony of *g-* and *d-*. The claim is only made for *g-* and *d-* as preinitials to be allophones, not for the initials *g* and *d* to be allophones. Thus, before the medial *-r-*, *g-* and *d-* are in contrast in both Old and Written Tibetan. See also Jacques 2014: 157.

⁴ The terms 'grave' and 'acute', coined by Jakobson et al. (1951: 29–30), have fallen out of fashion in general linguistic literature. However, there are two main reasons for resurrecting these terms. Firstly, the term 'grave' is typologically quite satisfying, since labials and velars have a phonetic connection in many languages of the world. Secondly, these terms are helpful for engaging in related Sino-Tibetan historical comparison, as these terms are used for Old Chinese by Baxter (1992), Baxter & Sagart (2014), and Hill (2019b), among others.

It not impossible to imagine that **g-* and **d-* were prefixes in Proto-Trans-Himalayan. One hypothesis proposed by Jacques (2008: 53–57, 2014: 158–160), who provides comparative evidence with Tibetan and Gyalrongic, is that **g-* comes from an animal noun classification prefix (perhaps **gV-*) and **d-* is a lexicalized indefinite possessor prefix in inalienable (kinship and body part) nouns (perhaps originating as **dV-*)⁵. For our purposes at the moment it is suffice to be agnostic concerning their function, and even entertain the notion that **g-* and **d-* were never prefixes. Whether they were prefixes or not does not have a direct effect on our present study, but for now I will just assume that these were prefixes at some earlier stage.

Over time, these prefixes lost their meaning and became fossilized as non-productive pre-initials, merging with preinitials of the same phonological value but unrelated to the prior prefix; **g-* and **d-* at this stage remained separate phonemes. Next comes a purely phonological change; Sa-skyā Paṇḍita's Law redistributes these fossilized prefixes through a new phonological constraint (*g-* only before acutes and never before graves, and *d-* only before graves and never before acutes), so that what was once in principle **dd*, **dts*, **ds*, **dz*, **dn*, **dsh*, **dj*, etc. is now *gd*, *gts*, *gs*, *gz*, *gn*, *gsh*, *gj*, etc., and what was once **gk*, **gg*, **gp*, **gb*, **gm*, etc. is now *dk*, *dg*, *dp*, *db*, *dm*, etc. Thus, *g-* and *d-* as preinitials are in complementary distribution, and become allophones of the same preinitial phoneme, while still remaining in contrast as initial and final position phonemes. If there was a written language for Pre-Proto-Bodish, then we would expect to find examples of at least some of the following forms: **dd*, **dts*, **ds*, **dz*, **dn*, **dsh*, **dj*, **gk*, **gg*, **gp*, **gb*, **gm*, etc. It is possible that part of the motivation for Sa-skyā Paṇḍita's Law is the avoidance of double consonants, e.g., *dd* and *gg*, which then became *gd* and *dg*, respectively. Thus, Sa-skyā Paṇḍita's Law is mostly a process of dissimilation.

If we combine Jacques' hypothesis (*g-* is from an animal prefix and *d-* is from a kinship/body part prefix) with Sa-skyā Paṇḍita's Law we can assume that in some Tibetan words with a *g-* preinitial before an acute initial (e.g., in some animal terms) no change occurred (*g-* > *g-*), but when we have *g-* before an acute initial for a kinship or body part word (like *gzhang* 'anus') then the *d-* > *g-* change occurred; in some words with a *d-* preinitial before a grave initial (e.g., body part terms) no change occurred (*d-* > *d-*), but when we have a *d-* preinitial before a grave initial in animal words (*dgo ba* 'Mongolian gazelle', *dbyi* 'lynx') then *g-* > *d-*. The term *dbyi* 'lynx' also has an alternative spelling *g.yi*, which hints to this sound change and shows the complementary distribution (*g-* cannot occur before *b*), which led to the ousting of *b* and the retention of *g-* in *g.yi*. In Sertha, 'eye' is pronounced by some varieties as *ɣəɲiŋ* and other varieties as *ɣəɲiŋ*; turning the original preinitial into a uvular or velar fricative initial of a new syllable by adding an epenthetic schwa.

3. Reflexes of *g-/d-* in Tibetic words and Gyalrongic loanwords

Table 1 tabulates some examples of *g-* in Tibetic and Gyalrongic lects, and Table 2 gives examples of *d-*. Since Gyalrongic languages have been in contact with Tibetic languages for over a millennium, and have thus borrowed extensively from Tibetic, Gyalrongic languages provide

⁵ This proposal has been called into question by Hill (2014: 629–630), because it is impossible to completely rule out chance operations (e.g., it would be preposterous to propose that the 'h' in 'head', 'hand', 'heart', 'heel', etc. indicate a *h-* prefix that is related to body parts in English). However, Jacques (2014: 158–160) is not proposing a body part prefix but an indefinite possessor prefix of inalienably possessed nouns, which is well attested in Gyalrong (Jacques 2021: 114). The *d-* preinitial in Tibetan is but a fossilized trace of that prefix. Since in Old Tibetan *g-/d-* are found covering a broad range of semantic domains, e.g., body parts, animals, colors, times of day, abstract notions, etc., it could be that there were several prefixes with velar or dental initials that merged in Proto-Tibetan (Bialek p.c.)

an excellent point of comparison. These tables are in no way exhaustive, but still provide ample evidence for this study.

The data in the columns ‘Balti’ and ‘WAT’ (Western Archaic Tibetan) comes from Bielmeier et al. (2018; Forthcoming). WAT examples give the lect name in parentheses after the WAT form. The data in the column ‘Amdo’ comes from Bielmeier et al. (2018; forthc.) and Hua (2001). Data from Bielmeier et al. (2018; forthc.) indicate the lect in parentheses after the WAT and Amdo form, but data from Hua do not indicate the lect (more explanation is given in the next paragraphs). In order to simplify the Amdo data from Hua, I only include examples from lects that have preinitials. If no preinitial was found I chose just one of the Amdo lects to represent them all, as lects that have lost the preinitial have no bearing on our study. Although Hua has chosen to use the voiceless glottal fricative *h-* as a preinitial, it is likely that the voiceless velar fricative *x-* would be a better choice for some if not all of these examples. In any case, there is no phonological contrast between [h-] and [x-] as a preinitial in these lects; and even if there is a reflex *h-*, this is still explainable by the hypotheses presented in this article.

The following are notes for the Amdo data from Hua (2001) in Tables 1 and 2, indicating which dialects the data reflects. If nothing is said about the Amdo forms below, then all the lects of Amdo from Hua have the same form. The Amdo for *gtsang ma* ‘clean’ is from Rebong (Tongren), Xunhua, Hualong, and Skakhog/Skachu (Hongyuan). According to Hua, in Xunhua *gser* ‘gold’ is *hse*, in Hualong *hsar*, and in Skakhog/Skachu (Hongyuan) and Themchen (Tianjun) it is *hser*; *gso* ‘nourish’ is from Bsangchu (Xiahe), Rebong (Tongren), Skakhog/Skachu (Hongyuan), Themchen (Tianjun). The data for *gzah hkhor* is from Skakhog/Skachu (Hongyuan); *hza* being the first part of *hzadawa* ‘Monday’. The word for *dpon* in Themchen (Tianjun) is *hwonbu*. Hua has recorded that *dgah* has become *hga* in Bsangchu (Xiahe), Rebong (Tongren), Xunhua, and Xualong, but in Skakhog/Skachu (Hongyuan) and Themchen (Tianjun) has become *rga*.

G.yukhog and Stau data come from my own fieldwork (Gates 2021). G.yukhog is traditionally considered to be an Amdo lect, but I display the data separately since it has close proximity to Stau and provides ample examples of uvular preinitials for a Tibetic lect. Geshiza data comes from Honkasalo (2019) and personal communication with Honkasalo (October 31 – November 1, 2021). Khroskyabs data comes from Lai (2017) and personal communication with Lai (October 31, 2021). Japhug data comes from Jacques (2015–2016).

In both Tables 1 and 2, blank places indicate that there is no relevant comparanda available; in the case of the Tibetic lects no cognates were found, in the case of Gyalrongic no loanwords were found.

Balti *g-/d-* reflexes are mostly *x-/ɣ-*, with the occasional *g-* reflex as *χ-* (e.g. *χsu* ‘feed’) and *ʈ-* (e.g. *ʈsanma* ‘clean’), and the occasional *d-* reflex as *ʂ-* (e.g. *ʂpuxma* ‘shoulder’). According to Bielmeier et al. (2018: 50, 52), in Skardu (Western Balti) and Khaplu (Eastern Balti) “*x* and *ɣ* are postvelar.” The most common *d-* reflex in WAT Purik lects is *ʂ-/r-*.

Amdo *g-/d-* reflexes tend to be velar fricatives, with the exception of Sertha and G.yukhog, which have uvular fricative reflexes. Sertha adds an epenthetic schwa after the *g-* uvular fricative reflex, as in *ʂəjak* ‘yak’.

In Gyalrongic languages, *g-/d-* loanwords have a uvular fricative preinitial *ʂ-/χ-* except for *lonbut^{hi}* ‘elephant’, in which the preinitial has been lost, and the forms for the etymon ‘like’, which all have the reflex *r-*⁶. The complementary distribution resulting from Sa-skya Paṇḍita’s

⁶ There are many instances in Tibetic and Gyalrongic lects where *g-/d-* preinitials have been lost altogether, but this is a trivial sound change that needs no further discussion.

Written Tibetan	Gloss	Balti	Amdo	G.yukhog	Stau	Geshiza	Khroskyabs	Japhug
<i>gtsang ma</i>	clean	<i>ḥtsaṃma</i>	<i>htsaṃma</i>	<i>ḥtsəmə</i>	<i>ḥtsōmə</i>	<i>atsəŋ</i>		<i>χtso</i>
<i>gser</i>	gold	<i>xser</i> (title: Amacha)	<i>γser</i> (Arik)	<i>ḥser</i>	<i>ḥsər</i>	<i>asər</i>	<i>χsə̀r</i>	<i>χsər</i>
<i>gso</i>	feed	<i>χsu</i>	<i>γsu</i> (Arik)	<i>ḥsu</i> ‘eat (hon.)’	<i>ḥsu</i> ‘eat (hon.)’			<i>χsu</i>
<i>gsur</i>	offering for <i>yi dwags</i>			<i>ḥsər</i>	<i>ḥsər</i>			
<i>gzah ḥkhor</i>	week	<i>γza</i> ‘rainbow’	<i>hza</i>	<i>ḥzænk^hor</i>	<i>ḥzænk^hor</i>			
<i>gnad ḥgag</i>	problem			<i>ḥnængəq</i>	<i>ḥnanga</i>			
<i>gnaḥ</i>	long ago		<i>γna</i> (Arik)	<i>ḥnæ</i>	<i>ḥnæ</i>		<i>ḥnâ</i>	
<i>glu</i>	song	<i>xlu</i>	<i>ḥlə</i> (Sertha)	<i>ḥlə</i>	<i>ḥlə</i>	<i>alə</i>	<i>ḥlâ</i>	
<i>gla</i>	hire/salary	<i>xla</i> ‘tribute’	<i>γla</i> (Rkangtsha)	<i>ḥlæ</i>	<i>ḥlæ</i>	<i>ala</i>		
<i>gshal</i> ‘wash’	flow			<i>ḥçer</i>	<i>ḥçer</i>			
<i>gdam pa</i>	leader			<i>ḥdæmbæ</i>	<i>ḥdæmbæ</i>			
<i>gdan</i>	carpet	<i>γdan</i> ‘area’	<i>γdan</i> (Rkangtsha)	<i>ḥden</i>	<i>ḥden</i>			<i>ḥdɔn</i> ‘cushion’
<i>g.ya</i>	rust	<i>xja</i>			<i>ḥja</i>			<i>ḥja</i>
<i>glang chen</i>	elephant	<i>xlaṃpotfo</i>	<i>γlaṃtɕ^hen</i> (Rkangtsha)	<i>ḥləpotɕ^he</i>	<i>ḥləpɔtɕ^he</i>	<i>aləṃbutɕhe</i>	<i>ḥləṃbutɕ^he</i>	<i>ləṃbutɕ^hi</i>
<i>gcan gzan</i>	carnivorous animal	<i>xtfanzan</i> ‘snow leopard’	<i>γtɕandzən</i> (Arik)	<i>ḥtɕændzən</i>	<i>ḥtɕændzən</i>	<i>xtɕænzæn</i>		<i>χtɕɔnzɔn</i>
<i>g.yu</i>	turquoise	<i>xju</i>	<i>ḥjɔ</i> (Sertha)	<i>ḥjə</i>	<i>ḥjə</i>		<i>ḥjikæ̀r</i>	<i>ḥju</i>
<i>g.yag</i>	yak	<i>xjaq</i> ‘crossbread’	<i>ḥjək</i> (Sertha)	<i>ḥjiqɔ</i>	<i>ḥja</i>	<i>ajə</i>		
<i>gzig</i>	leopard		<i>γzək</i> (Arik)	<i>ḥzuq</i>				

Table 1: Words with g-

Tibetan	Gloss	WAT	Amdo	G.yukhog	Stau	Geshiza	Khroskyabs	Japhug
<i>dpe</i>	metaphor/example	<i>xpe</i> (Balti)	<i>χwe</i> ‘story’ (Ndzo)	<i>κwe</i>	<i>κpi</i>	<i>api</i>	<i>χpi</i>	<i>χpi</i>
<i>dpaḥ bo</i>	hero	<i>ṣpao</i> (Chiktan Purik)	<i>χwawo</i> (Rkangtsha)	<i>κpawo/κwawo</i>	<i>κpawu</i>		<i>χpawu</i>	
<i>dpon</i>	official	<i>ṣponbo</i> (Chiktan Purik)	<i>χwonpo</i> (Rkangtsha)	<i>κwən</i>	<i>κpən</i>	<i>apən</i>	<i>χpən̂</i>	<i>χpiun</i> ‘monk’
<i>dmaḥ</i>	low		<i>γma</i> (Arik)	<i>κmæ</i> ‘low altitude’	<i>κmæ</i> ‘low altitude’			<i>κma</i>
<i>dmag mi</i>	soldier	<i>ṣmaqmi</i> (Kargil Purik)	<i>hmak</i>	<i>κmɑqmjɑ</i>	<i>κtami</i>	<i>ɑmɑmɑ</i>	<i>κmæ̂κ</i>	<i>κtɑκmi</i>
<i>dgah</i>	like	<i>rgak^han</i> ‘friend’ (Chiktan Purik)	<i>hga/rga</i>	<i>rgæ</i>	<i>rgæ</i>	<i>rga</i>	<i>rga</i>	<i>rga</i>
<i>dpung pa/ma</i>	shoulder	<i>ṣpuxma</i> (Balti)	<i>χwunja</i> (Rkangtsha)	<i>κwō</i>	<i>κpō</i>			

Table 2: Words with *d-*

Law can be seen in Gyalrongic *g-/d-* loanwords. For example, if Tibetan loanwords in Stau have the preinitial *ʁ-* followed by a labial, then *ʁ-* corresponds to Tibetan *d-*, otherwise the preinitial *ʁ-* corresponds to Tibetan *g-* (compare also Lai 2017: 56). In Geshiza, Proto-West-Gyalrongic **ʁ-* has evolved into *ɣ-* (but *ʁ* becomes *ɣ* as an initial; e.g., Stau *ʁə* ‘head’, Geshiza *ɣuə* ‘head’), and Proto-West-Gyalrongic **ɣ-* has changed into *w-* (Honkasalo et al., to appear). The preinitial *x-* in the Geshiza word *xɬænzæn* ‘carnivorous animal’ goes against this, demonstrating that some loanwords from Tibetan were borrowed into Geshiza following this sound change.

All Tibetic and Gyalrongic lects have the reflex *r-* for *dgah* ‘like’, though some Amdo lects use *h-* (Hua’s 2001 transcription).

4. Huang’s hypothesis of uvular preinitials in Old Tibetan

Huang (2012) concludes that uvular preinitials were present in Old Tibetan based on the following evidence. Firstly, Huang demonstrates the possibility that uvular preinitials were inherited from Proto-Tibeto-Burman. She uses evidence from cognate sets that have uvular stops, nasals and fricatives in Ngwi, Qiangic, and Tibetan languages to bolster this claim (2012: 158–167). It is important to note at this stage that Huang is referring to both uvular initials and preinitials. She only gives examples of uvular initials in Ngwi and Qiangic and gives examples of uvular initials and preinitials in modern Tibetic lects⁷. Secondly, Huang shows that the majority of uvular fricative preinitials in Amdo (in particular the Skakhog/Skachu lect) correspond to Old Tibetan *g-/d-* (2012: 167–168). Thirdly, she demonstrates how Tibetan *d-* comes from *g-* (2012: 168–169). Fourthly, Huang proposes that Tibetan *g-* (which later splits into *g-* and *d-*) comes from the merger of **g-* and **ɣ-* (2012: 169–170). Her conclusion from all of this evidence is that at least some portion of uvular fricative preinitials in Tibeto-Burman come from a Proto-Tibeto-Burman voiced uvular stop preinitial **ɣ-* (2012: 170–172).

We can summarize Huang’s logic as follows, illustrated in Figure 1. Since many Tibeto-Burman languages have uvulars in cognate sets, and most of these languages are too widespread geographically to have borrowed uvulars from each other, Proto-Tibeto-Burman probably had a voiced uvular stop **ɣ*. This voiced uvular stop has voiced and voiceless stop and fricative reflexes in the Tibetan, Qiangic, and Ngwi subgroups. If the initial **ɣ* existed, and modern Tibetic and Gyalrongic lects have uvular reflexes for *g-/d-*, then we can postulate the preinitial **ɣ-* in Proto-Tibeto-Burman. Thus, for Tibetan *g-/d-*, Proto-Tibeto-Burman **ɣ-* and **g-* merged to become *g-/d-* in Old Tibetan, but became a uvular fricative preinitial (among other reflexes) in some modern Eastern Tibetan lects. By implication, Qiangic languages that have Tibetan loanwords with a uvular preinitial for *g-/d-* borrowed the uvular preinitial with the entire word.

Huang’s postulation of a uvular preinitial to Tibeto-Burman assumes the that history of uvular initials and preinitials are tied together; however, this may not be the case. The conflation of uvular initials and preinitials does not strengthen her argument, but it does not negate it either.

A central part of Huang’s proposal hinges on whether or not all modern Tibetan lects descend directly from Old Tibetan (the language that was reduced to writing in 650 AD), or if at least some of the lects (such as Amdo, Kham, and Balti) split off before the creation of Old Tibetan. Huang favors the later hypothesis, claiming that Central Tibetan (卫藏, or U-Tsang) was

⁷ Thanks to an anonymous reviewer who pointed out that it is important to indicate when Huang is referring to uvular initials, preinitials, or both. There is an absence of uvular codas in her data.

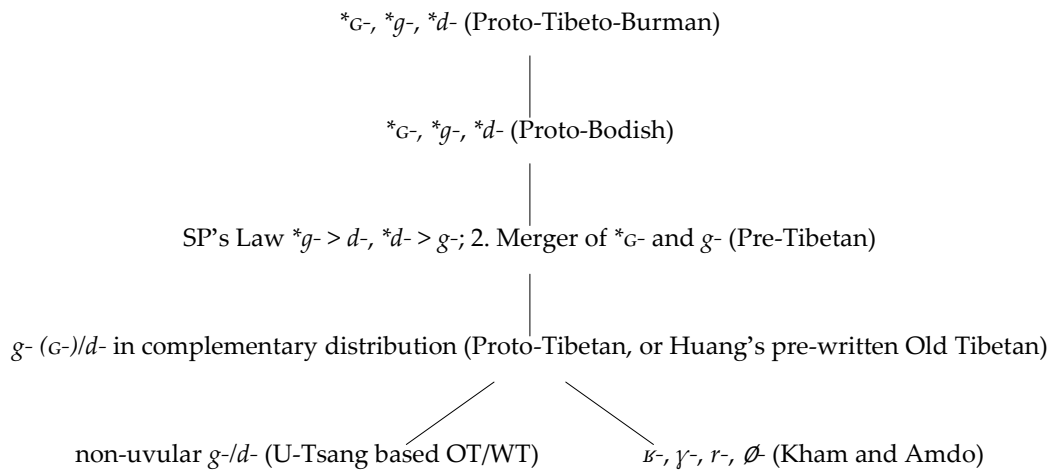


Figure 1. Huang's hypothesis

the variety used to create Old Tibetan (and thus also Written Tibetan), and that other lects split off before Old Tibetan was reduced to writing (2012: 166). Concerning uvulars in Tibetic languages, Huang states,

Naturally, it can also be said that uvulars in [modern] Tibetan vernaculars are the reflexes of the consonants that existed in Tibetan during the pre-Tibetan period, because the literary language of Tibetan reflects the phonetic forms of Old Tibetan — the Tibetan written script has no letters that reflect uvulars. The Tibetan written script was created in the seventh century; at that time, the Tibetan language was already divided into three lects: U-Tsang, Kham, and Amdo. The Tibetan written script was created on the basis of the Central Tibetan lect, that is, the U-Tsang lect. The Tibetan written script reflects that the U-Tsang dialect did not have uvulars at that time, and some of the Amdo and Kang lects have preserved uvulars to this day, which come from the Old Tibetan language before the seventh century (Huang 2012: 166, translation mine)⁸.

Of course, Huang is conflating Written Tibetan (or Classical Tibetan) to Old Tibetan (fortunately the two only differ slightly phonologically; Hill 2019b: 7), and it is not entirely clear if by 古藏语 'Old Tibetan' she actually means something like 'Pre-Tibetan'. Regardless, her claim places the split of Common Tibetan, the proto-language reconstructable from modern Tibetan lects by using the comparative method (Hill 2010: 112), before the writing of Old Tibetan, and that the written language of Tibetan descends from an old form of U-Tsang.

One way to avoid placing uvular preinitials as a distinctive uvular category in Old Tibetan is to use Huang's hypothesis, but modify it slightly. Pre-Tibetan would have uvular preinitials, but when certain lects like Amdo split from Pre-Tibetan they retained uvular preinitials, but Old Tibetan simply lost them. This modification of Huang's hypothesis is displayed in Figure 2.

In either scenario, Huang's hypothesis breaks up Tibetan (into Central, Amdo, and Kham) before Tibetan is reduced to writing. This hypothesis is an *a priori* hypothesis (unlike Hill's hypothesis discussed below in Section 5), because we know that the area for the Tibetan speaking population was small (and still not far from the Urheimat of the Yarlung Valley) and did not expand until after the writing system was created in 650 AD (Hill 2019b: 3; Bialek 2018).

⁸ “自然也可以说，藏语土语的小舌音是藏语在前藏文时期就有的辅音的遗存，因为反映古藏语语音面貌的藏语书面语——文并无反映小舌音的字母，藏文创制于七世纪，其时藏语已分化为卫藏、康、安多三大方言，藏文是以中部藏语即卫藏方言为基础创制的，藏文反映那时卫藏方言已无小舌音，而来自七世纪前更古藏语的部分安多和康方言清小舌塞音还保留至今。” I am translating 藏文 as 'Tibetan written script' and 古藏语 as 'Old Tibetan'.

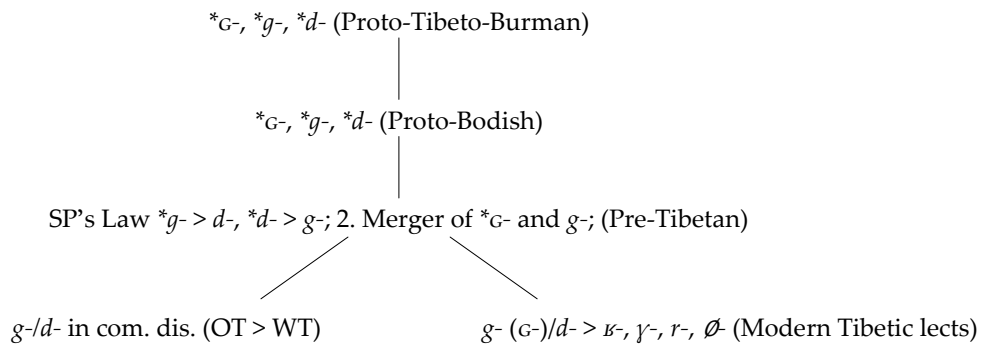


Figure 2. Huang's hypothesis modified

5. Hill's hypothesis

Hill claims that uvulars are not distinct in Old Tibetan (Hill 2010: 120). Based on Peiros and Starostin's Law, uvulars as a phonological category were lost through the merger of velars and uvulars in Proto-Bodish (or possibly Pre-Tibetan). At this point of in the discussion, I need to clarify that my focus is still on uvular preinitials. Peiros and Starostin's Law and Hill's claim deal with uvulars in general, that there was no uvular distinction in Proto-Bodish and Old Tibetan, respectively; but it stands to reason that if a uvular distinction did not exist in Old Tibetan, then it would not exist in either the initial or preinitial position.

Uvulars and velars were distinct in Trans-Himalayan (Hill 2019b: 32-34, 45). For our purposes, 'Trans-Himalayan' can be equated with Huang's 'Tibeto-Burman'. Hill's claim that there are no uvulars in Old Tibetan is supported by the fact that uvulars "were not orthographically distinguished from the velars" (Hill 2010: 120). The explanation that Hill gives for the presence of uvulars in modern Tibetic languages is that they are contact induced through Qiangic and/or Mongolic languages; languages in a region that he considers a "uvular-prone Sprachbund" (Hill 2010: 120).

Figure 3 is a summary of Hill's hypothesis regarding *g-/d-*⁹. Uvulars are lost long before Old Tibetan emerges (Peiros and Starostin's Law). Next Sa-skya Paṇḍita's Law, also occurring

⁹ To be accurate, Hill's (2010) hypothesis is not just regarding *g-/d-*, but regarding both uvular initials and pre-initials to counter Sun (2003: 782), who reports that in the Tibetic lect Zhongu certain "instances of uvulars are of mysterious origin, as they occur in lexical items of unknown ancestry (e.g., /χeŋə/ 'hole', /q^hetʂə/ 'neck'; /qeŋe/ 'hook', and in the suffixes /-qe/, /-qo/ attached to many nouns, as in /neqe/ <rna.ʔ> 'ear', /zeqo/ <zhwa.ʔ> 'hat')." Sun goes on to write, "It would be rash to attribute them indiscriminately to Qiangic substratal or areal influences, as very few of them are recognizable as Qiangic loanwords" (2003: 782-783). Sun also cites examples of uvular onsets in modern Zhongu words with clear origin in Old Tibetan. Sun suggests that this evidence (along with evidence of uvulars in other Tibetic lects) indicates the possibility that uvulars "predate standard written Tibetan." There are a couple of problems with this thinking. Firstly, how is the presence of modern uvulars more indicative of predating "standard written Tibetan" than coming after "standard written Tibetan" through natural sound change or through contact? Secondly, the presence of uvulars could still be the result of areal influences (such as from Qiangic languages) without the cited words being direct loanwords from Qiangic languages. For example, as long as some loanwords with uvulars were borrowed into Zhongu, these uvulars could effect change in other words, e.g., the ones Sun cites. Thirdly, there is the problem of the whether or not Old Tibetan had uvulars. While Sun can propose this possibility, most scholars agree that uvulars were not present in Old Tibetan. Finally, if Sun is proposing the same timing of the break up of Tibetan as Huang (i.e., before the Old Tibetan orthography was created), then once again there is the historical problem; all the evidence points to Tibetan speakers not leaving central Tibet (Yarlung, Ngaspo, and Rtsangbod) until after 650 AD (Bialek 2021: 342-351). Before 550 AD there is evidence that Pre-Tibetan speakers had already spread across the Dbus-Gtsang region (Bialek 2021: 342-351), but this is still not the eastward expansion that would come after 650 AD.

before Old Tibetan, puts $g-/d-$ in complementary distribution in Old Tibetan, which continues on later in Written Tibetan. The uvular fricative reflexes of $g-/d-$ found in Tibetic lects have been introduced through Qiangic (particularly Gyalrongic) or Mongolic languages. Hill places Peiros and Starostin's Law chronologically before Sa-skya Paṇḍita's Law, thus in Figure 3, Peiros and Starostin's Law is placed at the Proto-Bodish level and Sa-skya Paṇḍita's Law is placed at the Pre-Tibetan level. This is mainly for expository purposes; both laws could be placed at the same level at this stage of research, whether that be Proto-Bodish or Pre-Tibetan, as long as Peiros and Starostin's Law is placed before Sa-skya Paṇḍita's Law.

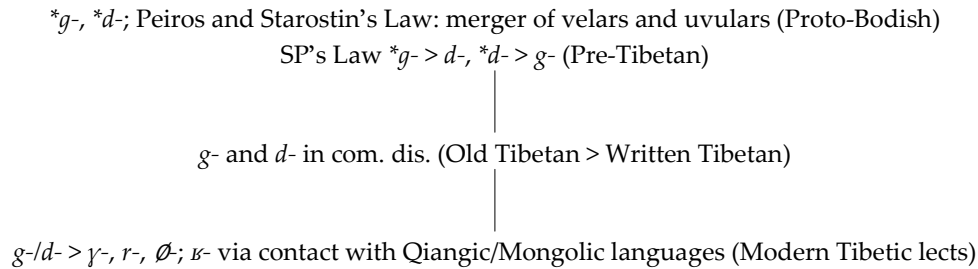


Figure 3. Hill's hypothesis

Hill's hypothesis is opposite to that of Huang's in that he claims that Common Tibetan descends from Old Tibetan (Hill 2010: 112). The Urheimat of Tibetan is the Yarlung Valley, and when the Old Tibetan orthography was created (650 AD) Tibetan speakers had not yet spread far from this region (see Bialek 2018 and 2021 for arguments in favor of this theory).

According to Hill's hypothesis, Gyalrongic languages borrowed words with $g-/d-$ into their languages and uvularized $g-/d-$ independently from Tibetic languages, but through contact with Gyalrongic languages Tibetic languages also uvularized $g-/d-$. In other words, $g-/d-$ (perhaps as $\gamma-/x-$) is substituted with $\nu-/χ-$ in Gyalrongic languages, and then these Gyalrongic languages have influenced some Eastern Tibetic lects to also pronounce $g-/d-$ reflexes as uvular fricatives. Although this is a reasonable hypothesis, Hill does not take into account the uvular fricative preinitials in Balti and WAT lects, which could not have developed into uvulars through contact with Mongolian and Qiangic languages as they are greatly removed geographically to the west of these languages. See Tables 1 and 2 for examples of uvular reflexes of $g-/d-$ in Balti and WAT.

6. Explanation 3

'Explanation 3' builds on Hill's hypothesis to address the real issue at stake: is the uvularization of $g-/d-$ regular or sporadic? Explanation 3 assumes that Hill is correct in the claim that there were no uvular preinitials in Old Tibetan, and claims that there is a regular sound change from $g-/d-$ to velar/postvelar/uvular fricatives (except before velar initials, where the change is to $r-$) in Amdo lects and to uvular fricatives Gyalrongic lects. For WAT lects, $g-$ regularly changes to velar/postvelar/uvular fricatives, but $d-$ changes to velar/postvelar/uvular or $\xi-/r-$. In the case of Amdo lects, the uvularization was triggered by contact with Qiangic and Mongolic, and in the case of WAT, the uvularization was triggered by contact with Indo-Aryan or Iranian languages.

However, the main emphasis of Explanation 3 is not on whether or not the change from $g-/d-$ to uvular fricatives was contact induced, but rather that the changes were regular and not

sporadic. More research would be needed to reveal what motivated the regular change; whether contact induced or language internal reasons. In fact, since we know that the change has been regular, the issue of language contact induced change is not a highly important question, and may actually be unanswerable. If the change was completely sporadic there would be more motivation for trying to find an answer for the irregularities, e.g., language contact or inter-dialectal borrowing. The inability to point to the exact internal or external causes for a particular sound change is common in historical linguistics¹⁰.

Explanation 3 suggests that uvular preinitials in Balti and WAT could be the result of language contact with Indo-Aryan or Iranian languages. The main evidence for this explanation is found in the Balti data found in Table 1. As was mentioned above, Bielmeier et al. (2018: 50, 52) state that in Skardu (Western Balti) and Khaplu (Eastern Balti) “*x* and *ɣ* are postvelar.” The postvelar/uvular nature of these preinitials in Balti is confirmed by Caplow (2016: 208) with her transcriptions of *ʰɖzar.ba* ‘to drip’ from *ɦɖzar ba/gzar* or *ba/bzar ba*, *ʰχnot.pa* ‘to harm’ from *gnod pa*; and especially *χlat.pa* ‘brain’ from *klad.pa/glad pa*, which has a minimal contrast with *ʰχlat.pa* ‘to be tired’ from *glad pa*. Note that *χlat.pa* ‘brain’ and *ʰχlat.pa* ‘to be tired’ can still be considered in minimal contrast because the stress that she transcribes is deemed “marginally phonemic” (Caplow 2016: 191).

That being said, although uvulars are present in Indo-Aryan languages, for most of those languages they are not a prominent feature. For most of the the languages surrounding Balti, uvulars are only marginally phonemic, and usually they only have one unaspirated voiceless uvular stop (e.g., *q* in Hindustani, Khowar, Phasto, Shina), which does not occur in the preinitial position. However, Wakhi has *χ*, *ɖ*, and *q*, and thus it may be a candidate for inducing uvularization of preinitials in Balti. The isolate Burushaski could also be a source of uvulars, having *ɖ*, *qʰ*, and *q*¹¹. Recent research is pointing to the heavy influence of languages like Burushaski on certain WAT lects (Kogan 2019). In the consideration of this hypothesis, it is important to bear in mind that for many of these Indo-Aryan languages, *q-* is only found in loanwords from Arabic¹².

Explanation 3 proposes the following sound changes; note that the preinitial *r-* is also included, since in some situations there is a merger of *r-* with *g-/d-* in modern Tibetic lects:

1. Pre-Tibetan: **g-* > *ɣ-/x-*.
2. Pre-Tibetan: **d-* remains a dental.
3. Saska Paṇḍita’s Law: preinitials *ɣ-/x-* and *d-* become allophones of the same phoneme.
4. After the creation of Old Tibetan, but before the breakup of Tibetic languages, *d-* > [^h*r-*] (allophone of *x-/ɣ-*).
5. Break-up of Tibetan (post-Old Tibetan): preinitial *g-* in Balti and Amdo became uvular from language contact or remain velar, and became uvular for Tibetan loanwords with *g-/d-* in Gyalrongic.
6. The preinitial *d-* > [^h*r-*] (allophone of *x-/ɣ-*) > *ʂ*, *z*, *r* (before velars, and for WAT all environments), but became uvular or velar in other lects (except for before velars), due to language contact.
7. *r-* > *x-/ɣ-* > *χ-/ɖ-* (in many lects, also contact induced and a tendency towards merging all preinitials into velar/uvulars). *r-* > *r-* elsewhere.
8. A more recent change is *x* replacing *ʂ-*, e.g., Balti *xpe* ‘metaphor’.

¹⁰ This discussion benefited from comments by an anonymous reviewer.

¹¹ I do not have enough information on the Gujar and Domaaki languages to know whether or not they have uvulars.

¹² Thanks to Arnaud Fournet for reminding me of this constraint.

7. Explanation 4: *h* and *g-/d-* in velar and uvular free variation

If all Tibetic lects must descend from Old Tibetan, and if there are no uvular phonemes in Old Tibetan, and if uvulars in western Tibetic lects didn't develop through language contact, then perhaps there is another explanation to be explored for the uvular reflexes of *g-/d-* in Tibetic cognates and Gyalrongic loanwords. This explanation is that [ɣ-] and [ʁ-] were free variants of *g-/d-*. Voicing assimilation also occurs, producing [x-] and [χ-] before voiceless initials and [ɣ-] and [ʁ-] before voiced initials. This fourth explanation is illustrated in Figure 4.

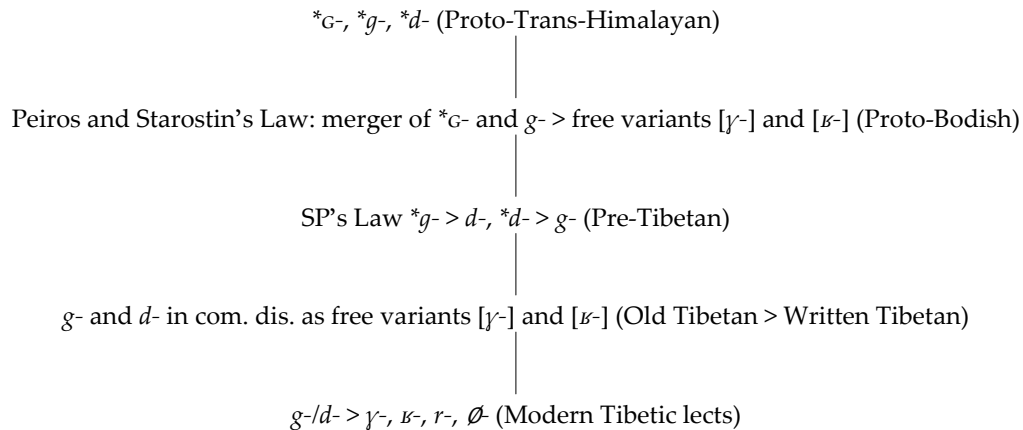


Figure 4. Explanation 4

As Figure 4 illustrates, the uvular reflexes of *g-/d-* in Tibetic descend directly from the Old Tibetan free variation of [ɣ-] and [ʁ-], which also explains why there are also velar reflexes for *g-/d-*. This explanation proposes that the reflexes of *g-/d-* found in Gyalrongic languages were borrowed in from Tibetic languages along with the entire word, albeit undergoing other changes specific to each language and lect.

This line of reasoning could also entail that *h* (ɣ) is in complementary distribution with C2 *g-* and *d-*, since *h* as [ɣ] never occurs in C2. *h-* does occur as N- in C2; the sound change from ɣ- > N- occurring before Old Tibetan (Hill 2005: 114–115, 127). Recently, Hill (2019a) has argued that *h* as a phoneme and *g-* as a preinitial/prefix have a common origin, and that *g-* has derived from *h*¹³. This would then imply, that while still not distinctive, there was free variation between [ɣ-] and [ʁ-] for *h*. Hill argues against *h* as a phonemic uvular consonant (Hill 2005, 2009). However, his claim still leaves open the possibility of *h* having the free variants of [ɣ-] and [ʁ-]; Hill admits that *h* “may have been articulated phonetically as a uvular” (2009: 124). While Explanation 4 is not contingent on *h* having uvular reflexes, it could be used to explain why some Tibetic languages also have uvular reflexes for *h* as a simple initial.

The problem with Explanation 4, while it is a possibility, is that it is not scientifically verifiable. There is just no way to falsify a proposal of free variation for a dead language with only written texts the sources of data¹⁴.

8. Conclusion

In summary, this article has explored four explanations for why there are uvular reflexes of *g-/d-* in some modern Tibetic words and Gyalrongic loanwords. Given Peiros and Starostin's

¹³ For a counter argument see Bialek (2020: 317, fn. 134).

¹⁴ I am grateful to an anonymous reviewer for this criticism.

Law, Sa-skya Paṇḍita's Law, and the data given in Section 3 (Tables 1 and 2), the most likely hypotheses are Hill's hypothesis in Section 5 (the result of contact with Mongolic and/or Qiangic languages), or Explanation 3 in Section 6, which proposes regular sound changes. Huang's hypothesis (explained in Section 4) rests too heavily on the hypothesis that Common Tibetan broke away from Old Tibetan before it was reduced to writing, a hypothesis that has been falsified. Explanation 4 (explained in Section 7), which states that *ḥ* and *g-/d-* were in velar and uvular free variation in Old Tibetan is not falsifiable, and thus cannot undergo scientific verification. All in all, Explanation 3 provides the most satisfying explanation to date. Explanation 3 provides a tentative set of sound changes and embraces the possible influence of Qiangic and Mongolic in the west and Indo-Aryan and Iranian in the east, but does not rely on contact with these non-Tibetic languages to support it.

References

- Baxter, William H., Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford University Press.
- Baxter, William H. III. 1992. *A Handbook of Old Chinese Phonology*. *Trends in Linguistics, Studies and Monographs* 64. Berlin: Mouton de Gruyter.
- Bialek, Joanna. 2018. The Proto-Tibetan clusters *sl-* and *sr-* and the periodisation of Old Tibetan. *Himalayan Linguistics* 17(2): 1–50.
- Bialek, Joanna. 2020. Old Tibetan verb morphology and semantics: An attempt at a reconstruction. *Himalayan Linguistics* 19(1): 263–346.
- Bialek, Joanna. 2021. Naming the empire: From Bod to Tibet. A philologico-historical study on the origin of the polity. *Revue d'Etudes Tibétaines* 61: 339–402.
- Bielmeier, Roland, Felix Haller, Katrin Häsler, Brigitte Huber, Marianne Volkart. 2018. *Comparative dictionary of Tibetan dialects (CDTD): Vol. 2: Verbs*. De Gruyter Mouton. doi: 10.1515/9783110554076.
- Bielmeier, Roland, Felix Haller, Katrin Häsler, Brigitte Huber, Marianne Volkart. Forthc. *Comparative dictionary of Tibetan dialects (CDTD): Vol. 1: Nouns*. De Gruyter Mouton.
- Caplow, Nancy J. 2016. Reconstructing stress in Proto-Tibetan: Evidence from Balti and Rebkong Amdo. *Linguistics of the Tibeto-Burman Area* 39(2): 180–221.
- Gates, Jesse P. 2021. *A grammar of Mazur Stau*. Paris: École des Hautes Études en Sciences Sociales dissertation.
- Hill, Nathan W. 2005. Once more on the letter ' '. *Linguistics of the Tibeto-Burman Area* 28(2): 111–141.
- Hill, Nathan W. 2009. Tibetan <*ḥ*> as a plain initial and its place in Old Tibetan phonology. *Linguistics of the Tibeto-Burman Area* 32(1): 115–140.
- Hill, Nathan W. 2010. An overview of Old Tibetan synchronic phonology. *Transactions of the Philological Society* 108(2): 110–125.
- Hill, Nathan W. 2011. An inventory of Tibetan sound laws. *Journal of the Royal Asiatic Society* 21(4): 441–457.
- Hill, Nathan W. 2014. Tibetan. In: Rochelle Lieber, Pavol Štekauer (eds.). *The Oxford Handbook of Derivational Morphology*: 620–630. Oxford University Press.
- Hill, Nathan W. 2019a. The derivation of the Tibetan present prefix *g-* from *ḥ*. *Acta Orientalia Academiae Scientiarum Hungaricae* 72(3): 325–332.
- Hill, Nathan W. 2019b. *The historical phonology of Tibetan, Burmese, and Chinese*. Cambridge University Press.
- Honkasalo, Sami. 2019. *A grammar of Eastern Geshiza: A culturally anchored description*. Helsinki: University of Helsinki dissertation.
- Honkasalo, Sami, Jesse P. Gates & Mathieu Beaudouin. To appear. *Preinitials and their development in Horpa and Tangut*. Ms.
- Hua, Kan. 2001. *Zangyu Anduo fanyan cihui* [藏语安多方言词汇 = A vocabulary of Amdo Tibetan dialects]. Lanzhou: Gansu minzu chubanshe (甘肃民族出版社).
- Huang, Bufan (黄布凡). 2012. Zàngmiǎnyǔ de xiǎoshéyīn [藏缅语的小舌音]. *Yǔyánxué Lùncóng* (语言学论丛) 45: 157–174.
- Jacques, Guillaume. 2008. *Jiāróngyǔ yánjiū* [嘉絨語研究 = Study on the Rgyalrong language]. Beijing: Minzu chubanshe.
- Jacques, Guillaume. 2014. On Coblin's law. *Studies in Chinese and Sino-Tibetan Linguistics* 2014: 155–65.

- Jacques, Guillaume. 2015–2016. *Dictionnaire Japhug-Chinois-Français, version 1.1*. Paris: Projet HimalCo. Online at: <http://himalco.huma-num.fr/>.
- Jacques, Guillaume. 2021. *A grammar of Japhug*. Berlin: Language Science Press.
- Jakobson, Roman, C. Gunnar M. Fant, Morris Halle. 1951. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge: MIT Press.
- Kogan, Anton. 2019. On possible Dardic and Burushaski influence on some Northwestern Tibetan dialects. *Journal of Language Relationship* 17/3–4: 263–284.
- Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*. Paris: Université Paris III dissertation.
- Sun, Jackson T.-S. 2003. Phonological profile of Zhongu: A new Tibetan dialect of Northern Sichuan. *Language and Linguistics* 4(4): 769–836.
- Wylie, Turrell. 1959. A standard system of Tibetan transcription. *Harvard journal of Asiatic studies* 22: 261–267.

Джесси П. Гейтс. Увуляризация *g-/d-* в тибетических языках.

В статье проанализированы четыре (относительно) взаимоисключающие гипотезы о возникновении увулярных пред-инициалей в языках тибетической группы, касающиеся в первую очередь современных увулярных отражений старотибетского *g-/d-*. Первая гипотеза принадлежит Хуан Буфань (Huang 2012), которая возводит увулярные пред-инициали к дотибетскому языковому состоянию, исходя из того, что распад тибетической ветви предшествует появлению письменного старотибетского языка в VII в. н.э. Второе объяснение принадлежит Н. Хиллу (2010), который утверждает, что увулярные пред-инициали не унаследованы от «дотибетского», а возникли благодаря контактам с цянскими и/или монгольскими языками. Третья гипотеза предполагает чисто фонетическое объяснение с постулированием регулярного развития *g-/d-* в велярные, поствелярные или увулярные фрикативные в диалектах Амдо (за исключением позиции перед велярными инициалами, где они развиваются в *r-*) и в увулярные фрикативные в гьярунгских диалектах. Наконец, четвертое объяснение заключается в том, что в старотибетском варианты *h* и *g-/d-* находились в состоянии свободного варьирования, т. е. современные увулярные пред-инициали восходят скорее к свободным аллофонам, чем к фонемам. Из этих гипотез первые три можно считать научными (поскольку они формально фальсифицируемы), четвертая (свободное варьирование) оказывается нефальсифицируемой. Анализ обстоятельств, связанных с распадом тибетического языкового единства, приводит к отказу от гипотезы Хуан, что оставляет в качестве единственно возможных объяснения Хилла и третье объяснение; при этом только последнее позволяет объяснить увуляризацию *g-/d-* в терминах регулярных фонетических изменений.

Ключевые слова: закон Сакья-пандиты; дотибетское языковое состояние; старотибетский язык; гьярунгские языки; увулярные рефлексy *g-/d-* в тибетских диалектах.