The development of zl in Tibetic languages

Kitty Wenying LIU | ORCID: 0009-0003-2430-1203
University of Cambridge, Cambridge, UK
wl399@cam.ac.uk

Received 21 August 2023 | Accepted 8 February 2024 |
Published online 3 April 2024

Abstract

This article assesses five proposals for the development of Tibetic zl, which has modern reflexes including Lhasa Tibetan /_(ⁿ)d/. My assessment considers on their ability to account for zl’s modern reflexes, their plausibility from the perspective of phonetics, and their congruence with typological observations. I conclude that, at present, Bialek (2018)’s proposal is the most plausible. However, future research may produce comparative evidence that support Gong (2016)’s proposal. At the end, I outline a methodology for investigating the role of functional load in Tibetic consonant cluster mergers.

Keywords


1 Introduction

The Tibetic onset cluster zl is pronounced /_(ⁿ)d/ in modern Lhasa Tibetan. This manifests word-initially as zla-ba ‘moon, month’ /_ⁿdawa ~ _dawa/ in word-initial position. Word-medially, the prenasalization is realized as

1 /_/ indicates low tone, and /‾/ high tone. I use a two-tone analysis of Lhasa lexical tone, following Bielmeier et al (2018), who in turn follow Qu & Tan (1983). This is a simplistic analysis, since other analyses of Lhasa tones often describe systems of four (e.g. Hoshi, 2003, contrasting on height and contour) or six tonemes (e.g. Huang, 1995, contrasting on height, contour,
nasalization of the nucleus or coda of the preceding syllable, as in \texttt{ni-}zla ‘the sun and the moon’ /\_\texttt{n}ǐ:\texttt{d}ə/, \texttt{rogs-}zla ‘help-mate’ /\_\texttt{r}on\texttt{d}a/, \texttt{nis-zlos} ‘twice-repeated’ /\_\texttt{n}ĩː\texttt{do}/ (examples from Gong, 2016). How zl became /\_\texttt{d}/ has long been discussed in the literature, since it is not immediately obvious how zl loses the fricative and lateral elements in its historical form, and gains nasal and plosive elements instead. There are numerous sound change pathways proposed in the literature attempting to account for some part of this development, with reference to both the Lhasa reflex /\_\texttt{d}/ and zl’s reflexes in other modern Tibetic languages. This article presents the first systematic evaluation of these proposals. I explore five different proposals for the development of zl (Sprigg, 1972; Hahn, 1973; Beyer, 1992; Gong, 2016; Bialek, 2018), and evaluate them in relation to their ability to account for zl’s modern reflexes, their phonetic plausibility, and their congruence with relevant typological tendencies for sound change.

The development of zl takes place within a general trend of consonant cluster reduction in Tibetic languages over the past millennium. This is apparent when one compares Tibetan orthography to their pronunciations in modern Tibetic (MoT). Tibetan orthography preserves complex consonant clusters, and is considered reflective of Old Tibetan (OT) phonology (Hill, 2010), which in turn is considered the common ancestor of MoT languages (Tournadre, 2014). As inferred from the orthography, OT phonotactics allow complex consonant clusters, especially at the syllable onset. Table 1 summarizes the phonotactic structure of an OT syllable, as inferred from the orthography, with examples of words that attest near-maximal syllable structure.

The extent and outcome of consonant cluster reduction differs between different MoT languages, with tonogenesis occurring in many innovative languages (Huang, 1995). For example, bsgrubs (perfective of ‘to complete’) from Table 1 is reflected as /b\dʐəp/ in Mkharmar, /\_\dʐəʔ/ in Kardze, and /\_\tup/ in Lhasa Tibetan (Bielmeier et al, 2018: 224). Cluster reduction and tonogenesis are closely linked, with tone distinctions “taking over” the lexical contrasts previously maintained by consonant cluster distinctions (Yu et al, 2010).
TABLE 1 Tibetan syllable structure as inferred from the orthography (adapted from Hill, 2010). The consonants individual shown here are phonemes rather than graphemes.

<table>
<thead>
<tr>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>G1</th>
<th>G2</th>
<th>V</th>
<th>C4</th>
<th>C5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obligatory?</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible graphemes</td>
<td>b</td>
<td>d, g, k, b, m, s, z, n, r, l</td>
<td>All consonant phonemes</td>
<td>j, r</td>
<td>w</td>
<td>a, e, i, o, u</td>
<td>g, n, d, b, m, s, y, r, l</td>
</tr>
</tbody>
</table>

Examples of words attesting near-maximal syllable structure:

<table>
<thead>
<tr>
<th>bsgrubs perfective of ‘to complete’</th>
<th>b</th>
<th>s</th>
<th>g</th>
<th>r</th>
<th>u</th>
<th>b</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>grwa ‘corner’</td>
<td>g</td>
<td>r</td>
<td>w</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Properties of consonants and consonant clusters (onset and/or coda voicing; whether the onset was sonorant or obstruent; whether the onset was complex; etc) provided conditioning environments for tonogenesis, which in turn caused consonants to be altered or lost because they were less necessary for maintaining lexical contrasts (Zhang, 1987; Huang, 1995). However, none of the five proposals for zl that I discuss consider tonogenesis in any detail, so I also largely ignore tonogenesis in this article, although the precise relationships between consonant cluster reduction and tonogenesis are doubtlessly fascinating.

Like other Tibetic clusters, zl has a range of MoT reflexes that depart from the historic form, with /_(n)d/ being the most innovative. The range of attested MoT reflexes for zl are shown in Table 2, grouped according to languages’ classifications in Bielmeier et al (2018). From the reflexes in Table 2, one can make the following three observations. Firstly, some MoT reflexes retain both /z/ and /l/, but these two segments always appear in reverse order compared to zl. This I have termed the Metathesis Problem. Secondly, many MoT reflexes have an excrescent /d/ compared to ot zl. This I have termed the Excrescence Problem. Thirdly, some Central Tibetic languages have a prenasalized reflex, including /_(n)d/ in Lhasa Tibetan as discussed above. This may be termed the Prenasal-
TABLE 2  Some attested pronunciations of zl, organised by language subgroup, and labelled with some of the languages they occur in (data from Sprigg, 1972; Bielmeier et al, 2018: 468–469; Gong, 2016; Bialek, 2018)

<table>
<thead>
<tr>
<th>Western Archaic Tibetan</th>
<th>Central Tibetan</th>
<th>Hor Tibetan</th>
<th>Kham Tibetan</th>
<th>Amdo Tibetan</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ldz ~ Iz/ (Balti, Sapi)</td>
<td>/(_n)d/ (Lhasa)</td>
<td>/dz/ (Gertse)</td>
<td>/dz/ (Nangchen)</td>
<td>/rdz ~ rd/ (Themchen, Golok)</td>
</tr>
<tr>
<td>/ldz/ (Kargil)</td>
<td>/_nd/ (Ngari)</td>
<td>/_dz/ (Dartsedo)</td>
<td>/hdz/ (Bayan)</td>
<td></td>
</tr>
<tr>
<td>/ld/ (Leh, Wanla)</td>
<td>/_l/ (Western Dropkas)</td>
<td>/rz/ (Labrang)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The literature has also remarked that in Lhasa Tibetan, zl shares the reflex /(_n)d/ with the clusters ld, md, and nd. zl does not share a reflex with other consonant-l (Cl) clusters, while ld, md, nd do not share a reflex with other consonant-d (Cd) clusters. This patterning is shown in Table 3. This shows that zl merged with ld, md, and nd.

The evidence from modern phonology in Table 3 does not shed any light on when this four-way merger occurred or on the internal ordering of the merger, since it seems unlikely for these four clusters to have merged at the same time, with zl being so phonetically different from nd and md. One can therefore hypothesize that either zl and ld merged first (since they share the lateral) and their reflex then merged with md and nd, or that ld, md, nd merged first (since they share the plosive) and their reflex then merged with zl after zl somehow gains a plosive element. The results of my evaluation for different proposals for the development of zl also sheds light on the internal ordering of this merger.

In Section 2 of this article, I discuss some pre-existing research and theoretical assumptions that make my discussions here possible. In Section 3, I evaluate the claims made by Sprigg (1972), Hahn (1973), Beyer (1992), Gong (2016), and Bialek (2018) based on their phonetic plausibility, their ability to account for zl’s range of MoT reflexes, and their congruence with typological observations about sound change. I conclude that Gong (2016) presents the most plausible account. In Section 4, I discuss the implications of this finding and prospects for future research. I specifically focus on the possibility of investigating the role of functional load in mergers involving zl, using the methodology of Wedel et al (2013a; 2013b) and Babinski and Bowern (2018). I outline the methodology for such a study, as well as difficulties that one would face.
### Table 3: Pronunciations of different consonant-\(l\) and consonant-\(d\) clusters in Lhasa Tibetan, illustrated with words and their modern Lhasa pronunciations (Hoshi, 2003; Sung & Rgyal, 2010; Gong, 2016)

<table>
<thead>
<tr>
<th>OT cluster</th>
<th>Lhasa reflex</th>
<th>Examples from Lhasa Tibetan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consonant-(l) clusters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zl</td>
<td>(_{n}d/)</td>
<td>⟨zla-ba⟩ /(_{n}dawa/⟩ ‘moon’</td>
</tr>
<tr>
<td>kl</td>
<td>(l)</td>
<td>⟨klog⟩ /(loː/⟩ ‘to read aloud’</td>
</tr>
<tr>
<td>gl</td>
<td></td>
<td>⟨gla⟩ /(la/⟩ ‘to hire’</td>
</tr>
<tr>
<td>bl</td>
<td></td>
<td>⟨blug⟩ /(luː/⟩ ‘to pour’</td>
</tr>
<tr>
<td>rl</td>
<td></td>
<td>⟨rlog⟩ /(loː/⟩ ‘to destroy’</td>
</tr>
<tr>
<td>sl</td>
<td></td>
<td>⟨sleb⟩ /(leː/⟩ ‘to reach’</td>
</tr>
<tr>
<td><strong>Consonant-(d) clusters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gd</td>
<td>(_{d}/)</td>
<td>⟨gdangs⟩ /(_{d}aŋ/⟩ ‘song’</td>
</tr>
<tr>
<td>bd</td>
<td></td>
<td>⟨bde-skyid⟩ /(_{d}iː/⟩ ‘peaceful comfort’</td>
</tr>
<tr>
<td>rd</td>
<td></td>
<td>⟨rdar⟩ /(_{d}a/⟩ ‘to sharpen’</td>
</tr>
<tr>
<td>sd</td>
<td></td>
<td>⟨sdod⟩ /(_{d}oː/⟩ ‘to stay’</td>
</tr>
<tr>
<td>ld</td>
<td>(_{n}d/)</td>
<td>⟨ldebs⟩ /(_{n}deː/⟩ ‘side’</td>
</tr>
<tr>
<td>md</td>
<td></td>
<td>⟨mda⟩ /(_{n}da/⟩ ‘arrow’</td>
</tr>
<tr>
<td>(n)d</td>
<td></td>
<td>⟨‘di⟩ /(‘di/⟩ ‘this’</td>
</tr>
<tr>
<td><strong>d</strong></td>
<td>(_{t}/)</td>
<td>⟨dag⟩ /(_{t}aː/⟩ ‘to purify, to clean’</td>
</tr>
</tbody>
</table>
2 Research assumptions

2.1 Tibetan orthography and OT phonology

Using orthographic representations in research about phonological change relies on the assumption that orthography reflects phonology in some systematic and knowable way. Recent research on Tibetan orthography suggests that Tibetan orthography reflects OT phonology in a generally transparent manner. Hill (2010) deduces a likely OT phoneme inventory from grapheme distributions. He concludes that the orthography transparently reflects OT phonology, with the following exceptions: obstruent aspiration is subphonemic in OT despite being encoded in the orthography, and the digraphs ⟨hr⟩ and ⟨lh⟩ represent /t/ and /l/ rather than true clusters. Bialek (2018: 22) also claims that, at the time of its invention in the 7th century, Tibetan orthography reflected phonetic similarities and differences more than phonemic contrasts. This mirrors findings on literacy and script invention, such as Olson (1996)’s claim that, when sound-based scripts are first invented, they are more likely to reflect perceptually salient phonetic differences than phonological patterns that require more metalinguistic introspection.

There has also been discussion surrounding the OT phonological form that underlies orthographic zl. Sprigg (1972) and Hahn (1973) both claim that zl represented underlying /zd/ in OT, but neither of them provide a systematic theory of how the orthography diverges from OT phonology. Without such a theory, the possible divergences become unconstrained, and posited divergences between orthography and OT phonology become ungrounded.

Other research suggests that zl does represent /zl/. Hill (2010) finds that OT has contrastive /l/ and /l̥/, and that OT onset clusters underwent voicing assimilation controlled by the root consonant (l, in the case of zl and sl). Building on this, one can conclude that /sl/ and /zl/ were not contrastive in OT (/s/ > /z/ by voicing assimilation), but /sl–zl/ is contrastive from /sl/. For clarity, I write these two phonologically contrastive clusters as /zl/ and /sl/. Hill (2010) argues that zl writes /zl/, while sl writes either /zl/ or /sl/. Hill uses evidence from causativising contexts, where s-/z- are allomorphs of the causative prefix: verbs with initial l /l/ receive z- (e.g. lo ‘a report’, zlo-ba ‘to report’), while verbs with initial lh /l̥/ receive s- (e.g. lhes-ma ‘a braid’, sle-ba ‘to braid’) (examples from Hahn, 2003: 145; cited in Hill, 2010). Hill suggests that both zl and sl may represent /zl/ because l /l/ is more common than lh /l̥/, so one should find more instances of /zl/ than /sl/. However, zl is less common than sl, suggesting

2 zl appears 153 times and sl appears 950 times in the OTDO corpus (Hoshi et al, 2006) under the ‘match partial’ search setting, retrieved November 2023.

Downloaded from Brill.com 04/16/2024 12:39:02PM
via Open Access. This is an open access article distributed under the terms of the CC BY 4.0 license.
https://creativecommons.org/licenses/by/4.0/
that not all instances of /zl/ are written /zl/, and some of them are likely written as /sl/. The fact that /zl/ and /sl/ may sometimes be homophonous is supported by the alternate spellings of some words, such as ‘moon, month’, attested as both /zla/ and /sla/ (Wang, 1996).

The above discussion shows that, while the topic of /zl/’s OT pronunciation is not without discussion, it likely did represent a voiced fricative-lateral sequence (/zl/). Thus, I take /zl/ to represent /zl/ in my discussion in Section 4.

Ascertaining the precise environments where /sl/ represents /zl/ rather than /sl̥/ requires detailed philological investigation, and may never be comprehensively clarified. The fact that /zl/ and /sl/ may sometimes be homophonous poses a challenge for the potential study on functional load described in Section 4.2.

In this article, I represent preconsonantal འ as prenasalized ང, following Jacques’s (2012) transcription system. There is considerable debate in the literature about the underlying phonological content of འ in OT. Comparative reconstruction using MoT forms suggest that OT འ was /ⁿ/ before consonants, /ɣ/ before vowels, and /∅/ when syllable-final. Hill (2019a) especially argues that འ in all three environments should be taken as /ɣ/ or /ɦ/ (transcribed ḥ) because the two other possible realisations can be derived from /ɣ~ɦ/. However, this debate lies outside of the scope of this article, and I primarily choose to represent འ as ང to highlight the merger between འད /ⁿd/ and མད /md/ due to shared nasality.

2.2 Characterizing historical explanation

What counts as explanation in historical linguistics? What are its limits and how do we provide diachronic explanations within those limits?

Explanation could entail identifying distinct synchronic states in a language and finding the most plausible pathway of development from an earlier state to a later state. Alternatively, it could entail answering why a change takes place here and now, rather than in a different language at a different time. This latter understanding of “explanation” directly invokes the “Actuation Problem” as articulated by Weinreich et al (1968). Walkden (2017) comprehensively assesses the Actuation Problem, and concludes that it is difficult or even impossible to solve endogenously (by appealing to linguistic features that surround the change) or exogenously (by appealing to contact, social factors, the acquisition process), because it is difficult to turn these bodies of information into a constrained account of why the same change has not taken place in a different language or at a different time.

Thus, it seems more realistic to account for the development of /zl/ in Tibetic through the first understanding of “explanation”, as a search for the most plau-
possible way to link two synchronic states. In this case, the relevant synchronic states are OT /zl/ and its various MoT reflexes. Thus, a successful account of the development of zl must be able to account for the cluster’s diverse modern reflexes. As discussed above, I summarise the divergences between /zl/ and its reflexes by the “Metathesis”, “Excrescence”, and “Prenasalization Problems”. It is not de facto an issue if any particular account for the development of zl does not attempt to address all three Problems, but it is problematic if its solution to one of the Problems is not borne up by zl’s modern reflexes. In my assessment, I also reference typological patterns where appropriate, because they form a broader crosslinguistic context for the observed reflexes of zl. Since typological trends are subject to chance and data availability, I only appeal to typological considerations when there is an overwhelming crosslinguistic trend for something, supported by a large sample size, and when this trend is echoed within Tibetic itself.

Another issue in diachronic explanation is that language change is shaped by the interaction of divers forces (e.g. Solé and Recasens, 2012), so it is methodologically crucial to both acknowledge the presence of interaction, and also to attempt to separate the effects of different factors in order to study the effects of each. Factors that affect phonological change may include phonetics, sociolinguistic variation, language contact, and phonological reanalysis of caregivers’ speech during first-language acquisition. Structural factors such as functional load may also influence the development of a language’s phonotactics or phonological inventory, such as Babinski and Bowern (2018)’s work on the effect of functional load on predicting patterns of phonological mergers.

In this article, I focus on the role of phonetics (articulatory, acoustic, and perceptual) in shaping the development of zl. Phonetics has the capacity for both describing the process by which sounds change, and for explaining why a sound changes in one direction rather than another, making it powerful for elucidating developments in historical phonology. For articulatory assessments, I loosely use the framework of Articulatory Phonology. Articulatory Phonology (Browman and Goldstein, 1992; Pouplier, 2020) is a schematized model of articulation that views speech production as the coordination and execution of articulatory gestures. The gestural targets required to make a given sound are synchronized in a gestural score. The execution of gestures depend on the interaction between neuromuscular commands from the brain, and physical factors such as inertia and each articulator’s mass. The relative timings of neuromuscular commands interact with articulators’ tendency towards their resting positions, thereby producing imprecision and variability in articulation (Clark et al, 2007: 85–86). Gestures for articulating different sounds may overlap, or one
gesture may under- or overshoot its articulatory target. Sound change occurs when phonetic variation becomes systematized over time (Salmons, 2021: 3–19). In acoustic and perceptual accounts, listeners drive sound change interpreting the speech signal as having different phonological content than what the speaker intended. This may be due to acoustic similarities between sounds (e.g. [m] and [n] have similar formant structure), or because listeners interpret certain phonetic material as the result of coarticulation, even though it was part of the intended phonological content (Ohala, 1981; Beddor, 2012). Articulatory and perceptual factors often work together, such as velar palatalization ([k] > [tf]) before front vowels, which is frequent both because articulating [ki] requires a similar gesture to articulating [tf], and because [k] and [tf] sound more similar before front vowels (Guion, 1998). It is true that articulatory and acoustic details differ cross-linguistically for what is classified as the same sound in the IPA (Ladefoged and Maddieson, 1996), so we cannot know the precise phonetics of Tibetic speakers from centuries ago. However, the fact that the human articulatory and auditory processing apparatuses are stable across populations means that it is still possible to assess proposals for sound change based on phonetic arguments.

Thus, to summarize from this section: I assess different accounts for the developments of zl with especial focus on their phonetic plausibility, their ability to account for zl’s modern reflexes, and whether they reflect relevant typological generalizations.

3 Proposals for the development of zl: Analyses

3.1 Sprigg (1972)
In an article on phonological reconstruction using Firthian Prosodic Analysis (FPA), Sprigg (1972: 565) analyses the diachronic development of zl, which is shown in Table 4 in Sprigg’s original formalisms. *lovS represents an underlying phonological form, and the forms in bold represent changes to its pronunciation. All changes between these forms use one of the three following processes, which Sprigg claims as established principles of Tibetic phonology: voicing assimilation within a cluster (sdlz > zdlz), the deletion of duplicate clusters in a cluster (dld > ld etc), and consonant insertion between continuant and obstruent segments (slz > sdlz etc).

*lovS, the “original form” of zl, is an onset cluster involving a liquid, occlusion, voicing, and alveolar friction (Table 5). The elements in *lovS represent “prosodies”, phonological features (e.g. tone, palatalization, voicing, occlusion, friction) which overlap in time. This reflects FPA’s commitment to
analyzing the speech stream “horizontally”, treating tone, voicing, friction, etc as “tiers” of speech, rather than the “vertical” segmentation of IPA-based phonologies, which bundles different phonological features together based on their temporal synchronization (/d/ is a bundle of voicing, alveolar contact, and plosion) (Moro and Ogden, 2022) (Figure 1). In this respect, FPA resembles Articulatory Phonology (AP), the phonetic framework that I use in this article (indeed, the diagram in Figure 1 comes from AP). However, FPA and AP’s different approaches to timing make them very different horizontal segmentation systems. AP is concerned with gestural coordination, so its horizontal tiers are different articulators, and AP pays close attention to the coordination of different articulator movements. In contrast, FPA is less concerned with coordination, and, because FPA represents phonic features relevant to the analysis at hand, allows different horizontal tiers that use the same articulators (e.g. plosion as distinct from friction) (Moro & Ogden, 2022). As part of its interest in the interaction between gestural targets, speech timings, and physical forces that act on articulators, AP is very aware of what phonetic sequences are possible to articulate at typical speech rates. Since FPA does not model these aspects of articulation, it allows Sprigg to postulate sound sequences like zdlz, which are difficult to interpret phonetically, since it is difficult to produce [zdlz] as a practicable syllable onset. However, as my discussion below shows, there are issues with Sprigg’s pathway even if one assumes the phonetic reality of clusters like [zdlz] at any point in zl’s development.
These methodological assumptions mean that the prosodies in *lovS do not have a specified order, so can be realized by any sequence of segments (in our “vertical” and segment-based phonology) which contain the phonological features specified in Table 5 are valid realizations of *lovS. Sprigg claims that *lovS is realized as slz, which becomes zdlz through the change processes that Sprigg pre-specifies. Here, some Tibetic languages lose the first of the duplicate z’s, while others lose the other z. Then, in the dlz pathway, the z is preserved into MoT, with modern reflexes like ldz and rdz. In the zdl pathway, the z is lost (Sprigg does not give a reason), leading to modern reflexes ld, nd, and d, which do not have a fricated element.

3.1.1 Metathesis
As discussed above, Sprigg only uses the three mechanisms of voicing assimilation, ex crescence, and deletion of duplicate sounds in his pathway, explicitly so as to avoid postulating metathesis, which “encourages one to ignore the articulatory aspects of the problem” (1972: 565). There are two issues with this
claim. Firstly, articulation-driven accounts of metathesis are amply possible. Secondly, despite acknowledging the Metathesis Problem (albeit opting for solutions without metathesis), Sprigg fails to substantively address the Problem. I address this second issue first, before returning to articulation-driven accounts of metathesis.

Sprigg addresses the Metathesis Problem by claiming that ʈ and ɬ never switch places, because there was a fricative on either side of the lateral to start with (*lovS > slz). Sprigg does not justify the presence of two fricatives, despite there being no two distinct fricatives in the orthography or any attested MoT reflex, so there is no motivation for positing slz other than to avoid metathesis of lateral and fricative. This is circular, and cannot adequately account the Metathesis Problem.

There is a wealth of research on phonetic analyses of metathesis, often incorporating both articulation and perception. Working within ap, Blevins and Garrett (2004) produce a typology of metathesis that divides metathesis into four types, where three are attributable to gestural overlap during articulation, and the fourth to the hearer misperceiving the ordering of sounds in the acoustic signal. The kind most relevant to the present discussion is “Perceptual Metathesis”, which has articulatory pinnings despite its name: some articulatory features (incl. labialization, nasalization, laterality, palatalization) persist for longer than the segment they are associated with, leading to a reanalysis of where the relevant segment. Hume (2004) proposes an “Indeterminacy-Attestation” Model of metathesis, which claims that metathesis occurs when the speaker produces an ambiguous speech signal (“indeterminacy”), which the listener interprets as a phonotactic structure that is attested in the language (“attestation”), but different from the phonotactic structure intended by the speaker. Factors creating an ambiguous speech signal include gestural overlap during articulation, acoustic resemblance between the phonetic realizations of different sequences, and a sequence being too brief to contain clear perceptual cues.

The metathesis of /zl/ > /lz/, which is proposed by Beyer (1992), Gong (2016), and Bialek (2018), appears easily explicable with reference to Blevins and Garrett (2004) and Hume (2004)’s work. The only articulatory difference between the target gestures [z] and [l] is the contact area between the tongue and hard palate, which is shown schematically in Figure 2.3 While Figure 2 is taken from English speakers, speakers of different languages show similar patterns of palatal contact for sounds that are transcribed using the same IPA symbols. Sometimes lateral segments do not have a complete central closure, but the lateral passageway is
of the tongue make contact with the palate, and the tongue is grooved to let through a central airflow. For [l], there is a central closure with the front of the tongue, and the side(s) of the tongue do not make contact with the palate, thus creating lateral airflow. From an articulatory perspective, the metathesis of /zl/ > /lz/ may arise from the persistence of laterality à la Blevins and Garrett (2004)’s Perceptual Metathesis, overlapping so much with [z] that [l] is interpreted as the first member of the cluster. Cross-linguistically, lateral segments do not always have a complete central closure, as exemplified in Figure 3. However, as long as the lateral escape is wider than the central escape, the airflow remains predominantly lateral (Ladefoged and Maddieson, 1996: 183–185), thereby creating the articulatory conditions for laterality, and the perceptual indeterminacy of the speech signal à la Hume (2004). This gives rise to the possibility of overlap between [z] and [l], where lateral passageway opens during [z]. This thereby creates the opportunity for reanalyzing [l] as the first member of the cluster. The fact that /lz or /ls/ are not preserved in Written Tibetan suggests that /lz/ was not a pre-existing cluster in OT, but this does not prohibit the appearance of /lz/ as a result of metathesis: /l/ is already allowed in wider than any central passageways, thereby ensuring that airflow is predominantly lateral (Ladefoged and Maddieson, 1996: 183–185).
cluster-initial position (lg, ldʐ etc), and /z/ in cluster-final position (rz, gz, etc), and the metathesis would not compromise existing lexical contrasts by creating mergers.

4.2 Hahn (1973)

Table 6 shows Hahn’s proposal (1973; Hahn & Hill, 2008:15–16), which offers a solution to the Excrescence Problem. Hahn only considers zl as part of one morphological alternation in the words log, ldog, and zlog, the analysis for which is shown in Table 7. Hahn claims that zl writes the cluster /zdl/, because /z/ and /d/ are required to express the transitive and verbal meanings of zlog ‘to make turn back’. Thus, he resolves the Excrescence Problem by stating that /d/ was already present in the phonological form underlying zl, but it was not represented in the spelling due to orthographic restrictions on writing consonant clusters.

There are a number of issues with Hahn’s account, most notably that it struggles to account for instances of zl in nouns, since verbal and nominal instances of zl do not have diverging MoT reflexes. Hahn’s analysis of verbal prefixes in Table 7 receive some support in more recent literature on OT morphology: there is a consensus that s-/z- expressed transitivity (Bialek, 2020; Jacques, 2020; Hill, 2023), but I have found no other analysis that treats d- as verbalizing.4 In any case, by tying the underlying phonological content of zl to verbal morphology, Hahn’s account fails to account more generally for the development of zl.

4 Analyses for the function of ot d- include marking the imperfective, passivizing, and transitive, with the latter two analyzed as allomorphs for g- (Bialek, 2020; Zhuang, 2022).
Table 6: Pathway proposed by Hahn (1973; English translation from the original German is Hahn & Hill, 2008:15–16)

OTib /zdʃ/ (written zl) > modern Lhasa /nd/

Table 7: Hahn’s analysis of the l ~ ld ~ zl alternation. He claims that this represents OT /l/ ~ /d-l/ ~ /z-d-l/, where /d-/ is a verbalising prefix, and /z-/ a transitivising one.

<table>
<thead>
<tr>
<th>Orthography</th>
<th>Gloss</th>
<th>Underlying structure</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>log</code></td>
<td>‘backward, reversed’</td>
<td>/log/</td>
</tr>
<tr>
<td><code>ldog</code></td>
<td>‘to turn back’ (intransitive)</td>
<td>/d-log/</td>
</tr>
<tr>
<td><code>zlog</code></td>
<td>‘to make turn back’ (transitive)</td>
<td>/z-d-log/</td>
</tr>
</tbody>
</table>

3.2.1 Orthographic restrictions?

Another issue with Hahn’s account is his argument that zl wrote OT /zdʃ/. Hahn argues that orthographic restrictions forbade dl and zdl, leading scribes to write the initial clusters in /d-log/ and /z-d-log/ as ld and zl. There is indeed a lack of dl or zdl in the OTDO corpus (Hoshi et al., 2006), but could be due to the lack of /dl/ and /zdl/ clusters in OT rather than orthographic restrictions against faithfully representing them in writing. It is unlikely that Tibetan orthography forbade dl and zdl, if /ld/ and /zdl/ were in the spoken language. Tibetan consonant clusters are written as either stacked graphemes (ld མ) or linear strings (bs བས), and it is hard to argue against simply writing dl མ or zdl མ as linear strings. Tibetan spelling also demonstrates considerable flexibility towards introducing novel consonant stacks for writing Sanskrit loans (as in Figure 4), suggesting that in the OT period there were no strict rules against what clusters could or couldn’t be written. The shapes of d མ and l མ are amenable to stacking, since other letters with descenders simply shorten their descenders when written as a superscript letter (e.g. gl ཁ, from g ཁ and l མ, as in gleb ཁ ‘to make flat’). Additionally, while Hahn is happy to assume divergences between OT spelling and phonology, Hahn does not posit systematic divergences between the two (see discussion in Section 2.1), which makes his argument for the phonological content of zl seem post-hoc and haphazard.
The development of zl in Tibetic languages

Tibetan transliteration  བོད་ཀྱི་རྫོང་ཁུང་ལུགས
Latin transliteration  subhāṣitaratnādhiḥmaṇḍastra

FIGURE 4  The title of Sanskrit scholar Sa-skya Paṇḍita’s book of aphorisms (from Bialek, 2022: 34). None of the stacks here are not used for writing native Tibetan words, but are all created to write Sanskrit loans. This includes the combination of consonants being stacked, and the use of ‘<’ as a subscript to represent /ā/. The novel stacks are, from left right, <bhā>, <tna>, <dhi>, <shā>, and <str>.

TABLE 8  Pathway proposed by Beyer (1992:77–79)

<table>
<thead>
<tr>
<th>/zl/</th>
<th>/lz/</th>
<th>/ld/</th>
<th>a) before front vowels: /ldj/</th>
<th>/ldz/</th>
<th>/Ndz/</th>
<th>/dz/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>b) before back vowels: /nd/</td>
<td>/d/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3  Beyer (1992)

Table 8 shows the pathway proposed by Beyer (1992: 77–79). Beyer suggests that all Tibetic languages undergo the metathesis (/zl/ > /lz/) and fortition (/lz/ > /ld/), and the resulting /ld/ undergoes phonological conditioning to produce non-affricated MoT reflexes before back vowels, and affricated MoT reflexes before front vowels due to palatalization. zl presumably merges with ld at the point where zl is realized as /ld/, and /ld/ merges with ñd when it becomes /nd/. Thus, Beyer offers some solution for all three of the Metathesis, Excrescence, and Prenasalization Problems.

3.3.1  The sound changes

The individual steps in Beyer’s pathway are phonetically plausible. The articulatory plausibility of metathesis has already been discussed above. The fortition of /lz/ > /ld/ is articulatorily plausible since [l], [z], and [d] are homorganic, and differ only in the region of palatal contact, as shown in Figure 5. [l] has a central closure between the tongue and the palate, whereas [z] has lateral closures but a central channel. [d] has full oral closure, and is easily produced in the transition between [l] and [z], when there is a period where both the tip and the sides of the tongue are close to the palate. If this near-closure is maintained for long enough that air pressure builds up in the oral cavity, the subsequent lowering of the tongue tip will create plosion as the intraoral pressure is relieved (Reetz and Jongman, 2009: 15). This results in the excrescence of [d], either as [ldz] (assuming [d] is subsequently lost) or as [ld] directly.

/ld/ > /ldj/ > /ldz/ before front vowels is a prototypical case of phonologically conditioned palatalization, which Bateman (2011) defines as a sound change where front or high vocoids cause surrounding consonants’ place of articula-
tion to move towards the palate, especially when the outcome is fricated. Front tongue position in the subsequent vowel is interpreted as part of the onset /ld/, resulting in [ldi] (= Beyer’s /ldji/), altering the tongue position to be higher and further forwards before the plosive’s release into the vowel. The raised tongue body in [di] compared to [d] means that the intraoral cavity has less volume, which means that there is a shorter period of time where there is a pressure differential between subglottal and intraoral pressure. This pressure differential is necessary to maintain voicing during the onset cluster, meaning that the palatalization of /ld/ > /ldj/ compromises the realization of voicing during the onset cluster. The pressure differential (and thereby voicing) can be maintained by partially releasing the oral closure, leading to the production of an affricate /ldʑ/. The need to maintain voicing may be partly due to the presence of /l/ as the first element of the cluster: having a voiced sonorant on either side of [di] means that it is preferable for voicing to continue rather than come to a pause. This may in fact be crucial to the development of affrication during this palatalization process, since Beyer does not posit palatalization and affrication for all instances of /d/ before front vowels, but apparently only for /ld/.

Beyer’s next step, /ld(ʑ)/ > /nd(ʑ)/, is likely driven by a combination of the articulatory and acoustic similarity between [l] and [n], and the pressure to maintain voicing during the cluster.\(^5\) While [l] does have lateral escapes for air-

\(^5\) I am grateful to a reviewer for suggesting the maintenance of voicing as a motivating factor here.
flow, this may not allow enough airflow to maintain an air pressure differential across the glottis for voicing during the cluster's whole duration. Thus, lowering the velum (i.e. nasalization) provides a bigger escape for airflow, allowing voicing to be maintained more easily. The acoustic similarity between [l] and [n] likely also contributed to the reanalysis. Both nasals and laterals have alternating formants and anti-formants, compared to other vocoids which do not have anti-formants, and obstruents which do not have formant structure (Johnson, 2011: 185–205; see Figure 6). [l] and [n] having homorganic articulation adds to their acoustic similarity, since formant structure is heavily determined by tongue position. These similarities may also underlie [l]–[n] mergers that have been observed in other languages, including their recent/ongoing merger in many Sinitic languages, even though the outcome there tends towards [l] (To et al, 2015; Cheng et al, 2022; Huang et al, 2023).
3.3.2 Accounting for MoT reflexes

Despite its success in the above evaluation, Beyer’s proposal appears untenable when one compares its predictions for MoT reflexes to attested ones.

Beyer’s pathway claims that whether zl has an affricate reflex in a MoT language is subject to phonological conditioning. This predicts that

a) affricate reflexes for zl are never found before back vowels, and

b) if a MoT language has /(ⁿ)d/ as a reflex for zl or ld, /(ⁿ)d/ only occurs before back vowels, and /(ⁿ)dʒ/ appears before front vowels.

Neither hypothesis is supported. Contra hypothesis a), Bielmeier et al (2018) attest many Tibetic languages that have affricate reflexes for zl before back vowels, exemplified in Table 9, which show that affricate reflexes for zl could not have arisen from palatalization before front vowels.

Contra hypothesis b), Lhasa Tibetan has /_(ⁿ)d/ as its reflex for OT ld before both front and back vowels,⁶ as shown in Table 10.

---

⁶ The ottcorpus (Hoshi et al, 2006) does not attest instances of zli and zle, suggesting that these clusters were rare or perhaps non-existent in OT. There are also no instances of zli or zle in Bielmeier et al (2018), my main source for reflexes across Tibetic languages. Therefore I use ldi and lde instead, assuming (congruent with Beyer’s pathway) that zl and ld have at this point merged.
The development of zl in Tibetic languages

Table 10 /骊(n)骊/ in Lhasa Tibetan, as the reflex of ld, before all vowels (Gong, 2016). /骊/ is realised as nasalisation in the preceding syllable.

<table>
<thead>
<tr>
<th>Root Word</th>
<th>Lhasa reflex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before front vowels</td>
<td></td>
</tr>
<tr>
<td>ldīng ‘to float’</td>
<td>ʨʰu-ldīng ‘floating (duckweed)’ /骊ʰǔːdiː/</td>
</tr>
<tr>
<td>lde ‘key’</td>
<td>pʰʲag-lde ‘key’ (honorific) /骊ʔẽːdiː/</td>
</tr>
<tr>
<td>Before back vowels</td>
<td></td>
</tr>
<tr>
<td>ldag ‘to lick’</td>
<td>ltɕe-ldag ‘lick’ /骊ʔemtʰaː/</td>
</tr>
<tr>
<td>ldog ‘to turn back’</td>
<td>go-ldog ‘opposite’ /骊ʰõːtʰaː/</td>
</tr>
</tbody>
</table>

Table 11 Pathway proposed by Gong (2016). The pathway splits so that pathway a) produces fricated MoT reflexes for zl, and pathway b) produces non-fricated ones.

zl > a) /骊z/ > /骊dz/  
b) /骊ɮ/ > /骊ɮd/ > /骊ld/ > /骊d/ > /骊(n)d/ > /骊(n)骊/

3.4 Gong (2016)

Table 11 shows Gong’s proposal. There are two aspects to it, which I will interrogate separately: firstly, his decision to derive fricated and non-fricated reflexes through two wholly distinct pathways; secondly, the contents of each pathway.

3.4.1 Split pathway

Gong proposes two distinct pathways because he does not believe zl’s fricated and non-fricated reflexes can be derived from each other. This is contra Beyer (1992) and Bialek (2018), who both propose /骊dz/ > /骊ld/. For Gong, /骊dz/ > /骊ld/ is impossible because /骊/ and /骊/ form the natural class of liquids, and are expected to undergo parallel diachronic developments. Because OT rdz does not have the reflex /rd/ in any MoT language, /骊dz/ > /骊ld/ should not be posited.

While Gong’s observation about rdz is true (cf. Bielmeier et al, 2018: 453–455), there are other issues with this claim. For one, one would expect OT r and l to have parallel outcomes in MoT, but this is not always the case. Table 12 shows some such examples from Lhasa Tibetan, demonstrating that one cannot assume r and l will consistently pattern in diachronic development. Based on Hill (2010)’s analysis of OT phonotactics (see Table 1 in Section 1), it is also possible to claim that r and l did not occur in the same synchronic phonotactic contexts, since /骊/ can appear as a glide consonant while /骊/ cannot.
Wider research into liquids suggest that their status as a natural class should not be taken for granted, even though “liquids” is a commonly used grouping in linguistic research. Existing research identifies a wide range of articulatory behavior corresponding to segments that have been termed ‘liquids’, and their coherence as a natural class falls more on shared phonotactic and diachronic behaviour (Ladefoged and Maddieson, 1996) and perceptual similarities (Ladefoged and Maddieson, 1996; Howson and Madathodyil, 2023). The lack of consistent diachronic patterning in other clusters suggests that we should not expect rdz and /ldz/ to develop in parallel either, so the lack of observed rdz > /rd/ does not necessarily prohibit /ldz/ > /ld/.

Other evidence for or against the possibility of /ldz/ > /ld/ is conflicting. On the one hand, Bialek (2018) observes recent or ongoing deaffrication changes for reflexes of zd in Amdo languages, such as /rdz/ > /rd/ and /hdz/ > /hd/ (see Section 3.5). She does not observe /ldz/ and /ld/ in alternation, but the attestation of these similar changes suggests that /ldz/ > /ld/ is possible. On the other hand, evidence for Bialek (2018) does not directly observe synchronic alternations for /ldz/ and /ld/ as she does for /rdz ~ rd/ and /hdz ~ hd/, leaving room for deaffrication to be limited to recent developments in Amdo, and for /ld/ to derive from processes other than the deaffrication than /ldz/. Thus, data from Bialek (2018) does not necessarily elucidate the relationship between fricated and non-fricated reflexes of zd outside of Amdo. Other evidence suggests that deaffrication is typologically uncommon, including apparently in the history of Tibetic. Crosslinguistically, deaffrication (i.e. affricate > plosive) are uncommon, as shown in Kümmel (2007: 96–98)’s typology of sound changes, and Bybee and Easterday (2019)’s study of synchronic consonant alternations as precursors of phonological change. In Bybee and Easterday (2019), there are zero cases of synchronic affricate > plosive fortition in their sample of 81 languages. This is especially striking when considering that fricative > plo-
sive changes are much more common in both these sources, including being the second most common fortition phenomenon found in Bybee and Easterday (2019), occurring in five languages, compared to approximant > fricative in twelve. Deaffrication has also been posited in the development of Sino-Vietnamese, but not without contention. Hashimoto (1978: 7) claims that Middle Chinese fricatives and affricates became plosives when words with these onsets were loaned into Vietnamese (*s, *ts > s¹, *d, *dz > s², with s¹ and s² displaying a tonal contrast). However, Phan (2013) working on the same topic seems to prefer the hypothesis that MC fricatives became Sino-Vietnamese plosives, while MC affricates became fricatives, so the Sino-Vietnamese case cannot function as proof of affricate > plosive changes in East Asia.

Within Tibetic, there appears to be a conspicuous lack of deaffricated MoT reflexes for orthographic clusters containing an affricate. This is shown in Table 13, using data drawn from Bielmeier et al (2018). The blank cells in Table 13 may simply be gaps in attestation, especially considering that Bielmeier et al (2018) only contains verbs, and that its status as a reference work means it cannot contain all details about each language. But even so, it is striking that none of these clusters have plosive reflexes, suggesting that there seems to be a general dispreference against deaffrication in the history of Tibetic. The pattern of reflexes in Table 13 may be explained through perceptual factors. All OT clusters with fricative reflexes in can be seen as homorganic, and transitions between consonants with a central oral closure to fricatives often produce an excrescent fricative (cf. English *princess, sometimes pronounced [ˈpʰɹɪn.tsɛs]). Thus, sequences like [rdz] and [ldz] in Tibetic may be interpreted as /rz/ and /lz/ by listeners, leading to the loss of the plosive because it is not perceived as part of the underlying representation by speakers (cf. Ohala, 1981).

Bialek (2018)’s evidence from Amdo does not argue decisively for or against Gong’s split pathway for zl. It can be interpreted as a widespread tendency for the deaffrication of /ldz/ across Tibetic, thereby arguing that non-affricated reflexes of zl in western and central Tibetic-speaking regions all derived from deaffrication. Bialek herself (2018) takes this approach, although she only explicitly discusses Western Archaic and Amdo Tibetic languages. Alternatively, Bialek’s evidence for deaffrication can be interpreted as a localized change in modern Amdo, so reflexes of zl from other regions can still follow Gong’s pathway, as in Table 14:

---

7 mdz is not homorganic per se, but if one takes the view that mdz merges with ndz after or, resulting in /nďz/, then mdz can also be regarded as homorganic.
### TABLE 13 Affricate syllable onsets and their MoT reflexes across all Tibetic languages that I could find data for (ignoring tonogenesis)

<table>
<thead>
<tr>
<th>Cluster</th>
<th>MoT reflexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affricate reflexes</td>
<td>Fricative reflexes</td>
</tr>
<tr>
<td>dz</td>
<td>/dz/, /ts/</td>
</tr>
<tr>
<td>mdz</td>
<td>/ndz/, /dz/</td>
</tr>
<tr>
<td>ndz</td>
<td>/ndz/, /νdz/, /dz/, /ts/</td>
</tr>
<tr>
<td>rdz</td>
<td>/rdz/, /ndz/, /nts/, /dz/, /ts/</td>
</tr>
<tr>
<td>mdz</td>
<td>/ndz/, /νdz/, /mdz/, /dz/, /dʒ/, /tc/</td>
</tr>
<tr>
<td>ndz</td>
<td>/ndz/, /dz/, /dʒ/, /tc/, /tʃ/</td>
</tr>
<tr>
<td>/tʃʰ/, /xtʃ/, /tʃ/</td>
<td>/γj/, /pʰγj/, /φαc/, /wc/</td>
</tr>
<tr>
<td>rdz</td>
<td>/rdʒ/, /dz/, /tc/</td>
</tr>
<tr>
<td>ldz</td>
<td>/rdʒ/</td>
</tr>
<tr>
<td>gts</td>
<td>/γts/, /γtʃs/, /xts/, /ts/</td>
</tr>
<tr>
<td>bts</td>
<td>/tsʰ/</td>
</tr>
<tr>
<td>rts</td>
<td>/γts/, /γrts/, /xts/, /hts/, /ts/</td>
</tr>
<tr>
<td>sts</td>
<td>/ts/</td>
</tr>
<tr>
<td>brts</td>
<td>/γts/, /φγts/, /ts/</td>
</tr>
<tr>
<td>tsʰ</td>
<td>/tsʰ/</td>
</tr>
<tr>
<td>mtsʰ</td>
<td>/tsʰ/</td>
</tr>
<tr>
<td>ntsʰ</td>
<td>/nʰtsʰ/, /ntsʰ/, /mtsʰ/, /tsʰ/, /ts/</td>
</tr>
<tr>
<td>gtc</td>
<td>/tc/, /tʃ/, /xtʃ/, /tʃtʃ/, /γtʃc/, /γtc/, /γtcʰ/</td>
</tr>
<tr>
<td>btc</td>
<td>/tʃ/, /tc/</td>
</tr>
<tr>
<td>ltc</td>
<td>/tc/, /tʃ/, /xtʃ/, /stʃc/</td>
</tr>
<tr>
<td>tcʰ</td>
<td>/tcʰ/, /tʃʰ/</td>
</tr>
<tr>
<td>mtcʰ</td>
<td>/tcʰ/, /tʃʰ/, /mʰtcʰ/</td>
</tr>
<tr>
<td>ntcʰ</td>
<td>/tc/, /tʃ/, /tcʰ/, /ntcʰ/, /tʃʰ/, /pʰtcʰ/, /γc/, /cʰ/, /xʰ/, /ʃ/, /x/</td>
</tr>
<tr>
<td>/γtc/, /ptc/</td>
<td>—</td>
</tr>
</tbody>
</table>

Bielmeier et al., 2018: 232–264 for TC, TCʰ, DZ, 427–455 for TS, TSʰ, DZ

### TABLE 14 Gong (2016)’s pathway, with extra changes in pathway a) to accommodate Bialek (2018)’s evidence for deaffrication in Amdo languages

<table>
<thead>
<tr>
<th>zl</th>
<th>&gt;</th>
<th>a) Western Archaic &amp; Amdo: /lz/ &gt; /ldz/ &gt; Amdo only: /rdz/, /hdz/ etc &gt; /rd/, /hd/ etc</th>
</tr>
</thead>
<tbody>
<tr>
<td>b) other Tibetic languages: /zlʒ/ &gt; /zlʒd/ &gt; /zl̥d/ &gt; /ld/ &gt; /n̥d/ &gt; /(n̥d)/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overall, then, there is conflicting evidence surrounding Gong’s proposal for a split pathway. I now discuss the contents of Gong’s pathways, focusing on the plausibility of the changes in pathway b), since the plausibility of changes in pathway a) have already been demonstrated in my discussion so far.

3.4.2 “Verschärfung”
In pathway b), for deriving non-fricated reflexes, Gong posits a process that he calls ‘Verschärfung’, parallel to a similar process in Rgyalrongic languages, involving the “gradual fortition of sonorants to stops or clusters containing stops” (2016: 143). He proposes \( \text{zl} \rightarrow \text{zl} \text{ʃ} > \text{zl} \text{ʃ} \text{d} > \text{zl} \text{d} > \text{ld} \). /zd/, /zɬd/, and /zd/ are not attested as reflexes in MoT (as Gong himself acknowledges), but this cannot by itself falsify Gong’s proposal: if we interpret these intermediate steps as brief and overlapping transitional stages rather than discrete phonologized production targets, then it is not surprising that they are not preserved as reflexes.

A phonetic interpretation of Gong’s Verschärfung is amply possible. For example, /zl/ > /zɬ/ may occur because [zɬ] may be articulatorily easier when considering the transition into the following vowel. [z] requires the sides of the tongue to form a closure with the alveolar ridge and palate, and the centre of the tongue to be grooved to produce friction. For [l], the tip of the tongue forms a central closure with the alveolar ridge, and one or both sides of the tongue are lowered for airflow. For [ʃ], tongue position is like for [l], except that the lateral channel(s) are narrow enough that airflow causes friction. In a sequence like [zɬa], the sides of the tongue need to lower first, and then the tip of the tongue lowers to produce the vowel. In [zɬa], because the sides of the tongue are closer to the palate, there is minimal tongue movement in the [zɬ] transition, leaving the whole tongue to move in one go in the [la] transition. This means that [zɬa] is easier to produce during fast speech, compared to [zla] which requires the movements of the sides and tip of the tongue to be coordinated sequentially. The excrescence in /zɬ/ > /zɬd/ is more easily attributed to perception than articulation, since the transition between [ɬ] and the following vowel only involves the tongue moving away from the passive articulators so [d], which requires full momentary closure, cannot ‘fall out’ in any way.

However, this does not guarantee the pathway’s ultimate plausibility. Gong posits Verschärfung “inspired by the Rgyalrongic evidence” (2016: 144), but it is worth considering the applicability of Rgyalrongic evidence here, because there may be insufficient parallelism between Rgyalrongic and Tibetic for Gong’s claims. A change like /l/ > /ɬ/ > [ɬd] appears to have taken place in some Rgyalrongic languages, including the Khang-gsar variety of Stau, which has [ɬd] as an allophone of /ɬ/, and the alternation /la/ ‘boil’ ~ */s-la/ > /zɬda/
Gong then draws a parallel between this Rgyalrongic fortition and the historical Tibetic fortitions like Li Fang-kuei’s law \( \text{\*rj} \to \text{ot r}\acute{g}\) (Hill, 2013), and claims that the Rgyalrongic Verschärfung process is applicable to Tibetic \( zl \) because the same process is already observed in the history of Tibetan. However, as pointed out by one of the reviewers for this article, Verschärfung observed in Khang-gsar Stau inserts a plosive before the vowel, but pre-\( \text{OT} \) fortitions like Li Fang-kuei’s law always inserts the plosive between two consonants (cf. Hill, 2019b: 17), so the Rgyalrongic and Tibetic processes are not parallel because they ultimately have different phonological environments.

This makes the Verschärfung of Tibetic \( zl \) seem like an over-elaborate proposal. By identifying \( /l/ > /z\breve{l}/ > [\breve{z}d] \) a mechanism of fortition in Rgyalrongic and applying it wholesale to Tibetic, Gong seems to neglect that “Verschärfung” isn’t a unitary change with internal steps, but a post-hoc, epiphenomenal label that one can use to describe the successive but separate changes \( /l/ > /z\breve{l}/ > [\breve{z}d] \). Verschärfung is thus a mechanism of fortition in the sense of “mechanism” advocated by Walkden (2021), a descriptive shorthand which does not have explanatory value and whose subparts have no inherent cohesion. As such, it cannot be posited to explain the fortition of \( /zl/ > /ld/ \) in Tibetic, least of all when the non-attestation of \( /z\breve{z}\breve{l}/, /z\breve{z}d\breve{l}/ \), or \( /zd\breve{l}/ \) in MoT means that Gong’s Verschärfung posits three phonetic/phonological stages that left no trace. It is possible to argue in support of Gong’s Verschärfung, and claim that \( /z\breve{z}\breve{l}/, /z\breve{z}d\breve{l}/, \) and \( /zd\breve{l}/ \) existed only briefly as overlapping phonetic realisations, rather than distinct diachronic stages, but there appears to be no inherent benefit to Gong’s Tibetic Verschärfung over simply positing a \( /zl/ > /zd\breve{l}/ \) fortition, other than to preserve the unity of Verschärfung as a mechanism, when it has no inherent unity.

### 3.4.3 Prenasalization

After Verschärfung, Gong claims that \( /ld/ \) first becomes a unitary segment \( /d/ \), and then merges with \( /n\breve{d}/ \) because both have a very negative voice-onset time. Gong considers this necessary, because \( /ld/ \) is a consonant cluster and \( /n\breve{d}/ \) a single segment, so the two cannot merge directly. Gong supports this by citing the possible existence of a unitary prelateralized phoneme \( /\breve{d}/ \) in Zenskar Kenhat, a Tibetan language spoken in Ladakh.

Nikolaev (2020) criticizes Gong’s analysis for being needlessly complex, especially when a merger between \( /ld/ \) and \( /n\breve{d}/ \) can be explained by acoustic and perceptual factors, such as those I have been discussing. The complexity of Gong’s proposal does not make it impossible, but phonetic factors may well have also been active alongside Gong’s structural ones. Nikolaev (2020) also suggests one may investigate the role of functional load in affecting the merger.
between /ld/ and /ⁿd/, which I address in Section 4.2. Here, it suffices to conclude that Gong’s analysis for /ld/ > /ⁿd/ is possible, but not wholly convincing.

### 3.5 Bialek (2018)

Table 15 shows Bialek’s (2018) pathway, focusing on the development of clusters in the phonologically archaic wat and Amdo branches. Since prenasalization does not occur in those branches, she does not address it, but her treatment of the Metathesis and Excrescence problems apply to all Tibetic languages. As referenced in my discussion of Gong (2016), Bialek presents alternations between fricated and non-fricated reflexes for zl in Amdo (Figure 7). This leads her to propose deaffrication in Amdo reflexes of zl, and to project deaffrication onto the reflexes of zl in all other subgroups within Tibetic. As established in my discussion on Gong (2016), despite its presence in modern Amdo, deaffrication is typologically uncommon, and thus one should be wary of positing it occurring multiple independent times as Bialek has done. The pervasiveness of ongoing deaffrication in Amdo Tibetic may be attributed to areal contact, but this explanation localizes deaffrication to present-day Amdo, without particularly justifying it in other Tibetic languages. Positing the independent occurrence of deaffrication for reflexes of zl in all other Tibetic subgroups is parsimonious, but in the apparent absence of deaffrication in other clusters (cf. Table 13), Bialek’s claim is dubious.

### 4 Discussion

In Section 4.1, I summarize my evaluations of the five pathways. The most plausible account of the development of zl is either that of Gong (2016) or Bialek (2018), depending on further research on the classification of Tibetic languages.
I conclude that, until there is such evidence reflecting the pattern of phylogenetic branching implicated in Gong’s pathway, we should provisionally accept Bialek’s pathway, which is overall well-supported and parsimonious. In Section 4.2, I outline a possible study on the role of functional load in investigating the mergers that occurred in the development of zl.

### 4.1 Accounts for the development of zl

Table 16 summarizes how the five pathways discussed in Section 3 accounts for reflexes of zl, framed in relation to the three Problems of its development.

Based on discussions in Section 3, the proposals by Sprigg (1972), Hahn (1973), and Beyer (1992) should all be rejected. Sprigg posits historical forms that are elaborate and difficult to interpret (zd/zl etc), and does not substantively address the Metathesis Problem or Prenasalization Problem, even though it would have been useful and articulatorily justifiable to posit metathesis in the development of zl. He does address the Excrescence Problem by proposing that consonant insertion was rife during the development of zl, but he does not discuss this in detail and simply assumes it. Hahn’s pathway is uninformative regarding phonological developments because it uses an analysis of verbal morphology to account for changes affecting zl, ignoring the fact that those changes also affected verbs. By postulating /zd/ as the phonological form underlying zl, it also fails to provide a systematic and persuasive account of orthography-phonology mappings. Beyer proposes phonetically plausible solu-

---

**Figure 7** Amdo languages with their reflexes for sl (middle column) and zl (right column), taken from Bialek (2018: 7), showing apparent ongoing deaffrication for reflexes of zl. Language name abbreviations are following Bielmeier et al. (2018).
### Table 16

Summary of how the five different pathways discussed in this section address the three Problems for the development of \textit{zl}. “—” indicates when a pathway does not address this Problem.

<table>
<thead>
<tr>
<th>Phonological content underlying \textit{zl}</th>
<th>The Metathesis Problem</th>
<th>The Excrescence Problem</th>
<th>The Prenasalisation Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sprigg (1972) \textit{zd}/—not justified</td>
<td>Metathesis does not occur, but its effects are generated through consonant insertion and deletion</td>
<td>/d/-excrescence is freely permitted between other segments in a cluster</td>
<td>Pathway includes /ld/ &gt; /nd/, but gives no other detail</td>
</tr>
<tr>
<td>Hahn (1973) \textit{zd}/—not justified</td>
<td>Not addressed</td>
<td>N/A—/d/ is already present in the OT form</td>
<td>Not addressed</td>
</tr>
<tr>
<td>Beyer (1992) \textit{zl}/</td>
<td>/zl/ &gt; /lz/ occurs in all Tibetic languages</td>
<td>Fortition from /lz/ to /ld/</td>
<td>Pathway includes /ld/ &gt; /nd/, but gives no other detail</td>
</tr>
<tr>
<td>Gong (2016) \textit{zl}/</td>
<td>/zl/ &gt; /lz/ for languages with a fricated MoT reflex</td>
<td>Fortition of /zl/ &gt; /zld/ &gt; /ld/, involving Gong’s “Verschärfung” process</td>
<td>N/A—no fricated MoT reflex also has prenasalisation /ld/ &gt; /nd/, then merges with /ⁿd/</td>
</tr>
<tr>
<td>Bialek (2018) \textit{zl}/</td>
<td>/zl/ &gt; /lz/ occurs for all Tibetic languages</td>
<td>/lz/ &gt; /ldz/ for all Tibetic languages</td>
<td>Not addressed</td>
</tr>
</tbody>
</table>

Sprigg (1972) nowhere addresses the Metathesis Problem in his proposal. He does not justify /zd/ not occurring, but its effects are generated through consonant insertion and deletion. Hahn (1973) also does not address the Metathesis Problem. Beyer (1992) presents a pathway that includes /ld/ > /nd/, but gives no other detail. Gong (2016) and Bialek (2018) pathways cannot be unequivocally rejected or accepted. Gong’s pathway for fricated reflexes, which involves metathesis and excrescence, is plausible. Bialek’s pathway for non-fricated reflexes, which involves the fortition of /zl/ > /zld/ and prenasalization, risks over-complication and does not give a satisfying explanation for the changes.

The pathways from Gong (2016) and Bialek (2018) cannot be unequivocally rejected or accepted. Gong’s pathway for fricated reflexes, which involves metathesis and excrescence, is plausible. His pathway for non-fricated reflexes, which involves the fortition of /zl/ > /zld/ and prenasalization, risks overcomplication and does not give a satisfying explanation for the changes.
proposal for a split pathway, which hinges on the impossibility of /ldz/ > /ld/,
receives conflicting evidence, since deaffrication processes are observed in
modern Amdo languages, but otherwise appears rare crosslinguistically and
in the rest of Tibetic. So, at best (for Gong’s pathway), the development of zl
followed a pathway like Gong’s, but the fricated reflexes achieve /zl/ > /zld/
and /ld/ > /ⁿd/ through different processes from Gong’s proposed Verschär-
fung and prelateralization. At worst, changes to zl did not follow a bifurcated
pathway, but something like /zl/ > /lz/ > /ldz/ > /ld/ > /ⁿd/. The latter option
is essentially what Bialek proposes, although she does not address the Pre-
usalization problem because she focuses on wat and Amdo languages. Recent
or ongoing deaffrication of reflexes of zl in Amdo is used to justify /ldz/ >
/ld/ in other Tibetic subgroups. The other steps in Bialek’s pathway are all
well-supported, and Bialek proposes the more parsimonious account. Overall,
choosing between Gong and Bialek’s proposals rests on whether a split path-
way is justified. Gong’s pathway has distinctive implications for classification
within Tibetic, and further research on resemblances between MoT languages
is required to bear upon Gong’s pathway.

Gong’s split pathway suggests that MoT languages with fricated reflexes for
zl are placed in a different phylogenetic branch to those with non-fricated
reflexes, since they have followed different sound change pathways for zl from
the start. This means that languages that are currently classified under the same
branch of Tibetic may be phylogenetically distant from each other (e.g. wat,
where some languages have a fricated reflex for zl and some have a non-fricated
one), and that the languages with fricated reflexes for zl cannot be said to rep-
resent a “prior” phonological state of the languages with non-fricated reflexes.
This does not necessarily contradict current classification practices for Tibetic,
since these groupings are based on both geographical proximity and shared
linguistic features (Bielmeier et al, 2018: 38). Tournadre (2014) further argues
that the effects of borrowing and migration makes a cladistic classification of
Tibetic languages impossible, so classifications should aim to represent geolin-
guistic continuua. Even so, it should be possible to evaluate Gong’s proposed
isogloss surrounding the fricated/non-fricated reflexes for zl in relation to MoT
languages’ similarities and differences in other respects. For example, some
wat languages have a fricated reflex for zl and some a non-fricated one. If Gong
is correct, one should expect other linguistic similarities within each of these
two groups, and differences between them.

Since Gong’s proposal for a split pathway is the more complicated, and
requires specific evidence to support it, it is better to provisionally accept
Bialek’s proposal until classification evidence for or against Gong’s proposal
can be produced. Intense and prolonged contact between different Tibetic lan-
guages means that this will require further research on language contact, areal effects, and word histories. If there is evidence to support Gong’s split pathway, it will still need different mechanisms to account for /d/-excrescence and prenasalization in the non-affricated pathway, and to include a deaffrication change in the affricated pathway to account for /rdz ~ rd/-type alternations Bialek (2018) observes in Amdo.

4.2 Mergers
For the four-way merger between zl, ld, "d, and md in Lhasa Tibetan and other innovative Tibetic languages, both Gong and Bialek’s pathways suggest that zl merges with ld first, and then their shared reflex /ld/ merges with md and "d to produce the shared reflex /nd/.

Nikolaev (2020) suggests that it may be interesting to investigate this merger from the perspective of functional load, following past work on functional load and phonological mergers like Wedel et al (2013a; 2013b) and Babinski and Bowern (2018). Considering functional load as a motivator of phonological change focuses on the role of phonological units as units of contrast within the language’s phonological system. Phonological units with lower functional load are more likely to be lost because they “do less work” in maintaining lexical contrasts in the language. This was first hypothesized by Martinet (1952), who was one of the early researchers who worked on functional load. In recent work, functional load has been quantified as the rate of occurrence of a phonological unit within the language (Wedel et al, 2013a; 2013b; Babinski and Bowern, 2018), or as part of an information-theoretic system where the loss of a phonological contrast results in an increase in system entropy (Surendran and Niyogi, 2006).

Wedel et al (2013b)’s system is easier to implement, because it allows one to calculate a phonological unit’s functional load from its number of occurrences in a language corpus. Wedel et al (2013b) and its follow-up studies (Wedel et al, 2013a; Babinski and Bowern, 2018) focus on segmental phoneme contrasts. They quantify functional load using these two measures: phoneme frequency, the frequency with which a phoneme occurs in a language, which may be presented as a percentage of the total frequency of all phonemes in the language; and minimal pair count for a given pair of phonemes, which is the number of minimal pairs in the language that are distinguished by this pair of phonemes (Wedel et al, 2013b). A higher phoneme frequency and minimal pair count translate to a phoneme having higher functional load in the language.

They find two phonemes are more likely to merge if they have a lower minimal pair count. In Babinski and Bowern’s (2018) study on Bardi, phoneme pairs with a minimal pair count of ≤ 4 always merge, while pairs with > 4 never
merge. They also find that, if there is a big frequency difference between the two phonemes in question, the less frequent phoneme is more likely to merge into the more frequent one. Wedel et al (2013a) find that these effects are stronger for minimal pairs where both lemmas are in the same syntactic category, since there are fewer morphosyntactic cues to differentiate the lemmas.

Wedel et al (2013a, 2013b) and Babinski and Bowern (2018) work on phonemes, but their methodology (based on its description in Wedel et al, 2013b) should also be applicable to phonological units of other sizes. Therefore, it is possible to apply their methodology to two-consonant onset clusters as units of contrast, as in zl is contrastive with ld or kʰr, like how the phonemes /z/ may be contrastive with /d/. Thus, a given cluster’s functional load can be quantified in terms of its cluster frequency, the frequency with which the cluster occurs in the language; and cluster contrast count, the number of lexical items that are distinguished by a given pair of contrasting clusters. These numbers can be obtained from corpora of Tibetan texts, such as the OTDO corpus of Old Tibetan (Hoshi et al, 2006).

The four-way merger of zl, ld, ᶷd, and md can be conceptualized as three mergers between two clusters each: zl + ld > /ld/, md + ᶷd > /ⁿd/, and the merger of their outcomes, /ld/ + /ⁿd/ > /ⁿd/. For the zl + ld > /ld/ merger, one would be investigating why zl and ld merge with each other, rather than zl merging with other Cl clusters and ld with other Cd clusters. Thus, one would obtain the frequency of zl, ld, and all Cl and Cd clusters from the corpus, and the pairwise cluster contrast counts between zl and ld, zl and Cl clusters, and between ld and Cd clusters. Based on the findings of Wedel et al (2013a) and Babinski and Bowern (2018), one should expect the cluster contrast count between zl and ld to be lower than those between zl and Cl or ld and Cd (hence zl and ld merging with each other rather than with other Cl and Cd clusters), and for zl to be less frequent than ld (hence /ld/ is the outcome of the merger).

This process can then be repeated for the md + ᶷd > /ⁿd/ and /ld/ + /ⁿd/ > /ⁿd/ mergers. md + ᶷd > /ⁿd/ involves looking at the frequencies and cluster contrast counts of md, ᶷd, and other Cd clusters. One should expect the cluster contrast count between md and ᶷd to be lower than those between each of them and other Cd clusters (hence md and ᶷd merging with each other rather than with other Cd clusters), and for md to be less frequent than ᶷd (hence /ⁿd/ is the outcome of the merger). For /ld/ + /ⁿd/ > /ⁿd/, “/ld/” and “/ⁿd/” which contribute to the merger cannot be counted directly in the corpus, but “/ld/” is counted as zl and ld, and “/ⁿd/” is counted as md and ᶷd. Again, one looks at “/ld/” and “/ⁿd/” versus other Cd clusters. One should expect the cluster contrast count between “/ld/” and “/ⁿd/” to be lower than those between each of them and other Cd clusters (hence “/ld/” and “/ⁿd/” merging with each other rather than
with other \(Cd\) clusters), and for “/ld/” to be less frequent than “/nd/” (hence /nd/ is the outcome of the merger).

There is one difficulty for this potential study, arising from Hill (2010)’s observation that orthographic \(zl\) and \(sl\) may both represent /zl/ in \(OT\), and hence orthographic \(zl\) and \(sl\) are not always contrastive. This means that one cannot take the cluster frequencies and cluster contrast counts involving \(zl\) and \(sl\) (the latter counted among the other \(Cl\) clusters that \(zl\) does not merge with) at face value. This can be corrected if one knows the respective rates of orthographic \(sl\) representing \(OT\) /zl/ and /sl/, but I have so far not found a way to do so. As it stands, orthographic \(zl\) is a lot less common than \(sl\) in the \(OT\) \(DO\) corpus, \(sl\) appearing five times as often as \(zl\) (Hoshi et al, 2006), but this trend may change when one accounts for the fact that some of these instances of \(sl\) represent /zl/.

However, the functional load methodology as outlined here may prove amply useful for other mergers in the history of Tibetic languages, since mergers are very prevalent and the relationship between orthography and \(OT\) phonology is largely transparent (Hill, 2010; Bialek, 2018). I hope to pursue such a project in the near future.

Acknowledgments

I would like to thank Dr Marieke Meelen for her help and support in the completion of the dissertation that this manuscript derives from. Many thanks also to the reviewers to this article, for their suggestions and advice.

References


Hill, N.W. (2019a). The derivation of the Tibetan present prefix g- from ḥ-. *Center for International and Regional Studies (Georgetown University), 72*(3), 325–332. https://doi.org/10.1556/062.2019.72.3.4


Ohala, J.J. (2012). The listener as a source of sound change. In M.-J. Solé & D. Recasens (Eds.), The Initiation of Sound Change (pp. 21–36). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.323.05oha


Wang, Y. (1996). 藏语 “zla-ba” 一词音义考 [Investigation into the pronuncia-
tion and meaning of the word “zla-ba”]. 民族语文 [Minority Languages of China], 1996(5).


