



# The Indo-Iranian labial-extended causative suffix

*Indic -(ā)páya-, Eastern Iranian \*-(ā)uaja-, and Proto-Caspian \*-āwēn-*

Zia Khoshirat | ORCID: 0000-0002-2291-0414

Department of Anthropology, University of California, Los Angeles, CA, USA  
ziaoddin@g.ucla.edu

Andrew Miles Byrd | ORCID: 0000-0002-1667-1744

Department of Linguistics, University of Kentucky, Lexington, KY, USA  
andrewbyrd@uky.edu

Received 25 February 2022 | Accepted 13 April 2023 |

Published online 9 August 2023

## Abstract

Alongside the expected reflexes of the Proto-Indo-European causative suffix *\*-éje/o-*, there appears in Indo-Iranian a second, expanded version that contains a labial consonant: Indic -(ā)páya- and Eastern Iranian (EIr.) \*-(ā)uaja-, the latter continued in Khotanese -ev-, Khwarazmian -(‘)wy-, and other modern EIr. languages. In this paper, we will argue that \*-(ā)uaja- is also the source of a causative marker in two closely related Caspian (Western Iranian) languages, Gilaki and Tati-Talyshi, through a reconstructable Proto-Caspian form \*-āwēn-. We propose that these three suffixes, -(ā)páya-, \*-(ā)uaja-, and \*-āwēn-, originated in Proto-Indo-Iranian, through the rounding of a root-final laryngeal to a labial sound in causative formations.

## Keywords

Indo-Iranian – Indic – Iranian – Caspian – causative – laryngeal – phonologization

## 1 Introduction<sup>1</sup>

As in many Indo-European languages, the Proto-Indo-European (PIE) causative suffix *\*-éje/o-*<sup>2</sup> is continued as a productive causative marker in Indo-Iranian (IIr.), developing as expected to *-áya/-aiia-* in the oldest IIr. languages. However, in Indo-Iranian there appears a second, seemingly expanded version of *-áya/-aiia-*, which we will refer to in this paper as the Labial-Extended Causative Suffix (LECS). The Indic LECS *-(ā)páya-* is found already within the Rigveda (Jamison 1983: 80), becoming the predominant causative marker in the Middle and later stages of Indic (Masica 1991). In Middle Eastern Iranian the suffix *-ēv-/(')wy-* is attested, which Cathcart (2015: 42, cf. Emmerick 1968: 187, Schwartz 1969: 447) traces back to *\*-(ā)uaja-*, though its lack of counterparts in Western Iranian, and in Avestan in particular, has raised questions about its reconstruction for Proto-Iranian.

In this paper, we will begin by showing that there is currently no compelling explanation for the origin of the Indic LECS *-(ā)páya-*, and that, following Cathcart (2015), at least some instances of the LECS found in Eastern Iranian must be derived from a reconstructed Proto-Iranian LECS *\*-(ā)uaja-*. We will also demonstrate that the Iranian suffix *\*-(ā)uaja-* is continued by two closely related Caspian (Western Iranian) languages, Gilaki and Tati-Talyshi. We will then argue that the Iranian LECS *\*-(ā)uaja-* and Indic LECS *-(ā)páya-* are most likely directly connected, both coming from similar sources in Proto-Indo-Iranian (PIIrr.) through the metanalysis of a root-final laryngeal as being part of the causative suffix. Our proposal, while speculative, provides a single origin point for the LECS, a suffix which is not found in any other Indo-European branch. Our proposal also accounts for the curious restriction of the LECS to original laryngeal-final roots in Sanskrit.

## 2 The labial-extended causative suffix in Indic

The enigmatic causative suffix *-(ā)páya-*, found in the earliest Sanskrit (Jamison 1983), is primarily used after long-vowel roots, such as *dhā-páya-* ‘make place’ (< *\*d<sup>h</sup>eh₂-* [LIV 136]), *sthā-páya-* ‘install’ (< *\*steh₂-* [LIV 590]), and *snā-*

<sup>1</sup> We would like to thank Allison Burkette, Jessica DeLisi, Dieter Gunkel, Farhang Karami, Ron Kim, Agnes Korn, Alexander Lubotsky, Martin Schwartz, Don Stilo, and two anonymous reviewers for their helpful feedback and discussions. All errors are our own.

<sup>2</sup> In this paper, we follow Bozzone (2020) in the reconstruction of PIE *\*-éje/o-* as a causative suffix formed to unaccusative base verbs.

*páya-* ‘wash (up)’ (< \**sneh*<sub>2</sub>- [LIV 572]), and becomes increasingly more productive through its spread to other root shapes such as *arpáya-* ‘make fit together’ (< \**h<sub>2</sub>er-* ‘id.’ [LIV 269]), *dípáya-* ‘make shine’ (\**deih*<sub>2</sub>- [LIV 108]), and *smāpáya-* ‘make smile’ (< \**smei-* [LIV 568]).<sup>3</sup> There is currently no clear explanation for the origins of -(*ā*)*páya-*, though certainly not due to a lack of trying. Over the last two centuries, scholars have put forth a myriad of possible PIE sources to account for this unexpected *-p-*, but as Ghosh (1982: 115) writes, none of these explanations has been satisfactory: “the existence of a suffix *-p-* in Indo-European remains uncertain, which renders futile speculation regarding its origin.”

Early scholarship attempted to explain the *-p-* through the composition of a verbal root with a following root meaning ‘make’ or ‘do’, which are frequently used in causative and factitive formations, both analytically (cf. English *make strong*) and synthetically (cf. late Lat. *forti-fic-āre* ‘id.’ and Lat. *faciō* ‘do, make’). In this way, Schleicher & Bendall (1847: 176) derive the *-p-* of *-paya-* from a root *ap-* ‘do, make’ (< \**h<sub>3</sub>ep-* ‘produce’ [LIV 298–299]), seen most prominently in forms such as Skt. *ápas-* ‘work, action’ and Lat. *opus* ‘work’, used together with the inherited *-aya-* causative.<sup>4</sup> Thus, *sthāpaya-* ‘install’ < \**sthā-* + *a(‘)p-* + *-aya-*. A similar proposal was made by Süitterlin (1906: 536–538) for *āp-* ‘obtain’ (< \**h<sub>1</sub>ep-* ‘seize, catch’ [LIV 237]). This sort of explanation, however, has understandably been viewed by many as unlikely (e.g., Brugmann 1895: 256), since there is no evidence that these roots could be compounded in such a way following the rules of early Indic morphology. Moreover, it is unclear why such a new compositional causative would have been created for long-vowel roots (cf. Ghosh 1984: 68), given the existence and productivity of the inherited *-aya-* suffix.

Other explanations lean into the idea that the source of this *-p-* comes from an “enlarged” version of a root or subset of roots. For Vendryès (1931: xvi), this *-p-* is found in verbs of motion. For Brugmann (1895: 256; cf. Persson 1912: 311) the *-p-* of the *-paya-* causative is a murky “altüberkommenes Determinativ”, with the *-p-* of the causative stem *dāpaya-* ‘shares’ (to *dā-* ‘share’) found elsewhere in purportedly related forms, including Gk. δάπτω ‘devour’, δαπάνη ‘cost’, Arm. *tawn* ‘celebration’, Lat. *damnum* ‘cost’, *daps* ‘sacrificial meal’, and OIce. *tafn* ‘sacrificial animal/meal’, all from \**dah*<sub>2</sub>*p-* ‘divide’ (LIV 104), with a possible *-p-* extension to \**dah*<sub>2</sub>(*i*)- ‘share’ (LIV 103). However, there is no evidence of a *-p-*

3 See Tedesco (1923: 389), Leumann (1940), Edgerton (1946), Insler (1987) for further discussion of the spread of the LECS outside of its original domain.

4 This root is not connected, as suggested by Schleicher, with Gk. ποιέω ‘make, do’, which is instead more likely to have been an earlier causative to \**kʷeij-* ‘pile up’ (LIV 378–379).

in this root outside of the Indic causative in question, and a *p*-extended form *dāpayīta* is attested rather late, in the Sūtras (Whitney 1885: 72).

Other examples of *-p*- extensions suggested as the source of the *-paya*-causative are even less compelling, such as *drā-* ‘run’ (< *dreh<sub>2</sub>*- ‘run’ [LIV 127]) beside Gk. δράπετης ‘deserter’<sup>5</sup> (note Skt. *dram*-, *drav*-, but no <sup>X</sup>*drap*-), and *snā-*<sup>6</sup> ‘bathe’ beside Lat. *Neptūnus* (cf. Ghosh 1982: 110), though this theonym is typically connected with Skt. *Apām Napāt*, Av. *Apqm Napāt*, and OIr. *Nechtan* (cf. Mallory & Adams 2006: 409) or derived from PIE \*nēbh<sup>h</sup>os ‘cloud.’<sup>7</sup> Hirt (1921: 132–133) proposes to derive the *-p*- in *-paya*- from a deictic particle *-p(e)* (the unstressed variant of \*pó ‘but, then again’; see Dunkel 2014: 622–624), as found in Latin *nem-pe*, *quippe* (< \*quid-pe) and Lithuanian *kaip*, *taip*, etc. (Ghosh 1982: 115–116), though it is not clear to us how such an adversative deictic particle could find its way into a causative formation. While such enlargements and particles may of course participate in the creation of new morphology through grammaticalization,<sup>8</sup> proposing that this is the case here is not really an explanation—it attempts to solve a mystery with another mystery, and, as above, does not address the fundamental question of why the *-paya*-causative was created in the first place, or again, why it was primarily used to form causatives in long-vowel roots.

The final, perhaps most widely held explanation is that the *-p*- of the *-paya*-causative has come into being by analogy. We of course know that analogy drives its spread, creating a more general, productive causative marker out of a suffix that was restricted to long-vowel roots (Insler 1987). But as the source of *-paya*-, more than one scholar has suggested to us that roots like *dīp-* ‘shine’, which forms a causative *dīp-aya*-, are the source of the Indic LECS, having transferred its *-p*- to the related root *dī-* ‘shine’ and thus providing a pathway for the insertion of *-p*- after long-vowel roots. In this case, however, as Ghosh (1982: 82) and Jamison (1983: 164<sup>122</sup>) point out, the root *dīp-*, which is attested rather late, is itself abstracted from the causative *dīpaya*-, and not the other way around. Likewise, one could argue that the *-p*- was perhaps extracted from the end of the very old secondary root *gup-* ‘protect’, but this seems unlikely, given that the causative is not attested until Epic Sanskrit (Whitney 1945: 37).<sup>9</sup> Similarly, Gaedicke (1880: 276) suggests that the *-p*- of *-paya*- comes from the

5 Ghosh (1982: 109). According to Beekes (2010: 331), the -π- is “unclear”. Reviewer #2 suggests that this form may have been a compound of δρᾶ- ‘run’ and πέτωμαι ‘fly, run’.

6 Ghosh (1982: 99–100).

7 Cf. IEW 316, Wodtko et al. (2008: 499).

8 See Campbell (2021) for discussion and references.

9 We thank Reviewer #2 for this suggestion.

final consonant of the root *lap-* ‘prattle’ via *lāpaya-* ‘make prattle’ (AV),<sup>10</sup> though it seems curious that the *-p-* would spread from such an uncommon source; indeed, we agree with Ghosh (1982: 72) that the number of roots ending in *-p* is “not important enough to determine the formation of a new suffix.” Charpentier (1912: 395) suggests that the *-paya-* causative was originally a factitive denominative formation built to nouns of the shape *-āpa-/apa-* (such as *kalāpa-* ‘collection’, *kaṇapa-* ‘type of weapon’), though the late evidence he puts forth makes it difficult to account for a formation that exists already in the oldest Sanskrit (Ghosh 1982: 113).

In short, there is no clear, obvious source for the *-p-* of the *-(ā)páya-* causative, analogical or otherwise.

### 3 The labial-extended causative suffix in Iranian

In Iranian, the inherited IIr. causative suffix is continued as *-aya-/aiia-* in the oldest languages. In Avestan, the causative suffix *-aiia-* productively attaches to a lengthened root; thus, the transitive present stem *s<sup>u</sup>ru-nau-* ‘hear’ forms passive *sru-iiā-* ‘be heard’ vs. causative *srāuu-aiia-* ‘make heard’ (Skjærø 2009: 89). The suffix *-aya-* is also found in Old Persian: *niy-ašādayam* ‘I established’ (Kent 1950: 30). In Khwarazmian, transitive verbs are regularly derived from Old Iranian causative *-aya-* with lengthened root vowel, e.g., *xwr-* ‘eat’ vs. *x'r-y-* ‘make eat, feed’ (Durkin-Meisterernst 2009: 349–350).

However, a new productive causative/denominative suffix appears in Middle Persian: *-ēn-* (Skjærø 2009: 213; Korn 2013: 45),<sup>11</sup> which attaches to intransitive stems (common; cf. *est-ēn* ‘to make stand, establish, place’), transitive stems (much rarer; cf. *ōzan-ēn* ‘to cause (somebody) to kill (somebody else)’), noun stems (e.g., *tarāzūg-ēn* ‘weight’), and adjectival stems (*āgāh-ēn* ‘make aware’). The suffix *-ēn-* is typically derived from *\*-aya-ana*, following Henning (1934).<sup>12</sup>

Aside from a few Tati varieties (Caspian), Takestani, Danesfehani, Khasneyni, and Ebrahim Abadi, where the causative suffix is *-ten-/den*, and in Tati Eshtehardi, where it is *-enden-* (Rasekh-Mohannad & Izadifar 2013), almost all

<sup>10</sup> From PIE *\*lep-* ‘babble’; see EWAiA 432–433 for discussion. Ghosh (1982: 71) argues that there is no root *lap-* ‘prattle’, rather *li-*, and that *lap-* was backformed from the causative *lāpaya-* (cf. *smi-* ‘smile’, caus. *smāpaya-*).

<sup>11</sup> The causative suffix *-ān-* rarely appears in Parthian, but this variant becomes more widespread in later Iranian.

<sup>12</sup> The reconstruction of Middle Iranian *-ēn- < \*ay(a)-ana-* goes back to Salemann, who believes that this innovation in Middle Iranian has a denominative origin (1901: 305).

other Iranian languages follow the Middle Iranian causative formants -(*V*)*n-* (Zolfaghari 2017). For instance, causatives in Khotanese are formed by adding the suffix *-āñ*:<sup>13</sup> *bam*- ‘vomit’ → *bam-āñ* ‘make vomit’ (Emmerick 2009: 391). In New Persian, a variant of this suffix continues to be used as a causative marker and is found in both intransitive and transitive stems, through the addition of *-ān/-on* in the past stems and of *-ān-(i)d/-on-(i)d-* in the present; cf. *ras*- ‘arrived’ → *ras-ān*- ‘made arrive’ and *rasid*- ‘arrive’ → *ras-on-(i)d-* ‘make arrive, take to’ (Windfuhr & Perry 2009: 448). In the northwestern Iranian language Dimli (aka Zaza), the inherited causative suffix surfaces as *-n-* (rarely *-ān-*), which is typically added to intransitive verbal stems; e.g., *geyr*- ‘go around’ → *geyr-n-* ‘show around’; *ters*- ‘be afraid’ → *ters-ān*- ‘frighten’ (Paul 2009: 556). In Kurdish, also a northwestern Iranian language, only intransitive verbs can be causativized by adding a suffix *ēn/-ān(d)-*, e.g., *tirs*- ‘be afraid’ → *tirs-ēn-/tirs-ānd-* ‘make fear, frighten’ (McCarus 2009: 604). In Kurdish Gurani, the causative suffix *-ān* attaches to past stems and *-en-* to present stems, e.g., *xaf-ān-∅* ‘made sleep’, *xaf-en-e* ‘make sleep’ (Bamshadi & Mirdehghan 2014: 13). Alongside the suffixing *-ēn-* to the present stem in Balochi, this language forms double causatives (or transitivizing transitives) by adding *-āēn-* in Southwestern Balochi, and *-āin-* in Eastern Balochi: *ras-āēn-ā* ‘cause to arrive’ (Jahani & Korn 2009: 659). In Sogdian, no special means is found other than a causative interpretation of transitive verbs (*xwēr* ‘feed’ vs. *xwar* ‘eat’), which was inherited from Old Iranian (Yoshida 2009: 304).

There also appears to be a cluster of Eastern Iranian (EIr.) languages which contain a LECS reminiscent of Indic *-(ā)páya-*. In Middle EIr. we find Khotanese *-ev-* (Emmerick 1968) and Khwarazmian *-wy-* (Samadi 1986), and in Modern EIr. Wakhi *-uv-* (Morgenstierne 1938: 497; Bashir 2009: 839), Parachi *-ew-* (Morgenstierne 1929: 86; Kieffer 2009: 694), Yidgha *-aw-*, Munji *-ov-*, Ormuri *-iw-*, and Pashto *-ew-* (Morgenstierne 1929: 101; Emmerick 1968: 375). In most languages, the LECS is the productive causative marker, but in Khwarazmian, it is quite rare, found at most with three roots: *pcr'wy-* ‘make warm’, *ÿwywy-* ‘make weep’, and *hnc'wy-* ‘to let rest’. In the first, *pcry-* ‘become warm’ → *pcr'wy-* ‘make warm’, the *-w-* may be inherent in the root itself (\**h₂reū-* ‘shine’; cf. Arm. *arew* ‘sun’, Skt. *ravi-* ‘sun(-god)'),<sup>14</sup> and so does not provide us with a clear case. Similarly in the second, *ÿy-* [xšai-] ‘weep, mourn’ → *ÿwywy-* [xšaiwaya-] ‘make weep’, the *-w-* may “point to a denominative origin” (Cheung 2007: 452; cf. Samadi 1986). The clearest instance is found to the root *hnc'ym* [han-čā-]<sup>15</sup> ‘to rest’ →

<sup>13</sup> From \*-ān-aya- (Konow 1932: 53).

<sup>14</sup> The initial *pc-* derives from the inherited prefix *\*pati-* (Cheung 2007: 192).

<sup>15</sup> Cf. Armenian *hangčim* ‘rest’ < \**kʷih-i-ske-* (LIV 393).

*hnc'wy-* [han-čā-waya-] ‘to let rest’, which likely derives from the laryngeal-final PIE root \**kʷyeh₂*- ‘rest’ (Samadi 1986: 91). In this causative, there is a clear, unexpected *-w-* added before the suffix *-aya-*.

Early scholarship considered these suffixes to be borrowed from the extended Sanskrit causative suffix *-(ā)paya-* or later Indic *-āvē-*. For example, Morgenstierne (1938: 497) argued that the consonantism found in the LECS within New Eastern Iranian languages, such as Wakhi *-uv-*, could not be derived from an Iranian /p/ but must have an Indic origin. This assumption was later supported by Bailey (1979), who asserted that Khotanese *-ev-* comes from Prakrit *-āvē-* < Sanskrit *-(ā)paya-*. On the other hand, Emmerick (1968: 187) argues against the borrowing scenario: “It is difficult to continue to believe that *-ev* is Ind[ic] in origin (< *-āpaya*) the more widespread it appears to have been in Ir[anian]. Of certainly Ind. origin, Kh[otanese] has *dukhev-*, *\*khijev-*, *upev-*, *samev-*, *\*suhev-*. In Ir. cf. Par[achi] *-ēw-*, Paš[h]tō, Orm[uri] *-aw-*, Munjī *-ōv-*, Yidgha *-iw-*, Waxī *-uw-*.” Schwartz (1969), following Emmerick 1968, also contends that the Khwarazmian suffix *-wy-* cannot be Indic in origin, as *-(ā)paya-* would have become *-by-* in this language.

For these reasons, Cathcart (2015: 42) has reconstructed an earlier causative/denominative suffix *\*-āuaia-* as a shared morphosyntactic property in Eastern Iranian. To our knowledge, there is currently no explanation for the origin of the unexpected *\*-(ā)u-*.

#### 4 The labial-extended causative suffix in Caspian

So far, we have seen how the Indic causative suffix *-(ā)páya-* is attested in the earliest Sanskrit and currently lacks a good explanation for how and why it came to be. We have also surveyed the Iranian languages, where a suffix *\*-(ā)uaia-* is to be reconstructed for Proto-Eastern Iranian, a suffix with both a similar shape and function to its Indic counterpart, and a suffix which also lacks a good explanation for its existence. In this section, we will present evidence that *\*-(ā)uaia-* is also continued in certain Caspian varieties, making the reconstruction of a suffix *\*-(ā)uaia-* for Proto-Iranian even likelier.

##### 4.1 Gilaki

Located in the northern region of Iran extended through the Caspian Sea litoral, Gilaki<sup>16</sup> is a sub-branch of the northwestern Iranian languages (Chris-

<sup>16</sup> The word Gilaki in this paper refers to (a) Gilaki as an epithet of a language and (b) the modern period of this language vs. an earlier one (c. 12th–15th cc.).



FIGURE 1 Map of Gilan, Iran

tensen 1930; Voegelin & Voegelin 1987; Windfuhr 2009a: 13–14; Rastorgueva et al. 2012), typically described as a Caspian or Dimli-Caspian language (Skjærøv 2017: 476). Scholars generally restrict Gilaki to the current administrative borders of Gilan (Gilaki meaning “the language of Gilan”; see figure 1) without considering the sociocultural and linguistic dynamics of the area over the centuries.

Based on this assumption, Gilaki is almost always categorized as a language with two (Stilo 2001, 2018; Austin 2008) and sometimes three (Samareh 1988) dialects, namely *biyeh pish* ‘eastern,’ *biyeh pas* ‘western,’ and *galeshi* ‘associated with shepherds; highland,’<sup>17</sup> spoken by only those speakers who dwell within the strictly defined borders of Gilan.<sup>18</sup> However, there are indigenous speakers

<sup>17</sup> One of the authors of this paper is a native of Lahijan, Gilan and was raised speaking a variety of Gilaki that is mostly spoken in the north of Lahijan (see figure 3 below).

<sup>18</sup> Thus, in this view, the categorization of the conventional language (and its dialects) sur-

spread out across the Caspian littoral (northern provinces), as well as in a small number of areas of northern Tehran, Alamout in Qazvin, Taleqan in Alborz, and the northern regions of Semnan.

In nearly all varieties of Gilaki, the structure of the causative resembles the causatives in many other Iranian languages, as discussed in section 3 above.<sup>19</sup> In the infinitive form, the Middle Iranian (MIr.) causative suffix *-Vn-*<sup>20</sup> attaches to the present stem of a verb followed by an infinitive marker (1.a). A past-tense marker is added in the past-tense form, followed by subject endings (1.b, 1.b'). For the future, the present causative stem follows the (conjugated) auxiliary verb of /xa/ 'want' (1.c, 1.c').

- |     |                                   |                             |                         |
|-----|-----------------------------------|-----------------------------|-------------------------|
| (1) | a. <i>kəf-Vn-en</i>               | vs. <i>kəf-en</i> 'to pull' | all dialects            |
|     | pull.PRS-CAUS-INF                 | pull.PRS-INF                |                         |
|     | 'to make pull'                    | 'to pull'                   |                         |
|     |                                   |                             | Roudbaneh               |
|     | b. <i>dəχuf-on-e-m</i>            |                             |                         |
|     | PRFX-extinguish.PRS-CAUS-PST-1SG  |                             |                         |
|     | 'I damped (fire), turned off'     |                             |                         |
|     |                                   |                             | Darmesar/Shoumshak/Asen |
|     | b'. <i>de-n-gərs-en-e-jämə</i>    |                             |                         |
|     | PRFX-NEG-turn_on.PST-CAUS-PST-1SG |                             |                         |
|     | 'I did not turn (it) on'          |                             |                         |
|     |                                   |                             | Rasht                   |
|     | c. <i>xa-jəm</i>                  | <i>və-tav-an-en</i>         |                         |
|     | want.PRS-1SG                      | PRFX-boil.PRS-CAUS-INF      |                         |
|     | 'I will make [the milk] boiled'   |                             |                         |

---

passes the lived experience of indigenous people of the area and their language identity. For example, in her book *Barresi-yi khurdi gütish-hä-yi mantiq-i Qasran* (an account of sub-dialects of Qasran region), Giti Deihim describes her encounter with a language in Lavasan (Rudbar-i Qasran), north of Tehran, which is not Farsi. Asking the indigenous people of Rudbar-i Qasran, they tell her that their language is Gilaki, "that they speak Gilaki" (2007: 3). However, she argues that while speakers call their language Gilaki, "[t]he language of Qasran region in contrast to what [indigenous] people say, is not Gilaki" (*ibid.*: 12).

<sup>19</sup> This paper focuses on the synthetic (suffixal) form of the causative. There is much literature that focuses on the other means of forming causatives in Iranian, such as periphrastic or compound verbs. See Stilo (2001), Rezayati Kisheh Khaheh & Sabzalipour (2008), Bashir (2009), Paul (2009), Windfuhr & Perry (2009), Yoshida (2009), and Sabzalipour (2015).

<sup>20</sup> In this study, the symbol ⟨V⟩ is used as a cover symbol for all variants of a vowel.

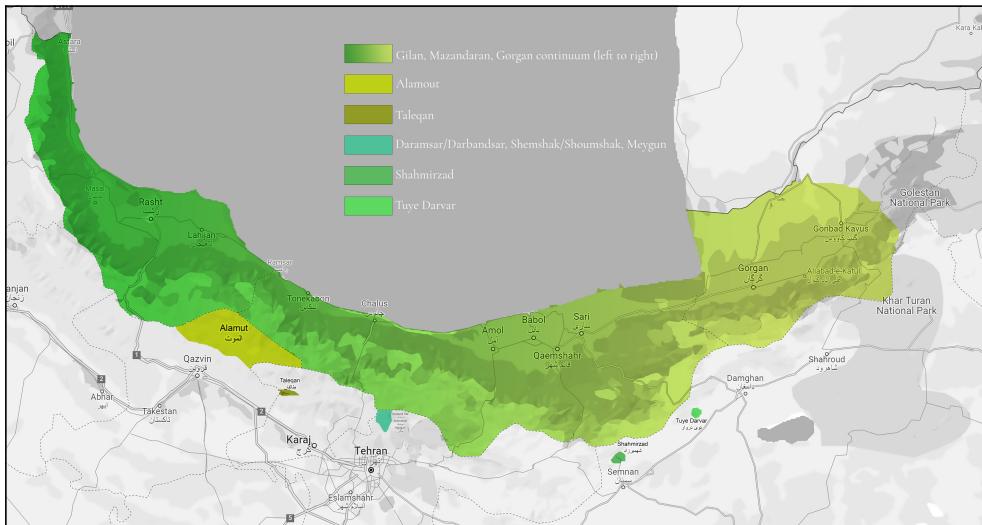


FIGURE 2 Geographical distribution of the Gilaki language

- c'. *xa də-mir-en-əm*  
want.PRS PRFX-drown.PRS-CAUS-1SG  
'I will make [something] drowned'

Let us now turn to the presence of the LECS in Modern Gilaki (Khoshhsirat 2018). In almost all varieties of Gilaki, the present causative stem is formed by adding the causative suffix to the verbal stem (2.a–c).

- (2) a. *kut-an-əm(ä)* Western Mazandaran-Eastern Gilan  
beat.PST-CAUS-1SG  
'I grind'
- b. *fo-sudʒ-on-əm* Roudsar  
PRFX-burn.PRS-CAUS-1SG  
'I cause to evaporate'
- c. *kuf-an-d-ambæ* Western Mazandaran  
kill.PRS-CAUS-PST-1SG  
'I killed'

However, in four<sup>21</sup> relatively adjacent locations (figure 3) in the north, south, and northeast regions of Lahijan<sup>22</sup> and Langaroud, the causative suffix in the present tense is not *-Vn-* but *-bVn-* or *-nVn-* (3.a–c).

Our investigation has discovered that the causative suffix *-nVn-* exists in the peri-urban area of Lasheidon Hokoumati. In the other three locations, peri-urban areas, rural areas, or areas outside the *haf məhəl* of Lahijan, the speakers use the LECS *-bVn-*.<sup>23</sup>

- (3) a. *fu-rus-be(:)n-əm* Roudbaneh  
 PRFX-scrub.PRS-CAUS-1SG  
 'I cause to scratch'
- b. *də-xuf-ne(:)n-əm* Lasheidon Hokoumati  
 PRFX-dry.PRS-CAUS-1SG  
 'I make dry out'

<sup>21</sup> One of the authors of this paper has heard an instance of the LECS in Siahkal, southwest of Lahijan, from a local speaker.

<sup>22</sup> Interestingly, the LECS is predominantly used in peri-urban and/or rural areas to the north and south of Lahijan but not in the city's historic urban neighborhoods, known as *haf məhəl* 'seven neighborhoods.' As in most Gilaki varieties, *-Vn-* is the causative suffix used by some speakers of Lahijan's *haf məhəl* for past and future tense (see also Jahangiri 2015: 376–377). The difference lies in the present tense, where *-Vn-* comes before *-n-*, the present-tense marker, which supposedly can be enregistered (Agha 2007) as an index of the speakers of the *haf məhəl*.

(1) *sudʒ-an-en-əm* *haf məhəl* of Lahijan  
 burn.PRS-CAUS-PRS-1SG  
 'I burn [something]'

Interestingly, in Chaboksar and Tonekabon some speakers add the present-tense marker *-n-* to the past tense causative stem (cf. Persian and Tajiki in Windfuhr & Perry 2009: 447–448):

(2) *sudʒ-ə-d-an-əm* Chaboksar (Ghasem Abad)  
 burn.PRS-CAUS-PST-PRS-1SG  
 'I burn [something]'

(3) *sudʒ-ən-d-an-əm* Tonekabon  
 burn.PRS-CAUS-PST-PRS-1SG  
 'I burn [something]'

<sup>23</sup> In our analysis, we interpret the final *-n-* of *-bVn-* and *-nVn-* as part of the LECS, given that it is also found in other LECS variants throughout the Caspian region and therefore may be reconstructed as part of the suffix, deriving from the more common Iranian causative suffix *-Vn-*. However, one could (synchronously) interpret the final *-n-* as a present marker, which perhaps explains why the LECS is only found in the present tense in Gilaki.

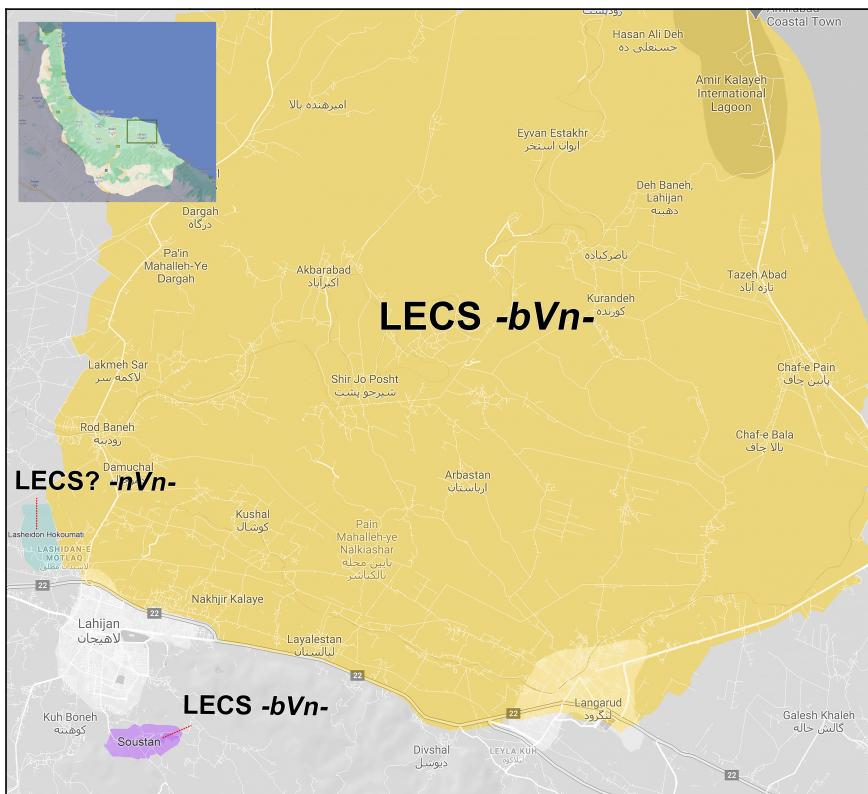


FIGURE 3 Distribution of LECS vs. -Vn- (Lahijan-Langaroud) in Gilaki

- c. *kə-be(:)n-əm*  
**pick.PRS-CAUS-1SG**  
'I pluck'

Soustan

Given that the meaning of the stems cited above is not always explicitly causative, a skeptical reader may question whether the function of the suffixes *-bVn-* and *-nVn-* is in fact a causative one. However, compelling evidence that both *-bVn-* and *-nVn-* are causative suffixes may be found in variant dialectal forms of the causative infinitive *dərgənVn* 'to thrust upon, to saddle with.' In these forms, the *-ən-* sequence of *gən-* in the present tense seems to have been reinterpreted through folk etymology as the causative suffix *-Vn-*, which in the relevant dialects has been replaced by the LECS (4.b-c, cf. 4.d), demonstrating that this suffix is indeed causative in function.

- (4) a. *dər-gən-Vn* all dialects  
 PRFX-hit.PRS(.CAUS)-INF  
 'to thrust upon, to saddle with'
- b. *dər-g-be:n-əm* Langaroud (Central/South)  
 PRFX-hit.PRS.-CAUS-1SG  
 'I thrust upon [someone]'
- c. *dər-g-ne:n-əm* Lasheidon Hokoumati  
 PRFX-hit.PRS.-CAUS-1SG
- d. *də(r)-gən-əm(ā)* Sakhtsar  
 PRFX-hit.PRS.CAUS-1SG

#### 4.2 Older stages of Gilaki and Tati-Talyshi

The Gilaki LECS is not a recent creation, but is attested in the earliest Gilaki texts,<sup>24</sup> namely *Tafsīr-i Daylāmī* (c. 13th century) and the Gilaki translation of *Maqāmāt-i Ḥarīrī* (c. 11th–14th c.).<sup>25</sup> In both Old Gilaki texts there is evidence for two types of synthetic causative suffixes. The first is the expected general causative suffix *-Vn-*, which occurs the most frequently.<sup>26</sup> The second suffix is of an unfamiliar shape, */-omen-/*, for which there has already been some speculation on its function and origin (Mohammadi 2015). There are many clear examples of this suffix, such as in the verb *ha-hizumanan* ‘incite, enkindle’ (*Tafsīr-i Daylāmī*),<sup>27</sup> which Mohammadi (2015: 155; contra Borjian 2009: 113)

<sup>24</sup> Since 2016, one of the authors of this paper has had conversations with specialists in the Caspian region texts, including Farhang Karami, Meysam Mohammadi, and Ali Zabihi, among others.

<sup>25</sup> Since Kia (1948) first published the old texts of the Caspian region, followed by Emadi Haeri’s prologue (in *Abū l-Fadl b. Ṣahrāwīr Daylāmī* 2008) and Borjian’s article (2009) on *Tafsīr-i Daylāmī*, the debate has not yet been resolved whether the language of these texts is Tabari (the old language of Mazandaran) or Daylāmī/Gilaki (the old language of Gilan). Despite all the disparities in the name of the language inside and outside academia, most scholars agree that the commentators or translators of these texts are presumably from somewhere between Lahijān and Shāhsāvār. While the name may be important to the indigenous people of the region, it is not relevant to the arguments of this section.

<sup>26</sup> Cf. *hukahanen* ‘to bring down’, *huazanen* ‘to descend’ (*Maqāmāt-i Ḥarīrī*: 146); *vagardanina* ‘[we] make them return’ (*Tafsīr-i Daylāmī*: 26), *rasanen* ‘to make [their work] done’ (*Tafsīr-i Daylāmī*: 39).

<sup>27</sup> It should be noted that Mohammadi (2012, 2015) transcribes the Middle Gilaki causative suffix as *-uman-* (e.g., *ha-hizumanan*). However, since the phonetic value of some vowels in the Early Modern Gilaki texts is not exact, other readings are possible. The variability

interprets as follows: “*hā* is a preverb, *hīz-* is a present stem, *-uman-* is a suffix which attaches to the present stem and forms a transitive stem, *-an* is a third plural marker.” In this example, Mohammadi considers */-omen-/* to be a transitivizing suffix that attaches to the present stem. While one might assume the final *-n* in */-omen-/* to be a present tense marker (as it functions in some Modern Gilaki varieties), this is unlikely, as there are instances of the *-n-* occurring in the infinitive such as /ha-hizomenan/ ‘to enkindle’ (*Maqāmāt-i Ḥarīrī*: 86R) and /arsomenen/ ‘to make exhausted/bored’ (*ibid.*: 100R). Three additional verbs with the suffix */-omen-/*, which Mohammadi (2012: 243) cites as having “transitive stem markers” should be added to the list: *basudžumanim* ‘we enkindle them’, *banifumanim* ‘we make them sit’, and *bagardumana* ‘[3SG] causes to turn’.<sup>28</sup> Of course, */-omen-/* may just as easily be interpreted as a causative marker that is added to intransitive stems.

There is also evidence for the LECS in two languages closely related to Gilaki: Tati and Talyshi.<sup>29</sup> Categorized as a subgroup of Northwestern Iranian languages, Tati and Talyshi (or what Stilo calls Tatic, grouping both languages under one name due to their close connection; see Stilo 1981, 2015: 413) have speakers distributed around (a) “the southwest Caspian littoral and higher altitude of the mountain slopes that face the coast both in Iran and the very southern areas of the Azerbaijan Republic” (Stilo 2015:413–414) and (b) “mountains passed that lead the Iranian Plateau ... [which] is the domain of Tati group” (*op. cit.* 414).<sup>30</sup>

In addition to the general Iranian causative suffix, we find the following causative markers in Talyshi: Northern Talyshi / Northern Tati *-ovn-*, *-own-*, *-öwn-* and Southern Talyshi / Central Tati *-amen-*, *-amən-*, *-əmən-* (Stilo 2015:

of the vowel readings is primarily a matter of orthography: in the modified Arabic script used in Early Modern Gilaki texts, the causative suffix is written و من (sometimes with diacritics), which can be interpreted as *oman*, *omən*, *omaen*, *omon*, *umen*, *umən*, *uman*, etc. but not as *āmen* or *amen*. Considering the variable presence of *ā*, it is plausible to assume that this letter represents [o]; thus, the four most likely pronunciations are [omen], [omən], [oməen], or [omən]. Given the /e/ vocalism present in the Modern Gilaki forms, we are inclined to posit [omen]. In our study, we place the suffix within slashes */-omen-/*, as phonemes are conventionally transcribed.

- 28 The sequence cited as *-uman-* in these three forms by Mohammadi should be interpreted as */-omen-/*, as discussed above in fn. 27.
- 29 The authors of this study are aware of the concerns of some speakers about the possibly derogatory etymology and origin of the name Tati (Clauson 1972; see also Stilo 2015: 413<sup>3</sup>). In some places, like Roudbar (Gilan), speakers use the name of their village or an identifiable place to which they belong as the name of the language. For example, some speakers of Jirandeh call their language Jirandehei ‘of Jirandeh’.
- 30 For more comprehensive background and analysis, see Stilo (2015).

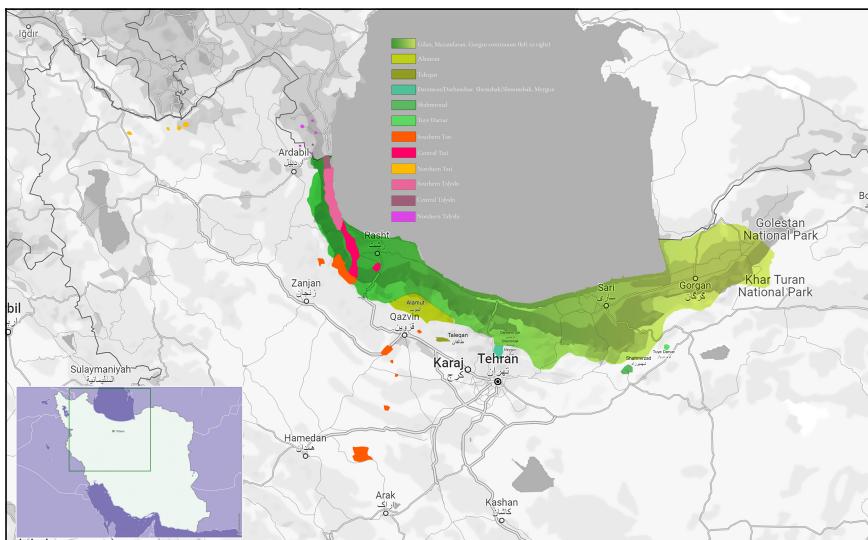


FIGURE 4 Geographical distribution of Tati-Talyshi and Gilaki  
ADAPTED FROM STILO 2015

432). We have also discovered that some speakers of Talyshi in Fouman (which is categorized as Southern Talyshi) use *-ovn-*.<sup>31</sup> Given the geographic proximity between Gilaki Foumanati and Talyshi Foumani, one could argue that */-omen-/* was originally an old Southern Talyshi suffix that diffused into Old Gilaki during the Middle Iranian period, though the /o/ vocalism makes such a borrowing unlikely (cf. Stilo 2015: 422–424). Indeed, it appears likelier that the Tati-Talyshi LECS and Gilaki LECS do not derive from each other, but rather from an earlier suffix in Proto-Caspian.

But what would such a suffix look like? The oldest evidence we have is found in Old Gilaki */-omen-/*, which shares the same consonants as Southern Talyshi / Central Tati *-amen-, -amən-, -əmən-*. However, the Northern Tati and Talyshi causative suffixes *-ovn-, -own-,* and *-öwn-* provide a clue as to how we should reconstruct the initial consonant of our hypothetical Proto-Caspian LECS, suggesting an earlier *-w-* in the suffix. This is further bolstered by the presence of both *-o-* and *-a/ā-* in the initial vocalism of the suffix, for the phonological change *\*-ā- > -o-* is conditioned by *\*-w-* in Gilaki, as illustrated through a comparison of New Persian *gāv*, Gilaki *gow*<sup>32</sup> ‘cow’ and New Persian *nāv*, Gilaki

<sup>31</sup> We are not sure if the individuals were speakers of Northern Talyshi dwelling in Fouman. According to Don Stilo (p.c., 08/07/2018), the existence of *-ovn-* in Southern Talyshi is unexpected.

<sup>32</sup> Contrast the variant *gāb* without the change of *\*-ā- > -o-* before [b].

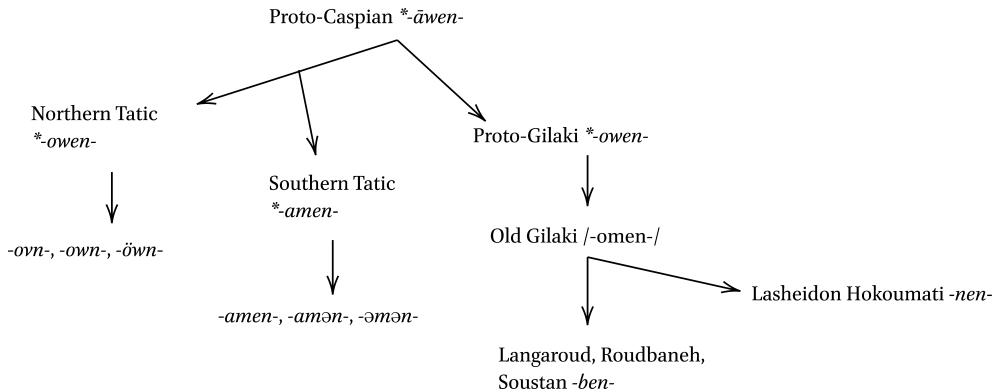


FIGURE 5 Evolution of the labial-extended causative suffix in Caspian

now ‘navy, boat’.<sup>33</sup> Inherited \*-ā- does not however become -o- before -m- in Gilaki, as shown by the change of PIr. \*zāmātar- ‘son-in-law’ (Av. zāmātar-) to Gilaki *damad* / *zama* ‘son-in-law, bridegroom’, not *xdomad* / *xzoma*.<sup>34</sup> Thus, the change in vocalism from \*-ā- > -o- in /-omen-/ requires the original Proto-Caspian form to have contained \*-w-, not \*-m-. As for the change of -w- to -m- in Southern Talyshi, Central Tati, and Early Modern Gilaki, we may attribute this to the not uncommon sporadic change of -v/w- > -m-: \*-owen- > -omen-.<sup>35</sup>

All of this leads to a reconstruction of \*-āwen- as the most likely reconstruction for the Proto-Caspian LECS. As for the Modern Gilaki suffixes *-bVn-* and *-nVn-*, the Old Gilaki form /-omen-/ provides a straightforward starting point for the suffixes *-bVn-* and *-nVn-*, with the change of *-mVn-* to *-bVn-* as an example of distant nasal dissimilation and the change to *-nVn-* as distant place assimilation or perhaps as the analogical spread of the more common *-n-* causative marker. These proposed changes are illustrated in Figure 5 above.

Nevertheless, some sticking points remain outside of Gilaki for this Proto-Caspian reconstruction. In our investigation, we encountered an unusual case in Harzani, a Northern Tati dialect (Stilo 2015: 439), where the LECS *-ovn-* is added to an intransitive stem to create a transitive stem, e.g., *toros-* ‘fear’ vs. *tors-ovn-* ‘scare’. However, if the second syllable of a disyllabic stem contains *-Vm-*,

<sup>33</sup> Cf. also PIr. \*hvāp- ‘sleep’ (Bailey 1979: 495; Cheung 2007) > MP *xwāb* ‘id.’ (MacKenzie 1971: 95), Gilaki *xow* ‘id.’ and OIr. *āp-* ‘water’ > NPers. *āb*, Gil. *ow* / *āb* ‘id.’.

<sup>34</sup> See Horn (1983: 532); Mallory & Adams (2006: 206–207). Note that inherited vowel length is no longer contrastive in Gilaki. For another example of inherited -a- before -m- not becoming -o- in Gilaki, cf. *damH-* ‘blow, breathe; swell’ > *dāmæstæn* (Cheung 2007: 56).

<sup>35</sup> Examples of *w/v* > *m* include *nəvā kudən* > *nəmā kudən* ‘should not be done’, \**kasayapa-* > *kävuf* > *kämuf* ‘turtle’, *gəvəndz* > *gəməndz* ‘deer’ (MP *gawazn*; MacKenzie 1971: 36).

this sequence is deleted when the LECS is added: Harzani *beræmesd-* ‘cry’ → *ber-öwn-esd-* ‘make cry’, vs. expected <sup>X</sup>*beræm-öwn-esd-*. It seems that the verb is reanalyzed/resegmented to *ber-æm-*, and the second segment is replaced by *-öwn-* through causativization.

Why does this verb undergo such processes? Perhaps, as Stilo suggests,<sup>36</sup> the Southern Talyshi and Central Tati causative formants *-amen-*, *-amən-*, *-əmən-* and the Northern Tati and Northern Talyshi *-ovn-*, *-own-*, *-öwn-* should all be derived from <sup>\*</sup>*āmen-*: <sup>\*</sup>*āmen-* > <sup>\*</sup>*amən-* > <sup>\*</sup>*omn-* > *-ovn-* ~ *-own-* ~ *-öwn-*.<sup>37</sup> Vafsi, a dialect of Southern Tati, provides several examples of this sound change: *dawan* < *daman* ‘skirt’; *zawa* < *zama* ‘son-in-law’; *div* < *dim* ‘face’; and *dæv-* < *dam-* ‘blow’. However, due to having undergone strong influence from Kurdish (where *-m-* regularly becomes *-v/w-*) and being geographically far from Northern Tati dialects, Vafsi is not a good candidate for the source of this sound change. Nevertheless, as Stilo (p.c) notes, there are sporadic examples of *-m-* > *-v/w-* in general Tati.

Similarly, Zarshenas & Asefi (2019: 9–10) propose a sound change *-m-* > *-v-* and suggest that *-ām(a)n-* and *-āv(a)n-* are in origin the same causative morpheme, with the former being older. According to them (*ibid*, 8, 11), following Henning’s reconstruction (1934) of Middle Iranian *-ēn-* < <sup>\*</sup>*ay(a)-ana-*, it is credible to derive the Talyshi causative suffix *-ām(a)n-* from the inherited causative <sup>\*</sup>*aya-* + middle participle *-mna-/mnā-*. While on the surface plausible, it should be noted that Henning (1934: 212) does not consider *-ana-* to have originally been a participle, but rather a nominal suffix. Indeed, as Agnes Korn points out to us,<sup>38</sup> the second segment *-ana-* cannot be a participle suffix, since in Old Iranian the suffix is *-āna-*, and <sup>\*</sup>*aya-āna-* would not have become *-ēn-* in Middle Iranian. Moreover, it makes little sense to use a middle voice participle in a causative formation. Taking these details into consideration, the reconstruction <sup>\*</sup>*aya-mnā-* for the Talyshi (and Tati) causative formants *-amen-*, *-amən-*, *-əmən-* and *-ovn-*, *-own-*, *-öwn-* appears unlikely. To add to this, recall that the oldest attested Caspian LECS is */-omen-/*, not */-āmen-/*, which all but rules out the derivation from <sup>\*</sup>*aya-mna/mnā-*. Thus, our proposed reconstruction of <sup>\*</sup>*āwēn-* is the likelier one, and the Northern Tati-Talyshi forms *-ovn-*, *-own-*, *-öwn-* are more archaic than their Southern Tati-Talyshi counterparts.

But what about the unusual case of haplography in the Harzani causative above (*beræmesd-* ‘cry’ → *ber-öwn-esd-* ‘make cry’)? Stilo (p.c.) suggests that the

36 P.c., 11/12/2021.

37 This may be supported by sporadic examples of the sound change *m* > *v/w* in Tati dialects and a few other Iranian languages like Bakhtiari (Southern Luri).

38 P.c., 12/01/2021.

sequence *\*-am-ām-* became *-ām-* through haplology, followed by the sound change *-m-* > *-v/w-*: *\*beram-āmn-/berām-āmn-* > *\*berāmn-* > *\*berovn-* > *berōvn-*. Likewise in Northern Talyshi *pandom-* ‘swell’: *\*pandām-āmn-* > *\*pand-āmn-* > *pand-ovn-* (with *\*ā* > *o*). Such a hypothesis requires the oldest form of the LECS to have had *-m-* consonantism, just as in Middle Gilaki */-omen-/* and the Southern Talyshi and Central Tati forms *-āmen-*, *-amən-*, *-əmən-*. This, of course, goes against our reconstruction of the Proto-Caspian LECS *\*-āwen-*. But one could explain the change of *berem-ōvn-esd-* → *ber-ōvn-esd-* as being, like haplology, an instance of dissimilation *-Vm-own-* → *-own-*, as the [labial] and [nasal] properties of the segment /m/ are also present in the sequence /own/, which triggers the deletion of a syllable.

Finally, how might EIr. *\*-āuāja-* be related to Middle Caspian *\*-āwen-*? The change of *-āja-* to *-ē-* is common in the Middle and Modern Iranian periods (Gray 1902: 117) and is also attested in forms such as Gilaki *pāre(:)roz* ‘the other day’ (< MIr. *parēr*; cf. Av. *paouruua+aīiarə*). The *\*e* of the second syllable was thus most likely originally long: *\*-āwēn-*. As for the suffix-final *-n-*, this is most straightforwardly understood as an analogical extension of the more common causative suffix *-(V)n-*. This type of secondary causative marking is found elsewhere in Iranian, such as in Mindjani *dəraw-åw-ān-* ‘to terrify’ (Morgenstierne 1938: 149). In short, *\*-āwēn-* may indeed be connected to these Eastern Iranian forms, all deriving from an earlier PIr. *\*-āuāja-*.

## 5 The labial-extended causative suffix in Proto-Indo-Iranian

Let’s review what we have discussed thus far. First, we have seen that Indic *-(ā)pāya-* is attested in the earliest Sanskrit texts, and there is currently no good explanation for the *-p-* extension added to the expected, inherited *-āya-*. A causative suffix *\*-(ā)waya-* appears to also be found throughout Eastern Iranian, most certainly in Khwarazmian *-(')wy-*, and possibly in Khotanese *-ev-* and other suffixes found in various modern EIr. languages, though some of these may be borrowed from Indic. Lastly, there is a labial-extended causative suffix (LECS) found in the earliest Gilaki texts, which is also continued in certain varieties of Modern Gilaki and is present in several closely related Tati-Talyshi varieties, all of which appear to descend from Proto-Caspian *\*-āwēn-*. The Proto-Caspian form itself very straightforwardly derives from an earlier *\*-āwaya-*, with the monophthongization of *\*-aya-* to *\*-ē-* and the analogical addition of the productive causative suffix *-n-*.

Could Iranian *\*-(ā)uāja-* and Indic *-(ā)pāya-* come from the same source? We believe that they do, given their remarkable similarity in form and func-

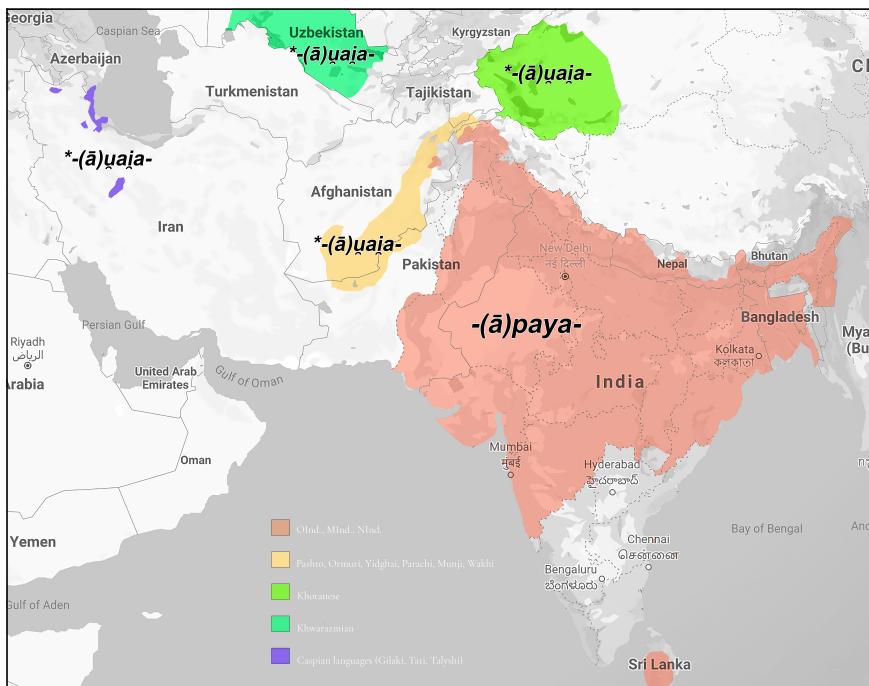


FIGURE 6 Attested distribution of the Indo-Iranian LECS

tion, especially considering the lack of any good explanations for the source of Iranian *\*-(ā)uaja-* and Indic *-(ā)páya-*. Though the account we set forth below is necessarily speculative, we believe that it provides the most plausible explanation for the existence of the LECS alongside the expected inherited causative suffix *-ája-* in both branches. It also accounts for the restriction of the LECS in the earliest Sanskrit to primarily long-vowel roots, as well as why in Khwarazmian it is found (solely?) in the inherited laryngeal-final root *\*kʷjeh₂-* (> *hncy-* ‘rest, relax’ → *hnc'wy-* [han-čā-waya-] ‘calm down’; Samadi 1986: 91), whose causative is reconstructed by LIV (393) for PIE, with a cognate in OCS *po-koj*, inf. *po-koiti* ‘let rest’.

Before the formation of Indo-Iranian, the Proto-Indo-European suffix *\*-éje/o-* was used to produce secondary causative stems to unaccusative base verbs (Bozzone 2020). This suffix was productively added to *o*-grade roots of all shapes, such as *\*mon-éje-* ‘make think’ and laryngeal-final roots like *\*kʷjoh₂-éje-* ‘make quiet’ and (secondarily, since neither root was unaccusative) *\*stoh₂-éje-* ‘make stand’, *\*doh₃-éje-* ‘make give’.<sup>39</sup> As PIE slowly evolved into Proto-

39 For simplicity of presentation, we assume in our reconstructions that the initial *-e* vowel

Indo-Iranian, it is typically thought that the three laryngeals  $*h_1$ ,  $*h_2$ ,  $*h_3$  merged into a single  $*H$  (Kobayashi 2004: 131; Lubotsky 2018: 1881), most likely a voiceless glottal fricative /h/ (Kümmel 2014). At this stage, our aforementioned laryngeal-final root causatives were pronounced as  $*kʷ̥iḥéje-$  ‘make quiet’,  $*stohéje-$  ‘make stand’, and  $*dohéje-$  ‘make give’.

The phoneme  $*h$  was ultimately lost in all Indo-Iranian languages, though it did leave traces in both initial and intervocalic positions (Beekes 1981; Lubotsky 1995, 2018: 1882; Cantera 2017: 487). With normal changes in vocalism, the loss of an intervocalic  $*h$  would produce the expected forms  $*kʷ̥iāja-$  ‘make quiet’,  $*stāāja-$  ‘make stand’, and  $*dāāja-$  ‘make give’, each containing marked sequences of hiatus. Contraction of the root and suffixal vowels was to be avoided, as speakers were accustomed to having a boundary between the root and the causative suffix. In short, the avoidance of vowel hiatus and avoidance of root/suffix contraction prompted the insertion of a consonant between the root and causative suffix (cf. Jamison 1983: 80), thereby leading to the creation of the LECS, the source of both Indic  $-(\bar{a})páya-$  and Iranian  $*-(\bar{a})uaia-$ .<sup>40</sup>

## 6 Possible phonetic paths for the creation of the LECS

But why does the LECS contain a labial consonant? In Proto-Indo-Iranian, mid vowels merge with /a/, and so PIE  $*o$  would become  $*a/\bar{a}$ , whose length depends on the shape of the syllable (Kobayashi 2004: 26–27). We postulate that in forms such as  $*stohéje-$  ‘make stand’, as the merger was taking place, the labiality of the  $*/o/$  vowel was transferred to the following laryngeal segments.<sup>41</sup> This produced a secondary phonemic split after the loss of the condi-

---

of the PIE causative suffix was not colored by preceding laryngeals, though this is far from assured. If coloring did occur in this context, this would complicate matters for forms such as  $*doh_3éje-$  ‘make give’, since it is clear that  $*-oHō-$  (at the very least  $*-oHó-$ ; see below) did not become  $*-oHʷo-$ . Of course, if coloring had occurred in this environment ( $*doh_3éje- > *doh_3óje-$ ), then one could reasonably assume that the suffix  $*-óje-$  was analogically replaced by the more common variant  $-éje-$  ( $*doh_3óje- \rightarrow *doh_3éje-$ ).

- 40 As Lubotsky has remarked (p.c., 02/20/2018), a further possible explanation is that by the time that  $*H$  had disappeared phonetically and caused hiatus, the causative morpheme  $*-aia-$  was no longer divisible from the root.
- 41 Cf. also Jasanoff’s Law (Jasanoff 1988: 73, Weiss 2020: 124), whereby the PIE sequence  $*-óHe# > *-óHu > *-óu$ . Jasanoff’s Law and the one proposed here are, of course, not connected to each other; rather, this process demonstrates how labialization can spread in similar environments.

tioning environment:<sup>42</sup> \**stohéje-* > \**stāhʷája-*, with the sequence \**hʷ* perhaps pronounced as a voiceless labiovelar glide [ℳ].

In Indic, the rounded laryngeal in forms such as \**stāhʷája-* underwent fortition to \**st(h)ápája-* > *sthápáyati* ‘makes stand’ (RV x.151.5; AV XII.3.51, XIX.46.4)<sup>43</sup> and spread to other stems to prevent further instances of hiatus (such as *arpáyati* ‘fits in’ to *r* ‘fit together’; Jamison 1983: 80) or other marked sequences (*jāyáyati* → *jápáyati* to *ji* ‘conquer’; Insler 1987: 64–65, cf. Jamison 1983: 135). In Iranian, the rounded laryngeal in forms such as \**stāhʷája-* exhibited a merger with the already existing voiced labiovelar glide \*-*ɣ-*, exactly as has happened in most English dialects today (*which, witch* = [wɪtʃ], etc.).

Our proposal explains the presence of the two causative suffixes in Indo-Iranian, as well as why -*p-* served as a “hiatus breaker” in roots with long ā (< \*-*eH-*) roots in Sanskrit (Jamison 1983: 80). However, a glaring problem lies with the most archaic Iranian languages, especially Gathic Avestan: if \*-āaīja- is to be reconstructed for Proto-Iranian, why is it not found in its oldest daughter language? Why do we not find Av. <sup>x</sup>*stāuuaiia-* (< \**stāhʷája-* < \**stoh₂éje-*) ‘make stand’ or the like? Ordinarily, we would expect such a form to surface as *stāaiia-*, with the intervocalic laryngeal \**H* lost (or perhaps not written; cf. Beekes 1988: 89) and hiatus maintained, which is commonly used as a metrical device (Kümmel 2014). But such a form does not exist. While causatives are certainly found in the Gāthās (*bānaya-* ‘make ill’, *srāvaya-* ‘make heard’, *vātaya-* ‘inspire’; Beekes 1988: 171), there are no causatives formed to long-vowel roots: there is no <sup>x</sup>*stāaiia-* or the like. Indeed, the only examples of a causative formed to a long-vowel root are found in Young Avestan, the clearest one being formed to this exact root: YAv. *stāiā-* ‘I put’, OP -*astāyam* ‘I put’ (< \**stoh₂éje-*; LIV 590), with the expected loss of laryngeal and merger of vowels.<sup>44</sup>

It is not immediately clear why forms such as YAv. <sup>x</sup>*stāvaiā-* ‘I put’, OP <sup>x</sup>-*astāvayam* are not attested. Perhaps vowel contraction was permitted in a common form such as ‘to put’ but was avoided elsewhere for morphological clarity, though this is not likely given that YAv. (*ā, fra*) *snaiāt* ‘wash off’, causative of *snā-* (< PIE \**sneh₂-* ‘bathe, swim’ [LIV 572]) also shows the loss of laryngeal in this configuration.<sup>45</sup> But it is possible that forms like YAv. *stāiā-* and

42 Cf. Beddor (2009), Miller & Trask (2015: 73).

43 A similar fortition has occurred in Gilaki; cf. Ir. *ahu* ‘gazelle’ > Gil. *apu*.

44 YAv. *juuaitānt-* (Yt. 19.11), a causative formed to *juua-* ‘live’, should not be considered relevant here, as it derives from the adjective *juua-* ‘living’ (< \**gʷih₃uo-*; Kümmel 2014), not \**gʷih₃-éje-*.

45 Metrically \**snaHaiia-* (Kellens 1990: 147<sup>41</sup>).

OP *-astāyam* provide a clue as to the earlier state of affairs.<sup>46</sup> For as Kümmel (2018) argues, whereas *\*h<sub>1</sub>* and *\*h<sub>2</sub>* merged as /h/ in Proto-Indo-Iranian and triggered the aspiration of preceding stops (cf. Ved. *path-*, Av. *paθ-* ‘path’ < *\*pn̥th<sub>1/2</sub>*), there is no evidence that *\*h<sub>3</sub>* did so, suggesting that *\*h<sub>3</sub>* remained distinct, with a pronunciation more akin to [y]. Following Kümmel’s reconstruction of the laryngeals,<sup>47</sup> there would therefore not have been two variants of the causative suffix in Proto-Indo-Iranian, but rather three:

- 1) PIE *\*-éje-* > PIIr. *\*-ája-*
- 2) PIE *\*-h<sub>1/2</sub>-éje-* > PIIr. *\*-hʷája-*
- 3) PIE *\*-h<sub>3</sub>-éje-* > PIIr. *\*-γʷája-*

With the loss of coda laryngeals occurring relatively early in Indo-Iranian, roots such as *\*dʰah-* ‘put’ and *\*day-* ‘give’ would frequently have been pronounced with long vowels by compensatory lengthening (*\*dʰā-* and *\*dā-*). Once speakers internalized these laryngeal-final roots as long-vowel roots after the loss of laryngeals in coda position, it was only natural to analyze *\*dʰāhʷája-*, the causative to *\*dʰā-* ‘put; do’, as *\*dʰā-hʷája-*, not the historically correct *\*dʰāhʷája-*. Likewise, speakers would naturally have analyzed *\*dāyʷája-*, the causative to *\*dā-* ‘give’, as *\*dā-yʷája-* and not *\*dāyʷ-ája-*.

---

	<i>*dá-dʰā-ti</i> ‘puts; does’		<i>*dá-dʰā-ti</i>
<i>*/dʰah-/</i> ‘put; do’	<i>*á-dʰā-t</i> ‘has put; has done’	→	<i>*/dʰā-/</i> <i>*á-dʰā-t</i>
	<i>*dʰāhʷ-ája-ti</i> ‘makes put; do’		<i>*dʰā-hʷája-ti</i>
	<i>*dá-dā-ti</i> ‘gives’		<i>*dá-dā-ti</i>
<i>*/day-/</i> ‘give’	<i>*á-dā-t</i> ‘has given’	→	<i>*/dā-/</i> <i>*á-dā-t</i>
	<i>*dāyʷ-ája-ti</i> ‘makes give’		<i>*dā-yʷája-ti</i>

---

FIGURE 7 Creation of the LECS through metanalysis in Proto-Indo-Iranian

It is in this way that the consonant was reanalyzed as part of the suffix, leading to its spread to other roots in later stages of the language. In Indic, *\*-hʷája-* was generalized as the go-to suffix for laryngeal-final roots, ultimately becoming *\*-pája-*. Indeed, it is striking that all of the long-vowel roots which form -páya- causatives in Vedic end in *-h<sub>1/2</sub>-*, with no good early examples of -páya- causatives to roots ending in *-h<sub>3</sub>*. The causative of *pā-* ‘drink’ (< *\*peh<sub>3</sub>(i)-*) is

46 Many of these ideas were suggested to us by Dieter Gunkel (p.c., 7/5/2021), to whom we are indebted for his thoughtful feedback.

47 Cf. Gunkel (2022).

famously *pāyáya-* (RV 5×, AV 2×), not *xpāpáya-*, and following Jamison (1983: 169, 172), the forms *jñapáya-* (AV, 1× to *jñā-* < \**ǵnoh₃-* ‘know’) and *dápáya-* (AV, 1× to *dā-* < \**deh₃-* ‘give’) should be considered secondary.<sup>48</sup> Thus, it is not without reason that the variant suffix *\*-γʷája-* (> *X-váya-?*) was lost and ultimately replaced by *-páya-* (< \*-*hʷája-*), a suffix that was exceedingly more frequent in the language.

But in Iranian, the variant *\*-uaja-* (< \**γʷája-*) was maintained and generalized in many Iranian languages, ultimately becoming extended to *\*-āuaja-*, with the long-vowel vocalism of the root being reanalyzed as part of the suffix through metanalysis (as occurred in later Indic). Perhaps *\*-uaja-* was extended to all long-vowel roots in Eastern Iranian and Proto-Caspian because the suffix-initial consonant was lost in the other variant *\*-hʷája-* (< \**h₁₂éje-*), once again leading to hiatus in the causative of many long-vowel roots. This is evident in more conservative varieties like Avestan and Old Persian, where *\*-hʷája-* became *\*-ja-* with laryngeal loss and vowel elision in the previously mentioned forms *\*stāja-* (YAv. *stāiā-*, OP *-astāyam*) and *\*snāja-* (YAv. (*ā, fra*) *snaiiāt*), which were both originally *\*h₂*-final roots. Or perhaps *\*-γʷája-* had been generalized to all laryngeal roots in Proto-Iranian for another reason, with the laryngeal in *\*-γʷája-* sharing a similar fate to other laryngeals in Avestan and Old Persian, i.e., loss and subsequent crasis. It is difficult to know for sure.

The sound changes we propose for Proto-Indo-Iranian, Proto-Indic, and Proto-Iranian, while phonetically natural, lack parallels outside of the causative formations in question.<sup>49</sup> One may attribute this to the fact that the sequence *\*-oh<sub>x</sub>é-* is relatively uncommon outside of the causative formation itself. For instance, in *o*-grade thematic nouns formed to laryngeal-final roots, the only case marker that begins in *-e-*, the vocative singular, would be either unaccented or barytone (i.e., *\*Coh<sub>x</sub>e* or *\*Coh<sub>x</sub>é*, not *Xcoh<sub>x</sub>e*); in any case we were unable to find any nouns of this shape, as *\*CVH-* thematic nominals (always?) occur in formations with accented *\*-ó-* and zero-grade root (*gʷih₃-ó-*, *\*dʰh₁-ó-*, *\*sth₂-ó-* etc.).<sup>50</sup> Elsewhere in the verbal system, the 3sg. of the perfect

<sup>48</sup> Indeed, one is tempted to view the Skt. infinitive *dāváne* ‘to give’ as continuing the earlier sequence *dāyʷánai* (\**doh₃énai*; cf. Av. *dāvōi* ‘to give’), with the same change in consonantism as seen in the Iranian LECS, though see Cowgill (1964: 354–356) for extensive discussion of related forms in other languages such as Gk. *δοῦναι* ‘to give’. There are only two other infinitives of the shape *-váne* in the RV (Macdonell 1916: 193), *turváne* ‘to overcome’ (*√tr*) and *dhúrváne* ‘to injure’ (*√dhvr*), both of which curiously contain *u*-vocalism in the root of their infinitival forms.

<sup>49</sup> Aside from Skt. *dāváne* (see fn. 48).

<sup>50</sup> See Wodtko et al. (2008) for further examples.

contains the sequence  $^*oh_xe$ -, but the  $^*e$ - is always unaccented ( $^*d^hed^hoh_je$  or  $^*d^hed^hóh_je$  'put', never  ${}^x d^hed^hoh_je$ ).<sup>51</sup>

One may wonder why the sound changes assumed here should be restricted to the specific sequence  $^*oHé$ - and not be present in similar sequences such as  $^*\bar{o}He$ -,  $^*oHó$ -, and  $^*\bar{o}Ho$ -.<sup>52</sup> The clearest reflexes of  $^*\bar{o}He$ - in IIr. are in word-final position, where  $^*\bar{o}He$ - > -au, as seen in the 3sg. perfect (e.g. Skt. *dadáu* <  $^*dedóh_3e$ ) and the nominative/accusative/vocative dual (Skt. *dváu* 'two' <  $^*duóh_je$ ), a change which resembles the one posited here, though again it cannot be related (see fn. 41 above). Like  $^*oHé$ -, the sequence  $^*oHó$ - was not especially common, given that in oxytone thematic formations to  $^*CeH$ - roots the root was in the zero grade. As for  $^*\bar{o}Ho$ -, this sequence is typically reconstructed in the genitive plural  $^*\bar{o}Hom$ , for which there is clear evidence in IIr. poetry of disyllabic scensions (Fortson 2010: 129; Kümmel 2013). Other forms, like Ved. *dás*, GAv. *dāh*- (<  $^*dóh_3os$ ; Wodtko et al. 2008: 62) also argue against the change of  $^*\bar{o}Ho$ - >  $^*\bar{o}H^wo$ -.

As of this moment, we do not have any obvious phonetic explanation as to why the proposed change was restricted to  $^*oHé$ - and did not take place in  $^*\bar{o}He$ - or  $^*\bar{o}Ho$ -. For this reason, perhaps it would be more prudent to posit an instance of irregular sound change, where the laryngeal was maintained solely in the causative suffix to avoid hiatus. A retention such as this would not be unprecedented. Like the survival of intervocalic -s- in the Greek aorist (Pre-PGk. *élusa* > Gk.  $\ddot{\epsilon}\lambda\bar{w}\sigma\alpha$  vs. expected  ${}^x \ddot{\epsilon}\lambda\bar{w}\alpha$  [cf. Hill 2014: 214]) or the past tense -ed [t] in English *passed*, despite its common deletion in words like *past* (Labov et al. 1968, Labov 1984, Labov 1989, Phillips 2006: 2), retention of root-final laryngeals in PIIr. causative formations could be seen as an instance of sound change blocked by the grammar.

Indeed, for a reductive change such as this one, the irregular, morphologically motivated retention of a laryngeal would be expected only if the causatives themselves were relatively infrequent, as reductive sound changes happen first in more frequent words (Phillips 2006: 58–76, Bybee 2011: 67); thus uncommon words or sequences are less likely to undergo sound deletion. While there of course exists no corpus of Proto-Indo-Iranian (or PIE) which we can consult to validate this assertion, there is reason to think that causatives were relatively infrequent in both Proto-Indo-Iranian and PIE itself. As Haspelmath et al. (2014: 614) write, "When a long form becomes frequent, it tends to be shortened in language change, and when a short form becomes rare, it

<sup>51</sup> See Fortson (2010: 104).

<sup>52</sup> We thank Reviewer #2 for bringing this matter to our attention.

tends to become longer." In the case of the causative suffix,  $*-éje/o-$  is in fact the **only** disyllabic derivational affix reconstructable for the PIE verbal system: cf. the monosyllabic present suffix  $*-u-$ , reduplication, the nasal infixes  $*-n(\acute{e})-$ ,  $*-n(\acute{e})\mu-$ ,  $*-n(\acute{e})h_x-$ , thematic  $*-\acute{e}/\acute{o}-$ , inchoative  $*-ské/\acute{o}-$ ,  $*-jé/\acute{o}-$ ,  $*-dé/\acute{o}-$ ,  $*-d^h\acute{e}/\acute{o}-$ ,  $*-té/\acute{o}-$ , the sigmatic aorist suffix  $*-s-$ , etc. (LIV 14–25). The very fact that the causative formant was longer than the others suggests that it was a rarer morpheme, compounded by the fact that PIE  $*-éje/o-$  was likely only added to unaccusative roots (Bozzone 2020), which include canonical S<sub>O</sub> verbs ('lie'), psych verbs ('fear'), motion verbs ('go'), verbs of emission ('fart'), and verbs of ingestion ('eat'). These verbs, for the most part, are inherently non-causal, and as Haspelmath et al. (2014) argue, the causatives of non-causal verbs are much less common than their base forms, again indicating that these formations would have been relatively rare compared to their base.<sup>53</sup>

In short, whether the sound change  $*-oHé- > *-oHwé-$  was entirely regular or was lexically restricted, the lack of parallel examples elsewhere in Indo-Iranian is not surprising and does not constitute a serious counterargument to the developments proposed here.

## 7 Final thoughts

In the many years that the authors have worked together on this topic, we have had a number of discussions with scholars who hesitate to accept a sound change such as the one we propose above, as it has no clear parallels. The proposed change, they argue, is ad hoc, and therefore our reconstruction of a PIIr. LECS is a non-starter. But it must also be noted that all of the other previous explanations of the Indic  $-(\bar{a})páya-$  causative are similarly ad hoc, rooted for the most part in either poorly understood enlargements or determinatives, or unfalsifiable (and often questionable) instances of analogical spread. These explanations also lack the many benefits of our proposal, as they are unable to account for (and do not acknowledge) the multiple LECS variants across Indo-Iranian; they do not explain why the LECS is found only in Indo-Iranian among the Indo-European branches; and they fail to motivate the restriction of the LECS to laryngeal-final roots.

<sup>53</sup> Of course, in attested Indo-Iranian causatives were productively formed to more than just unaccusatives, but at least in the case of Sanskrit, this extension to non-unaccusatives appears to have taken place "gradually and within the history of Vedic" (Bozzone 2020: 32, following Kulikov 2013). Thus, the proposed sound change in question would have taken place while the causative suffix was still primarily restricted to unaccusatives.

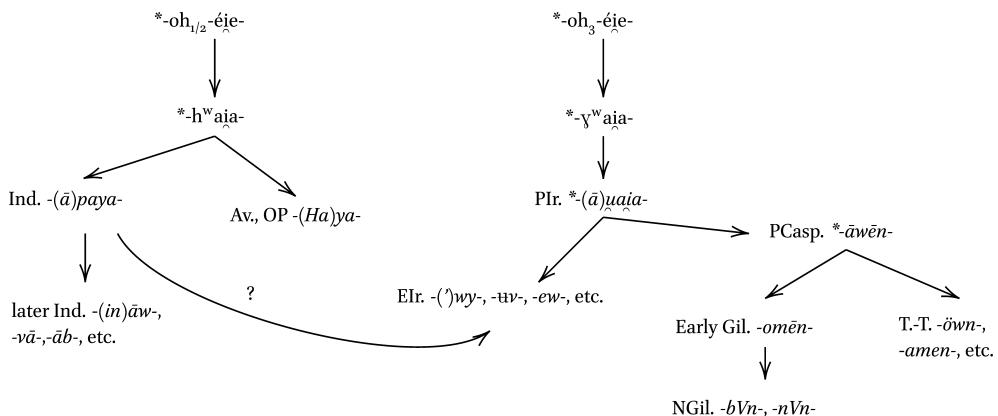


FIGURE 8 Evolution of the labial-expanded causative suffix in Indo-Iranian

But regardless of what one believes about the broader connections of Indic  $-(ā)\text{paya-}$  and Iranian  $*\text{-(ā)uaia-}$ , the evidence put forth in this paper makes it increasingly clear that, as pointed out by Schwartz (1969: 447, following Emerick 1968: 187), it is untenable to maintain that the Eastern Iranian  $*\text{-(ā)uaia-}$  derives from Indic  $-(ā)\text{paya-}$ ; rather, it represents a “genuine Iranian feature.” Given the strong evidence from the Caspian languages presented above, in Old Gilaki  $-\text{omen-}/$ , Gilaki  $\text{-bVn-}$  and  $\text{-nVn-}$ , Northern Talyshi, Northern Tati  $-\text{ovn-}$ ,  $-\text{own-}$ ,  $-\text{ōwn-}$  and Southern Talyshi, Central Tati  $-\text{amen-}$ ,  $-\text{amən-}$ ,  $-\text{amən-}$ , it appears likely that  $*\text{āuaia-}$  was continued in certain Western Iranian languages as well, reconstructible as the intermediate Proto-Caspian suffix  $*\text{-āwēn-}$ . We believe that all of these pieces of evidence are entangled—without one, it would be difficult to arrive at the correct conclusion. If not for the Modern Gilaki LECS, which has endured in a strictly demarcated area, the proposed connection between the Early Modern Gilaki LECS  $-\text{omen-}/$  and the Tati-Talyshi LECS would be much less clear. Indeed, reconstructing a PIIr.  $*\text{āuaia-}$  would be difficult on the basis of only an Eastern Iranian LECS, which is why the identification of its Northwestern Iranian cousin  $*\text{-āwēn-}$ , found in a handful of understudied languages spoken in northern Iran, is so important. What other archaisms may be found in these languages and in their diverse varieties? Or in the dozens of other understudied Modern Iranian languages? Partnerships and collaborations between native speakers of understudied languages and Indo-Europeanists (such as this one) might expand our field in new and exciting ways.

## References

- Abū l-Faḍl b. Šahrīdawīr al-Daylāmī al-Ǧīlānī. 2008 [c. 13th c.]. *Tafsīr-i daylāmī* [Daylāmi commentary]. Tehran: Markaz pāzūhish, kitābkhāneh, müzih va markaz asnād-i majlis-i shorā-i islami.
- Agha, Asif. 2007. *Language and social relations*. Cambridge: Cambridge University Press.
- Austin, Peter. 2008. *One thousand languages. Living, endangered, and lost*. Berkeley: University of California Press.
- Bailey, Harold W. 1968. *Saka documents. Text volume*. London: Humphries & Co. on behalf of Corpus Inscriptionum Iranicarum.
- Bailey, Harold W. 1979. *Dictionary of Khotan-Saka*. Cambridge: Cambridge University Press.
- Bamshadi Parsa, & MahinNaz Mirdehghan. 2014. *Radeh shenasi-i barhamnehesht va jahat-i fe'l dar kurdi gurani* [The typology of verb superposition and direction in Gurani Kurdish]. *Pazhouheshha-i zabani* 6(2): 1–20.
- Bashir, Elena. 2009. Wakhi. In Windfuhr (ed.) 2009b, 825–858.
- Beddoe, Pamela. 2009. A coarticulatory path to sound change. *Language* 85(4): 785–821.
- Beekes, Robert S.P. 1981. Intervocalic laryngeal in Gatha-Avestan. In *Bono homini donum. Essays in historical linguistics in memory of J. Alexander Kerns*, ed. Yoël L. Arbeitman & Allan R. Bomhard, 47–64. Amsterdam: John Benjamins.
- Beekes, Robert S.P. 1988. *A grammar of Gatha-Avestan*. Leiden: Brill.
- Beekes, Robert S.P. 2010. *Etymological dictionary of Greek*, vol. 1. Leiden: Brill.
- Borjian, Habib. 2009. *Motūn-i Tabari* [Tabari texts]. Tehran: Written Heritage Publishers.
- Bozzone, Chiara. 2020. Reconstructing the PIE causative in a cross-linguistic perspective. *Indo-European Linguistics* 8: 1–45.
- Brugmann, Karl. 1895. *A comparative grammar of the Indo-Germanic languages. Morphology. Part III: Verbs. Formation of the stem, and inflexion or conjugation*. New-York: B. Westermann & Co.
- Bybee, Joan. 2011. Lexical diffusion in regular sound change. In *Sounds and systems. Studies in structure and change. A festschrift for Theo Vennemann*, ed. David Restle & Dietmar Zaefferer, 59–74. Berlin & New York: De Gruyter Mouton.
- Campbell, Lyle. 2021. *Historical linguistics. An introduction*. Cambridge, MA: The MIT Press.
- Cantera, Alberto. 2017. The phonology of Iranian. In *Handbook of comparative and historical Indo-European linguistics*, vol. 1, ed. Jared Klein, Matthias Fritz, & Brian Joseph, 481–503. Berlin & Boston: De Gruyter Mouton.
- Charpentier, Jarl. 1912. Beiträge zur alt- und mittelindischen Wortkunde. *Indogermanische Forschungen* 29(1): 367–403.

- Cathcart, Chandra. 2015. *Iranian dialectology and dialectometry*. Ph.D. dissertation, University of California.
- Cheung, Johnny. 2007. *Etymological dictionary of the Iranian verb*. Leiden: Brill.
- Clauson, Gerard. 1972. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon.
- Cowgill, Warren. 1964. The supposed Cypriote optatives *duwánoi* and *dókoi*, with notes on the Greek infinitive formations. *Language* 40(3): 344–365.
- Deihim, Giti. 2007. *Barresi-yi khurdi gūyish-hā-yi mantiq-i Qasran [An account of the sub-dialects of Qasran region]*. Tehran: Farhangistān-i Zabān va Adab Farsi.
- Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Heidelberg: Winter.
- Durkin-Meistererst, Desmund. 2009. Khwarazmian. In Windfuhr (ed.) 2009b, 336–376.
- Edgerton, Franklin. 1946. Indic causatives in *-āpayati* (-āpeti, -āvei). *Language* 22(2): 94–101.
- Emmerick, Ronald E. 1968. *Saka grammatical studies*. Oxford: London.
- Emmerick, Ronald E. 2009. Khotanese and Tumshuqese. In Windfuhr (ed.) 2009b, 377–415.
- EWAIa = Mayrhofer, Manfred. 2001. *Etymologisches Wörterbuch des Altindoarischen*, vol. 3. Heidelberg: Winter.
- Fortson, Benjamin IV. 2010. *Indo-European language and culture. An introduction*<sup>2</sup>. Malden, MA: Wiley-Blackwell.
- Gaedike, Charles. 1880. *Der Accusativ im Veda*. Breslau: Koebner.
- Ghosh, Batakrishna. 1982. *Nominal and verbal formations in -p- in Sanskrit*. Tr. from the French by Dilip Kumar Biswas. Calcutta: Sanskrit Pustak Bhandar.
- Gunkel, Dieter. 2022. The first person singular of the athematic middle optative in Vedic and Indo-Iranian. *Journal of the American Oriental Society* 142(2): 251–266.
- Gray, Louis H. 1902. *Indo-Iranian phonology, with special reference to the Middle and New Indo-Iranian languages*. New York: Columbia University Press.
- Haspelmath, Martin, Andreea Calude, Michael Spagnal, Heiko Narrog, & Elif Bamyaci. 2014. Coding causal—non-causal verb alternations: A form—frequency correspondence explanation. *Journal of linguistics* 50(3): 587–625, 729.
- Henning, Walter B. 1934. Das Verbum des Mittelpersischen der Turfanfragmente. *Zeitschrift für Indologie und Iranistik* 9: 158–253.
- Hill, Nathan W. 2014. Grammatically conditioned sound change. *Language and linguistics compass* 8(6): 211–229.
- Hirt, Herman A. 1921. *Etymologie der neu hochdeutschen Sprache*. München: Vero.
- Horn, Paul. 1893. *Grundriss der neupersischen Etymologie*. Straßburg: Trübner.
- IEW = Pokorny, Julius. 2005. *Indogermanisches etymologisches Wörterbuch*<sup>5</sup>. Tübingen & Basel: Francke.
- Insler, Stanley. 1987. The Vedic causative type *jāpáyati*. In *Studies in memory of Warren Cowgill (1929–1985)*, ed. Calvert Watkins, 54–65. Berlin & Boston: De Gruyter.

- Jamison, Stephanie W. 1983. *Function and form in the -áya- formations of the Rig Veda and Atharva Veda*. Göttingen: Vandenhoeck & Ruprecht.
- Jahangiri, Nader. 2015. *Zabān-i gīlakī (gūyish-i Lāhijān): Vāzhabh'nāmah va pārahāt vīzhagī'hā-yi āvāyī va sākhvāzhabh'hā* [Gilaki Language (Lahijani dialect): dictionary and some phonological and morphoicla features]. Rasht: Farhang-e Ilya.
- Jahani, Carina, & Agnes Korn. 2009. Balochi. In Windfuhr (ed.) 2009b, 634–692.
- Jasanoff, Jay. 1988. The sigmatic aorist in Tocharian and Indo-European. *Tocharian and Indo-European studies* 2: 52–76.
- Kellens, Jean. 1990. *Les textes vieil-avestiques*. Vol. II: *Répertoires grammaticaux et lexique*. Wiesbaden: Reichert.
- Kent, Roland G. 1950. *Old Persian. Grammar, texts, lexicon*. New Haven: American Oriental Society.
- Khoshirsat, Zia. 2018. *The origin of the Gilaki causative -be:-*. M.A. thesis. University of Kentucky.
- Kia, Sadegh. 1948. *Vāzhehnāmeh-yi Tabarī* [Tabri disctionary]. Tehran: University of Tehran Press.
- Kieffer, Charles M. 2009. Parachi. In Windfuhr (ed.) 2009b, 693–720.
- Kobayashi, Masato. 2004. *Historical phonology of Old Indo-Aryan consonants*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Konow, Sten. 1932. *Saka studies*. Norway: Oslo Etnografiske Museum.
- Korn, Agnes. 2013. Looking for the middle way: Voice and transitivity in complex predicates in Iranian. *Lingua* 135: 30–55.
- Kulikov, Leonid. 2013. Constraints on the causative derivation in Early Vedic. Evidence for a diachronic typology of transitivity. *Poznań studies in contemporary linguistics* 9(1): 79–102.
- Kümmel, Martin. 2013. Zur Endung des Genitivs Plural im Indoiranischen. *Indogermanische Forschungen* 118: 193–212.
- Kümmel, Martin. 2014. The development of laryngeals in Indo-Iranian. Paper read at the conference “The Sound of Indo-European 3,” Opava.
- Kümmel, Martin. 2018. The survival of laryngeals in Iranian. In *Farnah. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*, ed. Lucien Beek, Alwin Kloekhorst, Guus Kroonen, Michaël Peyrot, & Tijmen Pronk, 162–172. Ann Arbor & New York: Beech Stave.
- Labov, William. 1984. The interpretation of zeroes. In Wolfgang U. Dressler et al. (eds.), *Phonologica 1984. Proceedings of the Fifth International Phonology Meeting, Eisenstadt*, 135–156. Cambridge: Cambridge University Press.
- Labov, William. 1989. The child as linguistic historian. *Language variation and change* 1: 85–97.
- Labov, William, Paul Cohen, Clarence Robins, & John Lewis. 1968. *A study of the non-standard English of Negro and Puerto Rican speakers in New York City* (Cooperative Research Report 3288). 2 vols. Philadelphia: U.S. Regional Survey.

- Leumann, Manu. 1940. Zur Stammbildung der Verben im Indischen. *Indogermanische Forschungen* 57: 205–238.
- LIV = Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, & Birgitte Schirmer. 2001. *LIV: Lexikon der indogermanischen Verben*<sup>2</sup>. Wiesbaden: Reichert.
- Lubotsky, Alexander. 2018. The phonology of Proto-Indo-Iranian. In *Handbook of comparative and historical Indo-European linguistics*, vol. 3, ed. Jared Klein, Matthias Fritz, & Brian Joseph, 1875–1888. Berlin & Boston: De Gruyter Mouton.
- Macdonell, A.A. 1916. *A Vedic grammar for students*. Oxford: Clarendon.
- Mallory, J.P., & Douglas Q. Adams. 2006. *The Oxford introduction to Proto-Indo-European and the Proto-Indo-European world*. Oxford: Oxford University Press.
- Maqāmāt-i Ḥarīrī. Malek Museum, MS 2487. Tehran.
- Masica, Colin. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- McCarus, Ernest N. 2009. Kurdish. In Windfuhr (ed.) 2009b, 587–633.
- MacKenzie, David N. 1971. *A concise Pahlavi dictionary*. London: Oxford University Press.
- Miller, Robert McColl, & Trask, Larry. 2015. *Trask's Historical linguistics*<sup>3</sup>. Oxford: Routledge.
- Mohammadi, Meysam. 2012. Barkhī barrisīha dar murid-i tafsī-i daylamī [Some analyses of Daylami Commentary]. *Nāmih-yi Bahāristān*, Duri-yi jadid [N.S.] 2: 234–257.
- Mohammadi, Meysam. 2015. Barrisī va naqd-i maqālī-yi “Tarjomih-yi Zaidī-i Qurān bi Zabān-i Tabarī” [Analyses and Critique of “Zaydi translation of Quran into Tabari language”]. *Nāmih-yi Bahāristān*, Duri-yi jadid [N.S.] 4: 146–191.
- Morgenstierne, Georg. 1929. The language of the Ashkun Kafirs. *Norsk tidsskrift for sprogvitenskap* 2: 192–289.
- Morgenstierne, Georg. 1938. *Indo-Iranian frontier languages*. Vol. II: *Iranian Pamir languages, Yidgha-Munji, Sanglechi-Ishkashmi and Wakhi*. Oslo: H. Aschehoug & Co. (W. Nygaard).
- Paul, Ludwig. 2009. Kurdish. In Windfuhr (ed.) 2009b, 545–576.
- Persson, Per. 1912. *Beiträge zur indogermanischen Wortforschung*. 2 vols. Uppsala: Almqvist & Wiksell.
- Phillips, Betty. 2006. *Word frequency and lexical diffusion*. London: Palgrave Macmillan.
- Rasekh-Mohannad, Mohammad, & Raheleh Izadifar. 2013. Barres-i sakht-i sababi dar zabān-i Tati [The study of casuative construction in Tati language]. *Adab Pazhūh* 5: 135–160.
- Rastorgueva, Vera, A. Kerimova, A. Mamedzade, L. Pireiko, D. Edel'man, & R.M. Lockwood. 2012. *The Gilaki language*. Uppsala: Acta Universitatis Upsaliensis.
- Sabzalipour, Jahangir. 2015. *Vizhigīhā-yi masdar dar zabānha-yi hashīyih-yi khazar (tati, taleshi, gilaki)* [The characteristics of the infinitive in languages of the Caspian littoral]. *Zabānhā va gūyeshhā-yi irānī*, 103–129.

- Salemann, Carl H. 1901. Mittelpersisch. In *Grundriss der iranischen Philologie* 1/1, ed. Wilhelm Geiger & Ernst Kuhn, 249–332. Straßburg: Trübner.
- Samadi, Mahlagha. 1986. *Das chwaresmische Verbum*. Wiesbaden: Harrassowitz.
- Samareh, Yadollah. 1988. *Tahlīl-i sākhtāri-yi fe'l dar gūyish-i gilaki-i kelārdasht* [The analysis of verb structure in Gilaki of Kelardasht]. *Dānishkadih adabiyāt va olūm ensāni dānishgāh Tehran* 20(1–4): 169–187.
- Schleicher, August, & Herbert Bendall. 1874. *A compendium of the comparative grammar of the Indo-European, Sanskrit, Greek and Latin languages*. London: Trübner & Co.
- Schwartz, Martin. 1969. Review of H.W. Bailey, *Indo-Scythian studies, being Khotanese texts*. Vol. VI: *Prolexis to the Book of Zambasta*. *Journal of the American Oriental Society* 89(2): 444–447.
- Skjærvø, P. Oktor. 2009. Middle West Iranian. In Windfuhr (ed.) 2009b, 196–278.
- Skjærvø, P. Oktor. 2017. The documentation of Iranian. In *Handbook of comparative and historical Indo-European linguistics*, vol. 1, ed. Jared Klein, Matthias Fritz, & Brian Joseph, 471–481. Berlin & Boston: De Gruyter & Mouton.
- Stilo, Donald L. 1981. The Tati language group in the sociolinguistic context of North-western Iran and Transcaucasia. *Iranian studies* 14(3–4): 137–187.
- Stilo, Donald L. 2001. Gilaki. In *Encyclopedia Iranica*, Vol. x, ed. Ehsan Yarshater, 660–668. New York: Bibliotheca Persica.
- Stilo, Donald L. 2015. The polygenetic origins of the Northern Talyshi language. In *Studies on Iran and the Caucasus*, ed. Uwe Bläsing, Victoria Arakelova, & Matthias Weinreich, 411–454. Leiden: Brill.
- Stilo, Donald L. 2018. The Caspian region and south Azerbaijan. Caspian and Tatic. In *The languages and linguistics of Western Asia*, ed. Geoffry Haig & Geoffry Khan, 659–824. Berlin & Boston: De Gruyter Mouton.
- Sütterlin, L. 1906. Die Denominativa im Altindischen. *Indogermanische Forschungen* 19: 480–577.
- Tedesco, Paul. 1923. “Geben” und “nehmen” im Indischen. *Journal of the American Oriental Society* 43: 358–390.
- Vendryès, Joseph. 1931. Procès-verbaux des séances. *Bulletin de la Société de Linguistique de Paris* 31: i–xxv.
- Voegelin, Charles F., & Florence M.R. Voegelin. 1987. *Classification and index of the world's languages*. New York: Elsevier.
- Weiss, Michael. 2020. *Outline of the historical and comparative grammar of Latin*<sup>2</sup>. Ann Arbor: Beech Stave.
- Windfuhr, Gernot. 2009a. Dialectology and topics. In Windfuhr (ed.) 2009b, 5–42.
- Windfuhr, Gernot (ed.). 2009b. *The Iranian languages*. London & New York: Routledge.
- Windfuhr, Gernot, & John R. Perry. 2009. Persian and Tajik. In Windfuhr (ed.) 2009b, 416–544.

- Whitney, William Dwight. 1885. *The roots, verb-forms, and primary derivatives of the Sanskrit language*. Leipzig: Breitkopf & Härtel.
- Wodtko, Dagmar S., Britta Irslinger, & Carolin Schneider. 2008. *Nomina im indogermanischen Lexikon*. Heidelberg: Winter.
- Yoshida, Yutaka. 2009. Sogdian. In Windfuhr (ed.) 2009b, 279–335.
- Zarshenas, Zohreh, & Nima Asefi. 2019. *Sākhtār-i afā'l-i sababi-yi talyshi az chish-mandāz-i zabānshenāsi-i tārikhi* [The structure of Talyshi causative verbs from the historical point of view]. *Zaban shenakht. Pazhūheshgāh olūm ensāni va motāle'at farhangī* 9(2): 1–14.
- Zolfagharī, Sima. 2017. *The Bakhtiari. An anthropological-linguistic lexical study of Haft Lang nomads of southwestern Iran*. Ph.D. dissertation, Leiden University.