Dialectal layers in West Iranian: a Hierarchical Dirichlet Process Approach to Linguistic Relationships

Chundra A. Cathcart

Department of Comparative Language Science/
Center for the Interdisciplinary Study of Language Evolution
University of Zurich

December 2019

Abstract

This paper addresses a series of complex and unresolved issues in the historical phonology of West Iranian languages, (Persian, Kurdish, Balochi, and other languages), which display a high degree of non-*Lautgesetzlich* behavior. Most of this irregularity is undoubtedly due to language contact; we argue, however, that an oversimplified view of the processes at work has prevailed in the literature on West Iranian dialectology, with specialists assuming that deviations from an expected outcome in a given non-Persian language are due to lexical borrowing from some chronological stage of Persian. It is demonstrated that this qualitative approach yields at times problematic conclusions stemming from the lack of explicit probabilistic inferences regarding the distribution of the data: Persian may not be the sole donor language; additionally, borrowing at the lexical level is not always the mechanism that introduces irregularity. In many cases, the possibility that West Iranian languages show different reflexes in different conditioning environments remains under-explored. We employ a novel Bayesian approach designed to overcome these problems and tease apart the different determinants of irregularity in patterns of West Iranian sound change. Our methodology allows us to provisionally resolve a number of outstanding questions in the literature on West Iranian dialectology concerning the dialectal affiliation of certain sound changes. We outline future directions for work of this sort.

1 Introduction

Isoglosses based on sound changes differentiating the West Iranian languages, a group comprising Persian, Kurdish, Balochi, and other speech varieties, have long been of interest to linguists. The West Iranian languages are traditionally divided into Northwest (containing Kurdish, Balochi, etc.) and Southwest (containing Persian and closely related dialects) subgroups, the latter of which can be defined by a small number of phonological and morphological innovations that have taken place before the attestation of Old Persian, its oldest member. At the same time, a comparable (if not larger) number of Persian innovations have taken place after Old and Middle Persian, and similar innovations can be seen in other West Iranian languages, showing the effect of complex areal networks that have existed during the development of these languages.

A number of reflexes can be identified as Southwest Iranian or Northwest Iranian on the basis of the languages in which they occur; however, language contact has complicated the picture significantly. In some cases, it is not clear what the "correct" outcome should be; for instance, in the word for 'spleen', Kurdish shows what is thought to be a typically SWIr outcome, while Persian shows a typically NWIr outcome.

Researchers have a set of diagnostics for marking individual words as loans in specific languages, but some of these diagnostics are better founded than others; in general, the picture is often so noisy, and these heuristics are so tightly intertwined, that all the facts cannot be qualitatively resolved within a traditional comparative-historical framework. We propose an alternative way of analyzing West Iranian data that integrates insights from the comparative method with probabilistic modeling. While previous research has tended to make hard decisions regarding a language's regular reflexes of sound change, this study avoids this approach; instead, we employ a quantitative approach intended to let regular behavior fall out of the data.

This paper investigates this variation in historical phonology across etymological reflexes and languages on a large scale. Specifically, we use a Bayesian probabilistic model to reduce the dimensionality of the data seen within and across languages into a set of latent, unobserved components representing dialect membership which can be shared by multiple languages. Our model is non-parametric, meaning that there is no upper bound on the number of latent features inferred. Both languages and phonological variants are associated with the presence of a latent feature. This allows us to identify potential networks of language contact across our dataset.

This methodology sheds light on a number of unresolved issues in the literature on West Iranian dialectology. We find, unsurprisingly, that West Iranian languages show admixture (to differing degrees) from two major dialect components, roughly corresponding to Northwest and Southwest Iranian dialect groups. We provisionally resolve a small number of questions regarding the dialectal provenance of certain types of sound change; while the impact of our results is potentially limited due to the relatively small size of our data, our results are interpretable, and our methodology is promising. We discuss future directions for models of this sort.

2 The West Iranian languages

The Iranian languages have traditionally been divided into East and West subgroups, but the genetic status of these labels is shaky. Historically, Bartholomae (1883:1) divided Old Iranian western and eastern variants, the former represented by Old Persian, and the latter by Avestan. The *Grundriss der iranischen Philologie*, particularly Wilhelm Geiger's contribution, provides a great deal of information on dialectology and subgrouping of contemporary Iranian languages. His chief distinction that cuts across Iranian is between "Persian" and "Non-Persian" dialects (Geiger 1901:414). East and West are used as purely geographic labels: at one point, Balochi is referred to as East Iranian (p. 414). There is not full agreement regarding which languages are western and which are eastern; the languages Ormuri and Parachi are considered to be western Iranian languages by some scholars (Grierson 1918, Oranskij 1977, Efimov 1986), but the consensus following Morgenstierne (1929) places them in East Iranian. The problematic nature of the geographic labels was noted by Bailey (1933). Sims-Williams (1996:651) states that East Iranian is not a genetic grouping, but a *Sprachbund*; most of its shared characteristics are retentions, rather than innovations, and the innovations that it shares are relatively trivial. Wendtland (2009) finds that there are no shared phonological or morphological characteristics between the East Iranian languages, and argues against Northeast and Southeast

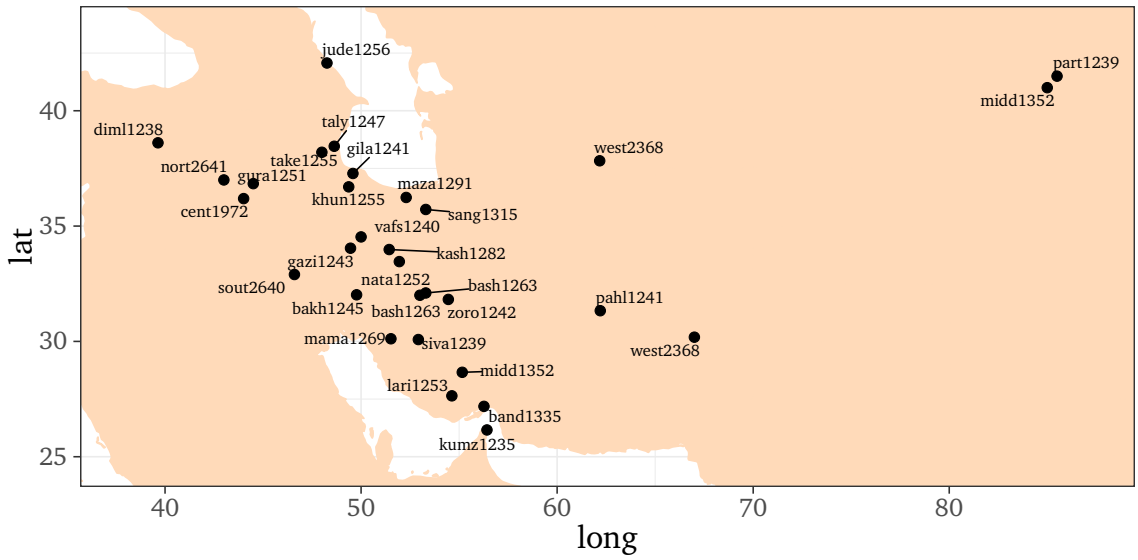

Figure 1: Approximate locations of languages in sample

subgroups (a division provisionally suggested in Morgenstierne 1926 and followed in Oranskij 1977, Kieffer 1989 and elsewhere). Cathcart (2015), Korn (2019) argue there are virtually no non-trivial innovations shared among West Iranian languages that could serve as diagnostics for subgrouping.

Regardless the genetic status of West Iranian, the label is meaningful, not only from a typological standpoint (West Iranian languages are highly convergent in their morphosyntax) but in terms of many of the diachronic trends displayed by West Iranian languages. Contact with non-Indo-European linguistic stocks, such as Turkic and Semitic, may have aided in shaping the linguistic profiles of West Iranian languages (Stilo 2005, 2018). Even if there are no good shared genetic innovations among West Iranian languages, the study of inter-dialectal West Iranian contact has the potential to shed light on the socio-historical development of Iran and surrounding regions.

The West Iranian languages are traditionally divided into Northwest and Southwest groups. The Southwest group, comprising Old, Middle and New Persian, as well as closely related dialects such as Bashkardi, Kumzari, Judeo-Tati, and others, is generally viewed as a genetic subgroup, defined by a small number of innovations. The Northwest group has fewer subgroup-defining innovations uniting it. The finer details of this distinction are not of particular importance to this paper, as our intention is for dialectal groupings to fall out of the behavior displayed by languages in our sample, which are listed in Table 1.

3 West Iranian historical phonological variation

West Iranian languages show a great deal of deviation from expected outcomes of historical phonology. This is clear in the oldest language; Old Persian contains a number of words which display the reflexes *s*, *d*, and *sp* and *zb* for Proto-Iranian **ć*, **ǰ*, **ć̲u*, and **ǰ̲u* instead of the expected outcomes *θ*, *d*, *s* and

Language	Alternative names/subdialects	Sources	Glottocode
Natanzi		Žukovskij 1922	nata1252
Zazaki	Dimli	Paul 1998b	diml1238
Pahlavi	*Book Pahlavi*, *Psalter Pahlavi*	MacKenzie 1971	pahl1241
Kashani		Žukovskij 1922	kash1282
Kurmanji Kurdish		Soane 1913, Thackston nd	nort2641
Bakhtiyari		Anonby and Asadi 2014	bakh1245
Mazandarani	Tabari	Nawata 1984	maza1291
Vanishani		Žukovskij 1922	khun1255
Feyli		Mann 1909	sout2640
Qohrudi	Soi	Žukovskij 1922, Mann and Hadank 1932	khun1255
Parthian		Durkin-Meisterernst 2004	part1239
Sivandi		Lecoq 1979	siva1239
Zoroastrian Dari	Yazdi, Gabri (pejorative)	Ivanow 1940, Vahman and Asatrian 2002	zoro1242
Gazi		Eilers 1978	gazi1243
S. Tati	*Challi, Eshtehardi, Takestani*	Yar-shater 1969	take1255
Gurani	*Kandulai*	Benedictsen and Christensen 1921, Hadank 1930	gura1251
N. Bashkardi		Skjærvø 1988	bash1263
Khunsari		Eilers 1976	khun1255
Rakhshani Balochi		Barker 1969	west2368
Vafsi		Stilo 2004	vafs1240
Talysh		Schulze 2000, Paul 2011	taly1247
New Persian		Steingass 1892	midd1352
Gilaki	*Rashti*	Rastorgueva et al. 2012	gila1241
Sorani Kurdish		Blau 1980	cent1972
Middle Persian (Manichean)		Durkin-Meisterernst 2004	midd1352
Zefrai		Žukovskij 1922	khun1255
Larestani		Kamioka and Yamada 1979	lari1253
Judeo-Tati		Miller 1892, Authier 2012	jude1256
Kumzari		Thomas 1930, van der Wal Anonby 2015	kumz1235
S. Bashkardi		Skjærvø 1988	bash1263
Sangesari		Azami and Windfuhr 1972	sang1315
Awromani	*Pawai*	Benedictsen and Christensen 1921	gura1251
Bandari		Pelevin 2010	band1335
Mamasani		Mann 1909	mama1269
Marw Balochi		Azami and Windfuhr 1972	west2368

Table 1: West Iranian languages in data set, along with alternative names and sub-variants (in italics), sources from which information was taken, and compatible glottocodes (Hammarström et al. 2017)

z. This has led scholars to draw a distinction between "proper Old Persian" and "Median" forms (cf. Hoffmann 1976:60 ff.), the latter label an allusion to the confederation which preceded the Achaemenid Empire (with which Old Persian is associated). Although we can reliably identify only a single form as explicitly Median (σπακα 'dog', recorded by Herodotus, which shows the sound change **ćų̂ > sp*), a number of Old Iranian onomastic items are generally assumed to be Median.

Words containing irregular historical phonological reflexes are common in Middle and New Persian as well, and are generally ascribed to contact with Northwest Iranian languages (although there are several probable loans from East Iranian as well). Northwest Iranian languages show the same degree of irregularity and contain a number of clear loans from various chronological stages of Persian, which is not surprising, given the sociopolitical influence of the Persian language in Iranian antiquity and onward.

It is likely that a number of mechanisms have worked together to create the complex patterns seen across West Iranian. These include (but are not limited to) language-internal factors such as the following:

- Poorly understood conditioning environments: we may not fully understand the factors influencing regular sound change within languages
- Analogical change, including paradigmatic leveling and extension, contamination, etc.

Additionally, inter-language factors like the following are almost certainly involved:

- Borrowing of lexical items
- Lexical diffusion of sound changes

As mentioned above, most explanations of irregularity appeal to lexical borrowing, often from an identifiable source such as Persian. However, it is additionally possible that more than one dialectal source of similar-looking reflexes was involved (e.g., the change **û̯- > b* may not have been restricted to Persian); furthermore, it is possible that under the umbrella of widespread multilingualism, speakers imposed sound changes from one dialect onto words from other speech varieties. In certain *Mischformen* it is quite clear that diffusion of sound changes was at work, though various cultural terms, secondary products, and technological innovations have been convincingly identified as loanwords. In many cases, however, it is not possible to distinguish between the two mechanisms. Additionally, it is not entirely clear whether similar-looking sound changes should be treated as unified, stemming from a single speech variant, or whether nearly identical sound changes were developed in parallel in different speech communities, possibly at different times. We identify some of these problems in the survey of sound change below, and propose a data-driven solution to some, but not all, of these issues.

4 West Iranian historical phonology

Below, we give a synopsis of historical phonological innovations in West Iranian languages (viewed through the lens of Persian, which has the best-documented historical record) and discuss outstanding problems. These developments are given in rough chronological order (where a chronology can be securely established), starting with innovations preceding Old Persian, and so on, focusing on some particularly vexing problems.

Dialectal differentiation is visible in the earliest attested West Iranian records, which consist of Achaemenid Old Persian inscriptions, as well as fragmentary Median records. At this stage, several phonological and morphological innovations that define Southwest Iranian as a subgroup can be identified (see below).

The *locus classicus* of West Iranian dialectal differentiation is Tedesco's (1921) study of Middle Persian/Parthian isoglosses in the Manichean texts of Turfan. Lentz (1926) discusses dialectal variation found in the Šāh-nāma. In many cases this variation can be periodized with respect to when variant forms were introduced, especially in the case of Persian (cf. Paul 2005). The isoglosses identified in these works have served as the basis for a large number of dialectological investigations. Over the past century, the list of variables has been supplemented (Bailey 1933, Krahnke 1976, Stilo 1981), and scholars have debated which features in particular are the most meaningful for West Iranian dialectology (Paul 1998a, Korn 2003, Windfuhr 2009), in terms of joint versus independent innovations.

4.1 Changes to PIr **ć, *ǰ*

The changes **ć* $>$ *θ* ($>$ Middle Persian, New Persian *h*), **ǰ* $>$ *d*, **ću̯* $>$ *s*, **ǰu̯* $>$ *z*, **θr* $>$ *ç*,[1] are found in a stratum of Old Persian (OP) vocabulary, and thought to be the expected outcome in Southwest Iranian languages. However, OP also exhibits a number of doublets or irregular reflexes of the aforementioned Proto-Iranian sounds, usually ascribed (as mentioned above) to Median admixture, though we know little about the true nature of the Median language, given the paucity of records.

This variation is well attested in Old Persian: PIr **ć* $>$ *θ* in one layer of vocabulary, but *s* elsewhere; PIr **ǰ* $>$ *d* in (likely) the same stratum, but *z* elsewhere. This variation is described further below.

4.1.1 PIr **ć-*

OP (or post-OP) initial *θ*- consistently corresponds to MP *s*-:[2]

OP *aθanga-* $>$ **θanga-* : MP *sang* 'stone'.

The fact that OP *θ* develops to MP *h* in most environments has led many scholars to assume that forms with MP *s*- are NW Iranian loans (cf. Gershevitch 1962a:2); however (Salemann 1901) takes initial MP *s*- to be the regular reflex of earlier *θ*-. The development of PIr **ć*- to *h*- is shown in a single form, *hadba* 'centipede' in Judeo-Shirazi, a dialect closely related to Persian but somewhat differentiated in terms of historical phonology (Morgenstierne 1932:55).

4.1.2 PIr **ću̯*

Reflexes of PIr **ću̯* are highly probative with respect to Iranian subgrouping; the "proper" Southwest Iranian outcome is taken to be *s*, while Khotanese Saka and Wakhi show *š*. Kurdish and Balochi, showing "transitional" behavior between Northwest and Southwest Iranian, appear to participate in the change **ću̯* $>$ *s* with Southwest Iranian, but not the changes **ć* $>$ *θ* $>$ *h*, **ǰ* $>$ *d*. In most Southwest Iranian dialects the change **ću̯* $>$ *s* must postdate the change **ć* $>$ *θ*. The rest of Iranian shows *sp*,

[1] This sound is usually interpreted as [s]; see Kümmel 2007 for an alternative view.

[2] The symbol $>$, used to represent descent between forms with regular sound change, is used somewhat liberally in this paper in that it may connect OP, MP, and NP forms that are not actually in a relationship of direct descent (given contact); $\rightarrow$ denotes an analogical change, possibly in combination with the operation of regular sound change.

or a sequence of sounds thought to descend from it, e.g., Ossetic *fs*; Khunsari, Gazi, Sangesari *asm* 'horse' $<$ **aspa-*. Zoroastrian Dari *sv* must be secondary rather than archaic, as PIr **sp* surfaces as *sv* as well.

The change **ćṷ* $>$ *sp* cannot be reconstructed to a hypothetical ancestor of the Central Iranian languages which share it (cf. Skjærvø 2009:50–51); without excluding Kurdish and Balochi from this group, but these languages cannot be placed in the Southwest Iranian group: in most Southwest Iranian dialects, the change **ćṷ* $>$ *s* must postdate the change **ć* (which probably had a phonetic value close to [s]) $> \theta$, as **ćṷ* became **θ* only in highly marginal dialects, e.g., Judeo-Shirazi *teš* 'louse', if from **θiša-* $<$ **ćṷiša-*. It is likely that changes to **ćṷ* represent a sort of areal diffusion among (originally) non-peripheral Iranian languages, albeit an old one which has operated prior to early Median and Scythian onomastic items. It is worth noting that similar fortition of OIA *śv* has taken place in the peripheral Indo-Aryan language Khowar as well as some Nuristani languages, though Morgenstierne (1926, 1932) cautions against connecting these developments with the Iranian one.

Persian shows reflexes containing *sp* at all chronological stages. New Persian also shows the cluster *sf*. Henning argues that this cluster cannot be secondary from earlier **sp*, and could instead be from a dialect in which PIr **ćṷ* "resulted directly in *sf*" (Henning 1963:71, fn. 13). Schwartz (2006:223) argues against the influence of Arabic (which resulted in certain sporadic *p* $>$ *f* changes, since Arabic lacks *p*) in certain words with *sf*. The circumstances under which NP *sf* came about remain unclear.

4.1.3 PIr **ćr*

A small number of Old and Middle Persian words show OP *ç*, MP *s* for PIr **ćr*, e.g., **ni-ćrai-* 'restore' $>$ OP *niyaçāray-* caus. (contaminated with *dāraya-*, according to Kent 1951:188), MP *nisāy-* 'conveying, dispatch'. Kent (1942:80) claims that OP **çunautiy* 'hear' 3sg ($<$ **ćrunauti*) yields NP *šunūdan*, but the latter form is better connected with **xšnau-* 'hear' (Cheung 2007:456). Elsewhere, Middle and New Persian show *s(V)r* and on occasion *š*:

PIr **ćrāṷaia-* 'hear' (caus.) $>$ MP *srāy* 'sing' $>$ NP *surū(dan)/sarāy* (Cheung 2007:357)
PIr **ćrauni-* $>$ NP *surūn* 'buttocks'
PIr **hṷaćrū-* $>$ NP *xusrū*, *xušū* 'mother-in-law'
PIr **aćru-(ka)-* 'teardrop' $>$ MMP *'rs* /ars/, *'sr*; NP *ars*, *ašk*

4.1.4 PIr **ćn*, **ǰn*

From the evidence available, initial **ćn-* $>$ *sn-* in most Iranian languages; medial **-ćn-* $>$ OP *-šn-*, e.g., **ṷać-na-* $>$ OP *vašna-* 'will, favor' (OAv *vasnā*, YAv *vasna* 'wish' inst.sg. may show the effects of analogy rather than a regular outcome; see Hoffmann and Forssman 2004:102).

PIr **ǰn* appears to have become OP *xšn-* ($>$ MP *šn-* $>$ NP *š(V)n-*) word-initially, and *-šn-* word-medially. From what we can tell, Northwest Iranian languages appear to have medial *-zn-* (on metathesis to *-nz-* in Median onomastic items, see Gershevitch 1962a), e.g., Parthian *gaznn* 'treasure' vs. NP *gašn* 'abundance' $<$ **gaǰna-*; Persian shows some forms with *-zn-*, possibly loans from Northwest or East Iranian. For initial *ǰn-*, however, these languages agree with Persian in reflecting **(x)šn-*, e.g., **ǰnā-sća-* 'know' $>$ Parthian *išnās*, Qohrudi *ešnās-* etc. (Cheung 2007:466). It is not clear how far this conditioned alternation can be projected into the proto-language; analogical change has blurred the picture.

4.1.5 PIr **ǰu̯*

The sequence **ǰu̯* is found in only a small number of Proto-Iranian etyma. Old Persian contains reflexes of only two of these forms, *patiyazbayam* 'proclaim' 1sg. impf. (< **pati-aǰu̯aiam* < **pati-a-ǰuH-aia-m*) and *hizāna-* 'tongue' (< **hiǰu̯āna-*). The latter form is believed by most scholars to be "proper" OP (Kent 1951), and the former a Median loan.[3] The Middle Persian word for tongue is *uzwān/izwān*, and cannot be taken as a direct reflex of the OP form (since *z* from other sources does not change to *zw*); Middle Persian however shows the development **ǰu̯* > *z* in *parzīr-* 'keep away' < **para-/pari-ǰu̯ar-* (Cheung 2007:475). New Persian and related dialects tend to show *zVb* or *zVw* (e.g., NP *zabān*, *zuwān* 'tongue' < **hiǰu̯āna-*).

4.1.6 PIr **ći̯*

The development of PIr **ći̯* in Persian is not entirely clear. Its fate is intertwined with that of the cluster **θi̯* (< PIIr **ti̯*), whose regular Old Persian reflex is thought to be *šiy* (Kent 1951:32); cf. **hu̯ai-paθi̯a-* > MP *xwēš* 'self'.

Old Persian attests this cluster only word-medially, where it surfaces as *θiy* (showing characteristic OP anaptyxis between consonants and glides), e.g., *viθiyā* 'house' loc.sg., possibly via paradigmatic leveling of the stem *viθ-* (< **u̯ić-*).

Old Persian does not directly attest this cluster word-initially; Middle Persian shows varying reflexes:

PIr **ći̯āu̯a-(ka-)* > MP *siyāh* > NP *syāh* 'black'

PIr **ći̯aina-(mr̥ga-)* > MP *sēn murw* 'a fabulous bird' > NP *sīmurɣ* (MacKenzie 1971:74). Young Avestan *saēna-* 'eagle' may show dissimilation of the first glide in the presence of the off-glide of the diphthong *ai*, which may also account for Persian *s*, but this development was clearly not pan-Iranian, since the initial consonant of Balochi *šēnak* 'falcon, hawk' (Korn 2005:129) cannot continue PIr **ć-*.

Word-internally, Middle and New Persian reflect variation between earlier **šii̯* and **θii̯*.

PIIr **matsi̯a-(ka-)* → **māći̯a-ka-* (Hoffmann 1976:637, fn. 25 attributes the long vowel to Vr̥ddhi) > MP *māhīg* 'fish' > NP *māhī*

PIIr **tusći̯a-(ka-)* > likely PIr **tući̯a-ka-* > MP *tuhīg* 'barren'

PIr **kaći̯apa-ka-* > MP *kašavag* 'tortoise' > NP *kašav*, *kašaf*; cf. Bandari *kāsapošt* 'turtle', Balochi *kasīp* ∼ *kasīb* 'turtle, tortoise', Kurmanji *kīsal*, Sivandi *kalapošt* (showing the East Iranian change **š* > *l*, according to Asatrian 2012; cf. Sims-Williams 1989:PP.), Bakhtiyari *kāsepošt* 'turtle', Zazaki *kese*

PIr **u̯aći̯ah-* > MP *wēš* 'more' > NP *bēš* 'more'; cf. Sivandi *vīštar*, Balochi (Marw) *geštir* 'greater, oftener'

PIr **kaći̯ah-* > MP *keh* 'smaller' > NP *kih* 'small'

PIr **maći̯ah-* > MP *mahy* (MMP <mhy>) 'bigger' > NP *mih* 'big'

Variation of this sort led Gershevitch (1962a:19–22) to argue that *θ* was an optional pronunciation of *s* in Old Persian. Others (cf. Hoffmann 1976:637, fn. 25) disagree; Klingenschmitt (2000:203) ascribes alternation between pre-OP **-ii̯a-* and **-i̯a-* to analogical suffix alternation, not phonological

[3] Verbal adjectives involving this form have found their way into Aśoka's Aramaic inscriptions (Schwarzschild 1960).

conditioning, with pre-OP **θi̯* yielding *šiy* and pre-OP **θi̯i̯* yielding *θiy*. Additionally, Cantera (2009) invokes a rhythmic law proposed by Klingenschmitt (*ibid.*) to account for phonological irregularities in Middle Persian nouns. Armed with these ideas, we can account for some of the variation within Persian, if we assume the pre-forms **māθ**íi̯a**-ka-* and **tuθ**íi̯a**-ka-* versus **kaćiápa-ka-* and **u̯áćiah-* (the stress placement assumed here follows Back 1978:30ff.), but this still does not explain *h* in reflexes of **kaćiah-* and **maćiah-*, which should have undergone the same development as **u̯áćiah-*, as noted by Gershevitch.

In Persian, as elsewhere in West Iranian, language contact, analogy, and prosodically conditioned change have interacted to bring about the complex variation seen in reflexes of PIr **ći̯* and related sounds. The limited knowledge we have of late Old Persian prosody can help to tease out the role of the last mechanism, but only to a certain extent. It is tempting to account for variation in West Iranian languages with no documented history in a similar manner, but this is purely speculative. An instructive example is the following thought experiment: Korn (2005:284) contends while discussing Balochi *kāsib/kasīp* 'turtle, tortoise' that "a genuine Bal. word should show *š*." However, given that Balochi *ī* reflects PIr **-i̯i̯ă̆-* (cf. Korn 2005:105), a pre-Balochi form **kaćiiapa-* is not inconceivable on historical phonological grounds, but perhaps overly speculative since we know virtually nothing about the phonotactics, syllabification, and stress pattern of Balochi's precursor. However, we also do not know whether *š* is the Balochi reflex of **ći̯* across the board (as assumed by Korn), or only in select environments. Ultimately, we may benefit from relaxing some of these assumptions and employing a probabilistic model that allows us to make generalizations regarding languages' diachronic behavior on the basis intra-language and inter-language distributions of sound changes.

4.2 PIr **θ*

PIr **θ* changes into OP *θ* (> MP, NP *h*) in most conditioning environments, though it may develop into MP word-initially, e.g., PIr **θaxta-* [cf. Khwarezmian *θɣd*] > NP *saxt* 'hard'. The change **θr* > *ç* [s] is well established, as is **θi̯* > *šy* (mentioned above), though numerous exceptions to these developments exist as well.

4.2.1 PIr **θn*

There are relatively few Proto-Iranian sources of the cluster **θn*, but these are realized as *šn* across the board in West Iranian, to the exclusion of the possible Median proper name in Akkadian *Pa-at-ni-e-ša-* = Med **Paθnī̯i̯ēša-* < **paθnī-aiša-* 'looking for a wife' (Tavernier 2007:273), and possibly the OP form (found in the Susa inscription of Darius) *kr̥nuvaka-* 'stonemason' (Schmitt 2009:133, 145), if from **kr̥t-nu-aka* following Kent (1942:80), though Kent (1951:180) is less certain regarding the presence of **-t-* and other scholars (e.g., Brust 2018:163) make no mention of etymological **-t-*.

PIr **araθni-* > OP *arašni-* 'cubit' > MP *ārešn* > NP *ārěš(n)*
PIr **dmāna-paθni-* > MP, Parthian *bāmbišn*
PIr **-i-θna-* > MP abstract noun suffix *-išn* > NP *-iš*; cf. Zazaki infinitive suffix *-iš* (Benveniste 1935:105)

Middle Persian *ārenč* > NP *āranj* 'elbow' is most likely a loan from a source closely related to Sogdian (cf. *ārinč*). If NP *āmvasnī* 'rival wife' is to be connected with **ham-paθnī-* (Tafazzoli 1974:119), it perhaps shows a secondary change **š* > *s* seen in other words.

4.3 PIr **št, *žd*

Old, Middle and New Persian (along with other Iranian languages) show variation between *st* and *št* for PIr **št*; later language attests variation between *zd* and *žd* (e.g., MP *mizd* 'reward', NP *mizd*, *mužd*). There is disagreement as to whether OP *st* for *št* is due to analogy (Kent 1951:34) or a sound change defining Southwest Iranian (Skjærvø 1989), and what the relationship of this behavior is to similar-looking developments in the later language. Lipp (2009:196ff.) states that OP *-st-* (found as a reflex of PIE **-k̂-t-*, **-ĝ-t-*) is due to analogy, while other developments are due to a phonological change predating Middle Persian:

1. PIE **h₃reĝ-to-* > PIr **rašta-* → OP *rāsta-* 'right' > MP *rāst* > NP *rāst*
2. PIr **mušti-* 'fist' > MP *mušt*, *must* > NP *must*
3. PIr **-išta-* (superlative suffix) > MP *-ist*; e.g., Phl *bālist*, MMP *bārist* 'highest' < **barǰišta-*; Phl *xwālist*, MMP *xwārist* 'sweetest' < **hu̯arǰišta-* (cf. Iron Ossetic *xorz*, Digor Ossetic *xuarz* 'good')

4.4 **r* + CORONAL change

4.4.1 Change to *l*

A number of West Iranian forms show a sound change whereby **r* + CORONAL sequences become *l*. This behavior is common in Middle and New Persian, perhaps representing a regular sound change which operated between Old and Middle Persian:

PIr **ǰ̥r**d**-* > MP *di**l*** > NP *di**l*** 'heart'
PIr **ů̯(a)**rd**a-* > MP *gu**l*** > NP *gu**l*** 'flower'
PIr **ća**rd**a-* > MP *sāl* (MMP *sār*) > NP *sāl* 'year'
PIr **p(a)**rd**anku-* > NP *palang* 'panther'; cf. Vedic *pŕ̥dāku-* (with meaning 'leopard' in the Paippalāda recension of the Atharva Veda, Zehnder 1999:59), SSog *pwrδ'nk* 'panther, leopard'[4]
PIr **b**r̥ǰ**ant-* > MP *buland* > NP *buland* 'high'
PIr **nard-* 'lament, moan' > MP *nāl-* > NP *nāl(īdan)* 'lament' (Cheung 2007:282)
PIr **ba**rǰ**ād(a)-*[5] > MP *bālāy* > NP *bālā* 'height'
PIr **ma**rǰ-*** > Phl *m'l-* /maːl-/ 'rub, sweep' > *mālīdan* 'rub, polish'
PIr **ŕ̥**ǰ**i̯a-* > MMP *"lwp/f* /aːluf/ > NP *āluh* 'eagle'
PIr **ǰa**rn**u-mani-* 'gold neck' > NP *dāl-man* 'black eagle' (Schwartz 1971:292, fn. 14)
PIr **g(a)r̥**n**a-ka-* (cf. Old Indic *gaṇá-*) > NP *gal(l)a* 'flock' (Schwartz 1971:292, fn. 14)
PIr **p̥**rt**u-* → **p̥rθu-* > MMP *pwhl* /puhl/ > NP *pŭl* 'bridge'
PIr **parćṵ-* → **parθ(a)ů̯-a-ka-* > MP *pahlūg* 'side, rib' > NP *pahlū* 'side'
PIr **čaθů̯(a)r-ćat-* (cf. Emmerick 1992:309) > PSWIr **čaθů̯**r**θat-* > MP *čehel* > NP *čihil* 'forty' (Emmerick 1992:309), Judeo-Tati *čül* (Authier 2012:88), cf. Zazaki *čewres* (Paul 1998a:61)

In some cases, this development has operated across an intervening vowel, likely unstressed:

PIr **ća**r(a)-d**āra-* (cf. Klingenschmitt 2000:194) > Phl 'Träger der Mund; Oberster' *sālār* (MMP *sārār* 'leader') > NP *sālār* 'leader' (cf. NP *sar-dār*, perhaps a later compound)

[4]This form is likely a Wanderwort, but seems reconstructible to Proto-Iranian; see Lubotsky 2001.

[5]Perhaps a de-instrumental *d*-stem built to PIE **b^{h}erǵ^h-eh$_1$-*, cf. Latin *mercēs*, *mercēdis* 'wages', *herēs*, *herēdis* 'heir' (Weiss 2009:304-5).

PIr **pa**ri-d**āna-* > NP *pālān* 'pack-saddle' (cf. Sogdian *pyrδnn* 'saddle', Sims-Williams cf. 1989:181)
PIr **pari-daiǰa-* > NP *pālēz* 'garden' (Cheung 2007:53)

However, this development is not exceptionless: it does not operate in forms like NP *padarzah* 'a wrapper in which clothes are folded up', if from **pari-darǰ-aka-* (Cheung 2007:63, marked as a loanword perhaps due to *z* < **ǰ*). It is unlikely that language-internal factors (viz., different conditioning environments) could account for the variation seen within Persian.

The lateralization of **r* + CORONAL sequences, from what we can see, post-dates Old Persian.[6] However, there are a large number of exceptions to this rule within Middle and New Persian; for example, NP *buland* is in a doublet with *burz*, thought to represent a Northwest Iranian form (Beekes 1997:3). For some etyma, Persian lacks *l*, while a non-Persian reflex displays it, e.g., NP *supurz* 'spleen' versus Kd *sipiɫ* '*id.*' < PIr **spr̥ǰan-*. The uncertainty surrounding this behavior can be summed up by the following comment by MacKenzie (1961:78) on the outcome of PIr **rd/rǰ* in Kurdish: "I do not think it is possible to be certain which is the true Kurdish development, but whether we consider the many words with *l/ɫ* as native or loan-words their preponderance is significant."

Gurani contains the forms *zil* 'heart' (< **ǰr̥d-*) and *wilī* 'flower' (< **u̯r̥da-*, suffixation unclear), which cannot be Persian loans; in the first, the change to OP *d* predates lateralization of **r/r̥d-*. In the second form, it is most probable that the change to *g-* in MP *gul* was triggered by the following **-r̥-*, which subsequently underwent lateralization (e.g., **u̯r̥da-* > **gVrd-* > *gul*, see §4.6).[7] Whether *l* in these forms owes itself to Persian influence as opposed to some other source is unclear.

4.4.2 **rn* > *rr*

The change **rn* > *rr* is attested in Middle Persian and onward, as seen in the following examples:[8]

**hu̯arnah-/*farnah-* > MP *xwarrah* > NP *farr* 'glory'
**u̯arna-ka-* 'wool' > Phl *wlk'*, MMP *wrg* /warrag/ > NP *barrah* 'lamb'
**parna-* > Phl *pl*, MMP *pr* /parr/ > NP *par(r)* 'feather'
OP *kr̥nuvakā* <*k-r-n*^{u}*-u-v-k-a*> 'stonemason' > Phl <*k(y)lwk'*>, MMP <*qrwg*> 'artisan', generally transcribed as *kirrog* on the basis of Armenian *krōgpet*, though late OP **uva* would seem to yield MP *ū* (Back 1978:80ff.)
**d(a)r-n-* > Phl *dl-* /darr/ > NP *darr-* 'to rend, tear up'
darrah
**us-pr̥na-* > Phl *spwl*, MMP *'spwr* /aspurr/ 'accomplished' (Klingenschmitt 2000:228)

It has been suggested that the changes **rn* > *rr* and **rn* > *l* are interconnected, and that *l(l)* ∼ *r(r)* variation in reflexes of **rn* represents dialectal variation within West Iranian (Schwartz 1971:292, fn. 14, who adduces Shirazi *vol* 'spider', partially reflecting **u̯arna-*, as well as *gallah* 'flock' and *dāl-man* 'black eagle' in favor of this sound change).

Middle and New West Iranian languages as a whole show an overwhelming tendency toward the change **rn* > *r(r)*. The word for 'lamb' (< **u̯arna-*) shows this behavior across the board:

[6] There are no good direct precursors of the forms listed above, but we see forms like OP *ardata-* 'silver' < PIr **arǰata-*, cognate to Yazdi *ālī* (Kent 1951:171).

[7] If Semnani *val(a)* reflects full-grade **u̯arda-*, it could in theory be a Persian loan, since changes affecting MP *wa-* post-date the **rd* > *l* change; however, there is no concrete evidence that Persian continues **u̯arda-*.

[8] The *rr* sequence found in NP *xurram* 'joyful, lucky' and some other forms may be secondary (cf. Horn 1893:106; Hübschmann 1895:55).

MP *warrag* > NP *barra*; Parthian *warrag*; Awromani *værā*; Balochi *gwārag*; S Bashkardi *vark*; Gurani *varāla*, *valala* (< **u̯arna-la-*?); Zazaki *vorek*; Shirazi *vol* 'spider'

Some Balochi and Parthian forms show the change **rn* > *n(n)*; Zazaki shows *rn* only via analogical maintenance or restoration, but otherwise $r \sim \bar{r}$ (Korn 2005:133–4).

4.5 $r \sim l$ variation

Proto-Indo-European **l* surfaces as *r* in the vast majority of Iranian languages. PIE **l* > **r* is often given as a Proto-Iranian sound change in most handbooks, yet there are a number of exceptions to this development (Schwartz 2008), indicating that PIE **l* has been conserved in some peripheral dialects. Northwestern dialects also contain morphological variants with *l* lost by Persian with congeners in Indic, e.g., Kashani *engulī*, Mazandarani *engel* (cf. Old Indic *aṅgúli-*) against NP *angušt* (cf. OInd *aṅgúṣṭha-*) 'finger' (Horn 1893; Krahnke 1976:226–8).[9]

However, some cases of West Iranian *l* may be secondary rather than archaic (Hübschmann 1895:262ff.). It is not clear, for example, where forms like S. Tati (Ebrahim-abadi) *nālbanda* ~ (Sagz-abadi) *nārbanda* 'elm' (Yar-shater 1969:71) = NP *nārvan* belong.[10] Similarly, one finds S. Tati *kelma* 'worm' = NP *kirm*; S. Tati *anjila* (Yar-shater 1969:71), Vidari *inǰil* (Baghbidi 2005:36) = NP *anjīr* 'fig' (forms elsewhere in Iranian point to **r*, e.g., Sogdian *ančēr*, *anjēr*; Gharib 1995:37). For 'worm', the evidence clearly points to an Indo-Iranian etymon **kr(i)mi-* containing *r*, whatever the exact shape (cf. Rg Vedic *kṛ́mi-*, Atharva Vedic *krími-*; Old Irish *cruim*), where any instances of *l* in Iranian languages should be secondary (e.g., Ossetic *kælm* shows expected **r* > *l* change in anticipation of **i* or **i̯*). Likely innovations are also found in Kurdish *valg*, Judeo-Tati *velg* (Miller 1892), etc. = NP *barg* < **u̯arka-* (Horn 1893:47); this variant surfaces in the Dari dialect of New Persian as *balg* (Korn 2005:160). Non-archaic *l* can also be found in NP *šikār* 'hunt' vs. Bandari, Bakhtiyari *eškāl*, S. Bashkardi *šekāl* 'mountain sheep', if from a verbal root **skar-* with no good Indo-European cognates (Cheung 2007:346). Ultimately, $r \sim l$ variation across West Iranian is due not only to preservation of original PIE **l*, but also a secondary change to *l* from original **r*, especially evident in loans originally from non-Iranian languages, e.g., Judeo-Isfahani *kelews* 'celery' = NP *karafs* (Stilo 2007) ← Arabic. We can be sure of the directionality in cases where there is secure evidence from outside of Indo-Iranian, but in the absence of such information, it can be difficult to tease apart primary and secondary *l*; it is equally unclear whether all variant pronunciations stem from the same dialectal source.

4.6 Changes to PIr **u̯-*

Reflexes of PIr **u̯-* are characterized by a high degree of irregularity across West Iranian.[11] Developments within Persian serve to demonstrate the complexity of these developments. Proto-Iranian **u̯-* surfaces as Middle Persian *g-* before **r̥* and **i̯*, but is otherwise unchanged in Middle Persian (with a few stray exceptions; see below):

[9] If the term for 'shepherd' in languages spoken in the Caspian region, *gāleš*, comes, as suggested by Asatrian (2002), from **gau̯a-raxša-* 'cow protector' (cf. Old Indic *go-rakṣa-*), then the presence of *l* is in agreement with PIE **h₂leks-*, pointing to another possible archaism.

[10] The elm appears to be the frequent target of folk etymology in Iranian languages (Henning 1963:70); it is possible that Tati speakers have conflated the tree's name with *nāl-band* 'smith, farrier' (Arabic *naʿl* 'horseshoe'), on the basis of some perceived but non-obvious connection to horseshoes.

[11] See Schwartz 1982 on certain conditioned reflexes of this sound.

PIr *u̯r̥tka- > MP *gurdag* > NP *gurdah* ‘kidney’
PIr *u̯r̥ka- > MP *gurg* > NP *gurg* ‘wolf’
PIr *u̯r̥pa-ka- > Phl *gwlbk′* /gurbag/ > NP *gurbah* ‘cat’ (cf. YAv *urupi-* ‘dog, fox (?)’ < *u̯r̥pi-)
PIr *u̯(a)r̥da-[12] > Phl *gwl* /gul/ > NP *gul* ‘flower, rose’
PIr *u̯r̥šna-ka-[13] ‘hungry’ > MP *gušnag*, *gursag*; NP *gušnah*, *gurusnah*; Bakhtiyari *gosne*; Balochi (Marw) *gušnag*; Gaz *vašše*; Larestani *gošna*; Mazandarani *vašnā*; Sivandi *feše* ‘qui a faim’ (showing the Central Iranian development *u̯- > *f-*, Asatrian cf. 2012); Taleshi *veši*; S Tati *gošna*; Zaz *veyšān*
PIr *u̯i̯āna- > OP **viyāna* (?) > MP *gyān* > NP *jān* ‘life, soul’
PIr *u̯i̯āka- > OP **viyāka* (?) > MP *gyāg* > NP *jāh* ‘place’

Generally speaking, PIr *u̯ĭ- > MP *wi-* > NP *gu-*:

PIr *u̯i-nāća- > MP *wināh* > NP *gunāh* ‘sin’
PIr *u̯i-čāra- > Phl *wc’l-*, MMP *wyc’r-* /wiza:r-/ > NP *guzār(dan)*
PIr *u̯i-dāna- > Phl *wyd’n*, MMP *wy’n* /wija:n/ > NP *giyān* ‘tent’ (cf. OInd *vi-dhā-* ‘furnish, spread, diffuse’?)
PIr *u̯a/inǰēča-ka- (?) > Phl *wncšk′*, Psalter *wnčšky* /windʒiʃk/ > *binjišk* ~ *gunjišk* ‘sparrow’, cf. Baxtiyārī *bingišt* (Schapka 1972:236); cf. Challi *veškenj* (Yar-shater 1969:69)

However, some exceptions exist:

MP *wiškar* > NP *bišgar(d)* ‘hunting ground’
MP *wiyābān* ‘desert’ > NP *biyābān*

In the following forms, PIr *u̯V(C)r- > MP *wV(C)r-* > NP *gV(C)r-*:

PIr *u̯arāǰa- > MP *warāz* > NP *gurāz* ‘boar’
PIr *u̯art- ‘turn’ > OP *v-r-t-* > Phl *wlt-*, MMP *wrd-* /ward/ > NP *gard(ūn)* ‘wheel, chariot’ (Cheung 2007:424–5)
PIr *u̯aǰra- > Phl *wlz* /warz/ > NP *gurz* ‘mace’ (cf. Bal *burz* ‘club’, Elfenbein 1963:25)

This development is blocked in the following words; most have a grave (i.e., labial, labiodental, or velar) consonant later in the word, but MP *war* ‘breast’ does not:

PIr *u̯arna-ka- > Phl *wlk’*, MMP *wrg* /warːag/ > NP *barra* ‘lamb’
PIr *u̯aǰr(a)-ka- > OP *v-z-r-k* /vazr̥ka-/[14] > Phl *wc(w)lg*, MMP *wzrg* /wuzurg/ (cf. Pazand *guzurg*, Bailey 1933:56) > NP *buzurg*
PIr *u̯arka- > MP *warg* ~ *walg* > NP *barg* ‘leaf’
PIr *u̯afra- > MP *wafr* ‘snow’ > **βafr-* (cf. Judeo-Persian <bp̄r> [bafr], Paul 2013:50) > NP *barf*
PIr *u̯ara- > MP *war* ‘breast’ > NP *bar*
PIr *u̯ar-ma- (? cf. Horn 1893:298) > MP *warm* ‘pond’ > NP *barm*

[12] YAv *varəδa-* ‘rose’ points to (and Semn *val(a)* ‘flower’ seems to point to — perhaps also Pth *w’r* /wa:r/) *u̯arda-, while the Persian forms, along with Gor *wilī*, may point to *u̯r̥da- (MacKenzie 1961:77 gives the former etymon for all these forms).

[13] See Klingenschmitt 2000:208 regarding the reconstruction of this form (departing from the earlier reconstruction of Hübschmann 1895:92), as well as the double reflex *š* ~ *r(V)s*.

[14] Schmitt (1989:69) and others give this reading, departing from *vazraka-* (found in Kent 1951), on the basis of the later forms.

MP *warm* 'memory' > NP *barm*
MP *wardag* 'captive, prisoner' > NP *barda*
PIr **u̯arǰa-* > Phl *wlc* /warz/ 'work, agriculture' > NP *barz* 'a sown field, agriculture' (cf. NP *varzīdan* 'sow a field', with *v-*)

Elsewhere, PIr **u̯-* > MP *w-* > NP *b-*:

PIr **u̯ahāna-ka-* (Gershevitch 1952) > Phl <wh'n(k)>, MMP <wh'n(g)> > NP *bahāna* 'reason, pretext'
PIr **u̯ahāra-* > MP *wahār* > NP *bahār* 'spring'
PIr **u̯ata-* > Phl *wt'*, MMP *wd* /wad/ > NP *bad* 'bad'
PIr **u̯at-čaka-* > MP *waččag* > NP *bačča* 'child'
PIr **u̯ana-* > Phl *wn* /wan/ > NP *bun* 'tree'
PIr **u̯āta-* > Phl *w't'*, MMP *w'd* > NP *bād* 'wind'
PIr **u̯īćati-* > Phl/MMP *wyst* /wi:st/ > NP *bīst*
PIr **u̯ahi̯a-* > MP *wah* > NP *bah-*
PIr **u̯aći̯a-* > MP *wyš* /we:ʃ/ > NP *bēš* 'more'
PIr **u̯ahišta-* 'best' > Phl *whšt*, MMP *whyšt* /wahiʃt/ > NP *bihišt*; cf. the name of a 4th cent. CE Christian martyr, *Gu(hi)štāzād* (Peeters 1910), the first member of which < **u̯ahišta-*
PIr **u̯a/inǰēča-ka-* (?) > Phl *wncšk'*, Psalter *wnčšky* /windʒiʃk/ > *binjišk* ~ *gunjišk* 'sparrow', cf. Baxtiyārī *bingišt* (Schapka 1972:236); cf. Challi *veškenj* (Yar-shater 1969:69), Bandari *jüšk*
PIr **u̯rῐ̄nǰi-* > Phl *blnc* /brindʒ/ > NP *birinj* ~ *gurinj* 'rice' (cf. AV+ *vrīhí-*)

As is apparent, none of the sound laws sketched above is exceptionless. It is almost certain that contact between closely related dialects is responsible for some of the doublets seen above. But it is also clear that succinct generalizations regarding the behavior of PIr **u̯-* in different conditioning environments are hard to come by. This issue has not received a systematic treatment in the literature. Lentz (1926:280–1) seems to consider **u̯-* > *b-* the regular Southwest Iranian outcome. MacKenzie (1971:76) takes the change **u̯-* > *b-* as a feature shared by Persian and Northern and Central Kurdish dialects, whereas "[i]n most other W.Ir dialects *w-* is little modified in this position, while in Bal. it has developed into *g(w)-*."

Attempts to establish the regular behavior of PIr **u̯-* for non-Persian West Iranian languages have proved as difficult as for Persian. Early Judeo-Persian records, thought to typify a link between Middle and Modern Persian, present an equally challenging picture (Paul 2013:35ff.). An errant strain of Middle Persian shows *g-* for expected *b-*, e.g., Pazand *guzurg* : NP *buzurg* (Bailey 1933:56). A large number of West Iranian languages leave **u̯-* more or less unmodified (surfacing as *v*, *w* or *f* but more importantly not merging with PIr **g-*, **b-*), but forms with *g-* and *b-* still preponderate. For instance, while Zazaki usually shows *v-* (e.g., *vā* 'wind'), the word for 'blood' is *gūnī* < **u̯ahuni-*[15] (Paul 1998b). South Tati *varga* 'leaf' sits alongside *behār-* 'spring' (Yar-shater 1969:95, 103, 110). The Kurmanji dialect of Kurdish shows a preference for *b-* where other languages do not, e.g., *burāz*, *vurāz* 'boar' : NP *gurāz*; *birsī*, *birchī* 'hungry' : NP *gurusnah* (Soane 1913, Thackston nd), but elsewhere agrees with Persian, e.g., *gurg*, *gūr* 'wolf'.

If a regular outcome can be established for a given non-Persian language, there is a tendency to assume that any words containing deviations from it are loans from Persian (though this approach

[15]The word for blood shows irregular historical phonology across west Iranian. NP *xūn* (MP *xōn*) has either undergone a metathesis between **u̯* and **h*, or was subject to the same irregular *x-*prothesis as MP *xāyag*, NP *xāya* 'egg', *xirs* 'bear'. Parthian has *guxn*, with unexpected *g-*; Sivandi has *fīn*.

is in general avoided by Korn 2005). For instance, Marw Balochi *burz* 'mace' (< **u̯ajra-*; note the metathesis identical to Persian) does not show expected *g(w)-*, hence, Elfenbein (1963:25) marks it as a "Persic" loan. However, there is no reason to expect NP *b-* in a reflex of a Middle Persian word with an initial syllable of the shape **u̯ar(C)-*, unless a grave consonant is found later in the word (and if the sound law sketched above is accurate). The Northern Kurdish dialect Kurmanji does, as mentioned above; this behavior can be found sporadically in other non-Persian languages as well (e.g., Mamasani *burâz* 'wild pig', Mann 1909:184). Given this evidence, these languages may be more viable donors for Balochi *burz* than Persian (the metathesis found in both of the forms is another question entirely).

4.7 Metathesis

Over the course of Persian history, more than one metathesis development has taken place (Hübschmann 1895:266–7), involving the re-sequencing of word-final and some word-internal clusters ending in *r* (and on occasion *l*). By the advent of Middle Persian, we see *narm* 'soft' < **namra-* and *warz* 'club, mace' < **u̯ajra-*. Fricative + *r/l* clusters (as well as some fricative + fricative clusters) have undergone metathesis after Middle Persian attestations:

PIr **u̯afra-* 'snow' > MP *wafr* > NP *barf*
PIr **taxra-* 'bitter' > Phl *taxl*, MMP *tahr* > NP *talx*; Phl *taxlīh* 'bitterness' > NP *talxī* (the latter change could be analogical)
PIr **čaxra-* 'wheel' > MP *čaxr* > NP *čarx*
PIr **aćru-* 'tear' > Phl *ars*, MMP *ars*, *asr* > NP *ars* (alongside *ašk* < **aćru-ka*)
PIE **?sup-sḱé-* 'sleep' (Rix et al. 2001:612) → MP *xwafs-* > NP *xusp(īdan)*

Other West Iranian languages vary as to whether they show metathesis in the same words; this variation is often language internal:

- PIr **u̯afra-* > Bal *barp*; Gaz *vaf*, *-varf* (in compounds); Gur *varwa*; Khun *varf*; Lar *vafr*, *barf*; Maz *varf*; Siv *varf*; Tal *var*; S Tati *vara*; Zaz *vewr*; Judeo-Tati *vähr* 'snow' can be found in the materials of Miller (1892:59), but Authier (2012:323) gives *verf*.
- PIr **taxra-* > Bal (Rakhshani) *ta(h)l*

Language contact must have played a role in bringing about intense variation, but the exact mechanisms are unclear. Metathesis is generally associated with Persian, since it can be documented in Persian's history. However, it is not clear whether the presence of metathesis in a non-Persian language is due to wholesale lexical borrowing or lexical diffusion (i.e., the adoption of the pronunciation *rC* for earlier *Cr*). Lexical borrowing from Persian tends to be assumed in the literature. For languages with *varf* : NP *barf*, it is assumed that the loan is from Middle Persian, or some period predating the change of MP *w-* to NP *b-*; for instance, Eilers (1978:749) derives Gazi *värf* 'snow' from MP *varf* [*sic*]. However, this is unlikely to be the case. If we take Judeo-Persian to be representative of the link between Middle and New Persian (cf. MacKenzie 2003), then Judeo-Persian forms like <*bp̄r*> [bafr] (Paul 2013:50) make it clear that metathesis postdates the merger of MP *w-* with *b-*, and that an intermediate stage **warf* was unlikely. Additionally, *w-*, *v-*, etc. cannot be secondary from earlier **b-* in the forms given above, since most of the languages mentioned show *b-* for original PIr **b-*.[16]

[16]Lenition often affects earlier intervocalic labial consonants, e.g., Qohrudi *vīxōvā* = NP *bē-xᵛāb* 'sleeplessness' (Žukovskij 1922:79), NP *bē-* < MP *abē* < **apa-ika-* (Durkin-Meisterernst 2014).

This detail aside, there are other reasons to question the account of lexical borrowing from Persian: first, this metathesis may not be a solely Persian development. Since most West Iranian languages (with exceptions, e.g., Yarshater 1962) lost final syllable nuclei, it is likely that many languages had words ending in *-xr*, *-fr*, etc., clusters which posed articulatory and perceptual problems, and were resolved in a variety of ways, including metathesis. Second, many of the above forms can be analyzed only as *Mischformen*, vitiating a lexical borrowing account. Instead, it is possible that speakers in a situation of heavy multilingualism imposed pronunciations from forms in one language upon their cognates in another, a well-documented phenomenon in situations of multidialectalism, generally affecting less frequently uttered words (Phillips 1984, Stollenwerk 1986, Wieling et al. 2011).

4.8 Changes affecting **dr*

Gershevitch (1962b:78–9) discusses reflexes of the word for 'spade', demonstrating that some modern West Iranian languages reflect a form **barda-* (metathesized from **badra-*, which is internally derived from **badar-*). The source of metathesis in **barda-* is unclear. (Schwartz 1971:297–8) shows that Iranian languages continue a doublet in the word for 'grape', **angudra-* (> MP, NP *angūr*) ∼ **angurda-(ka-)* (> NP *angurda*), the latter being secondary and a likely East Iranian loan into Persian and other languages. It is not clear whether the metathesis in **barda-* is a related phenomenon.[17]

4.9 Prothetic *x-*, *h-*

Two separate protheses have operated during the history of Persian. The first involves sporadic insertion of *x-* before an initial vowel, and predates Middle Persian; the second involves sporadic insertion of *h-* before an initial vowel, and predates New Persian.

These developments can be seen elsewhere in West Iranian, e.g., *xotkā* 'duck' (language unmarked by Asatrian 2012:113) < **āti-ka-*; Kumzari, Bandari, Larestani *xars* 'tear' < **aćru-* (cf. Bakhtiyari *hars*, Zazaki *hesri*). Korn (2005:155–9) provides a detailed treatment of this issue, and makes a strong case that some items showing initial *h-* in both Balochi and Kurdish are due to contact, though elsewhere, the sporadic presence of *h-* may be a sort of hypercorrection, as in many English dialects (Wells 1982:252–6), and not necessarily due to wholesale lexical borrowing (further bolstered by the fact that many Iranian languages lose initial *h-* under varying circumstances, e.g., **hiǰu̯āna-* 'tongue' > MP *izwān*, *uzwān*).

4.10 *č* ∼ *š*

Some quasi-systematic variation between *š* and *č* is found in forms across West Iranian. In some cases, original *č* becomes *š* due to the interference of Arabic, which lacks a phoneme *č* (in the relevant dialects), as in *šatranj* ∼ *šatrang* 'chess' < MP *čatrang* (← Old Indic *catur-aṅga-*).

In other forms, as noted by Horn (1901:71), *č* is secondary, e.g., Zor Yazdi *čūm* 'supper' = NP *šām* 'evening' (1st member < **xšap-*, cf. YAv *xšāfniia-*, Bartholomae 1904:553); Kashani *čiltúk* 'unhulled rice' = NP *šaltok*; Kashani *cepū́n* 'herdsman', Kurdish *čuwān* (*čōpān* 'butcher') = NP *čubān*, *šubān* <

[17] Bailey's (1973) derivation of the ethnonym Baloch from **baδlaut-čī* < **u̯adra-u̯at(č)ī* '[land] having water [channels]' (cf. the Greek toponym *Gedrosia*) is criticized by Korn (2005:47) on the grounds that there is no parallel for **dr* > **δl* > *l*. However, this form may speak to a near-identical metathesis to **badra-*, **angudra-* etc., though the change **dr* > **rd* > *l* is a not a common Balochi development.

**fšu-pāna-* (Horn 1901). Martin Schwartz (p.c.) points out that reflexes of the latter etymon may have undergone influence from NP *čūb* 'staff, crook'.

4.11 **t* > *r*

The change **t* > *r* in North Tati dialects was noted by Henning (1954:173). This change is seen in other languages, e.g., Judeo-Yazdi *čer-* 'go' (< **čiuta-*), Judeo-Isfahani *čer-* 'know' (< **čait-*), Kumzari *spīr*, North Bashkardi *espīr* 'white' < **ćuaita-*. Some Central Dialects show variation between *šīr* ∼ *šīt* for 'milk', though this may be due to the continuation of separate etyma **xšīra-* and **xšuifta-*.

4.12 Other developments

Above, a number of developments thought to be of interest to West Iranian dialectology were discussed. In this study, it is not possible to consider all possible meaningful changes, including vowel fronting (Krahnke 1976), *p* ∼ *f* variation (e.g., S Tati *fercel* 'dirty' : Bakhtiyari *parčal*), and other isoglosses. A hope is that as digitization efforts grow, fully data-driven approaches will allow us to take into account a wider range of innovations (see §9 for details).

4.13 Key Issues

The foregoing sections served to illustrate the difficulties posed for the traditional comparative method by West Iranian sound change. Along the way, some problematic analytical decisions made by scholars have been highlighted, which are restated here:

- Elfenbein (1963) assumes that Marw Balochi *burz* 'mace' is a Persian loan, given unexpected *b-*, but it could easily be from another language (§4.6)
- Eilers (1978) assumes that Gazi *värf* is a loan from Middle Persian **warf*, but no such form existed, given the relative chronology between the developments **u-* > *b-* and **-fr* > *rf*; if the metathesis shown by the Gazi form is due to Persian influence, lexical diffusion rather than lexical borrowing was likely involved (§4.7)
- Korn (2005) assumes that PIr **ći* > Balochi *š* in all conditioning environments, and hence, that Balochi *kāsib/kasīp* 'turtle, tortoise', is a loan, but we cannot be sure this is the case (§4.1.6)

It is hoped that the qualitative points made or revived here — namely that some of the segmental and prosodic contextual factors involved in West Iranian sound laws are indeterminate, that not all donor languages are necessarily Persian, and that pure lexical borrowing is not the sole mechanism of contact— are convincing on their own merits. Still, it remains difficult to resolve many of the questions raised above within the constraints of the traditional comparative method. In general, it is difficult to maintain a bird's-eye view of the many innovations and archaisms that cut across the West Iranian lexicon; while discussing one type of variation, another type is ignored (the above discussion is no exception). The remainder of this paper develops a probabilistic methodology designed to relieve historical linguists of the need to make hard decisions regarding phonological outcomes in a dialectal group, and instead let regularities fall out of the data.

5 Mixed Membership Models

As described above, West Iranian languages show admixture from an unknown number of latent (i.e., unobserved) dialectal components, each with its own individual sound laws and analogical changes. The key aim of our work is to learn which underlying components have contributed various features to the noisy pattern observed. A number of statistical techniques exist for the purpose of reducing the dimensionality of multivariate categorical data; mixed-membership models of this sort learn clusters that capture co-occurrence patterns of features in a data set in a way that the human eye cannot easily manage to do. These include certain classes of so-called generative models, which attempt to tell a story specifying one or more latent parameters which are thought to have generated the observed data. The latent parameters specified in a generative model can be estimated, usually within a Bayesian framework, which infers their posterior distributions. Bayesian modeling allows prior distributions to be imposed over these parameters, which serves as a sometimes-necessary means of ensuring that the model embodies realistic behavior.

Topic modeling, which seeks to identify the topics present in a set of documents by associating the words found in them with one or more topics, is well-known application for Bayesian mixed-membership models. Latent Dirichlet Allocation (LDA) is one such model (Blei et al. 2003); it assumes a fixed number of topics. It assumes that there is an overall distribution over possible topics, that each document has a specific distribution over topics, and that each word in each document is distributed according to a particular topic. The posterior global distribution over topics, document-specific topic distributions, and word-specific topic associations can then be inferred. LDA is highly similar to the Structure algorithm of population genetics (Pritchard et al. 2000), which has been used in some linguistic applications (Reesink et al. 2009, Syrjänen et al. 2016).

It is often unreasonable to assume that an exhaustive list of possible topics has been drawn up. LDA has a non-parametric extension, the Hierarchical Dirichlet Process (HDP, Teh et al. 2005, 2006), which allows for a potentially infinite number of topics. Over the course of the inference procedure, the model will return the number of topics which best explain the data. (It should be noted that if the procedure is entirely unsupervised, topics will receive meaningless labels such as "Topic 1" rather than "History," and that these labels require further interpretation.)

We wish to extend the HDP model to the problem of admixture in the vocabularies of Iranian languages. By aggregating the patterns of variation in reflexes of a number of Proto-Iranian etyma, we may be able to identify components in the lexicon of each language which conceivably can be explained via historical language contact. We assume that there exists a countably infinite set of areal components which underlie the variation reflected synchronically in West Iranian languages, and that we can recover their associations with variants and representation within languages.

An advantage of Bayesian models of this sort over classical methods for categorical data analysis is that they are generally robust to uneven or missing data — this is critical, given the patchy coverage for some Iranian languages. At the same time, mixed-membership models can potentially be sensitive to skews in data coverage. If a large number of features bearing on a particular isogloss are well attested in the data, but others are not, the algorithm used to infer component distributions may learn a distribution based on the former, even when the latter are highly relevant (but under-attested).[18] For this reason, we have taken pains to cast a wide net in our selection of features whilst maintaining parity in terms of the number of data points pertaining to each feature.

[18] A general practice is to remove uninformative and redundant features as well.

6 Feature selection and representation

For the upcoming analysis, words exhibiting the relevant Proto-Iranian sounds and sound sequences were collected from grammars and dictionaries by searching for the relevant semantic field, yielding a dataset of 1229 words. It is acknowledged that this means of data collection is highly limited, as some languages are better etymologized than others, and it would be preferable to take a top-down approach to data collection using a digitized etymological dictionary or etymological database, when such resources are developed.

As mentioned above, we are interested in teasing apart the effect of areal contact and conditioning environments within West Iranian. As a concrete example, the presence of *b* in Sorani Kurdish *baran* (< **u̯ār-*) versus *g* in Sorani Kurdish *gurg* (< **u̯r̥ka-*) is due either to contact (e.g., the language has taken the words over from different donor languages) or different conditioning environments in the two words triggering the changes **u̯* > *b-* and **u̯* > *g-*. Information regarding conditioning environments is key to the feature representation which serves as model input. However, explicitly stipulating conditioning environments requires too many assumptions. We use the ETYMON ITSELF as a proxy for conditioning environments; stating that **u̯* > *b-* in the etymon **u̯r̥ka-* WOLF is akin to stating that the change is triggered by the following **-r̥-* and/or the following **-k-*. This would be a highly uneconomical analysis for a traditional historical grammar; however, any redundancy that this representation entails will be picked up by the model as part of the dimensionality reduction that it carries out.

A potential concern is that morphological variants of the same etymon are reflected in our sample; as mentioned above, different languages many continue different variants of a historical doublet **u̯r̥da-*/**u̯arda-* 'flower'. A similar concern is that of homophony between reconstructed etyma, namely formally identical items that cannot be unified semantically (e.g., **u̯arma-* > NP *barm* 'pond' and **u̯arma-* > NP *barm* 'memory'). We leave the first problem untreated, with the hope that if a number of morphological variants of a single etymon are reflected in the data, this variation will be detectable in the model's output, namely via uncertainty in component level sound change distributions concerning this etymon.[19] We address the second problem by merging formally identical but semantically disparate reconstructions with one another, rather than treating them as instantiating different conditioning environments.

For the purposes of our model, each unobserved dialect component has a collection of SOUND CHANGE PARAMETERS associated with it. We envision this as a CATEGORICAL probability distribution over the POSSIBLE OBSERVED OUTCOMES for each PIr sound of interest in each etymon (our proxy for the CONDITIONING ENVIRONMENT). These parameters can be visualized as follows, for a given dialect component (probabilities are hypothetical):

Under the Neogrammarian hypothesis, sound change is exceptionless (Osthoff and Brugmann 1879, Bloomfield 1933, Hoenigswald 1965, Davies 1978). The probability of a sound change operating in a given speech variety is strictly categorical: one outcome will occur with 100% probability, all others with 0% probability. Our model RELAXES the Neogrammarian hypothesis, allowing sound change probabilities to be non-categorical. The first purpose is practical: rigid categorical-valued variables which assign zero, rather than infinitesimal probability mass to an outcome, will cause

[19] A pervasive issue in the historical morphology of Indo-Iranian languages is the widespread use of the **-aka-* suffix. The *k* of this suffix has been elided in most modern West Iranian languages, including New Persian (Pisowicz 1985), making it difficult to determine whether certain forms in fact reflect **-aka-*. In general, we do not make a distinction between suffixed and unsuffixed forms, unless there is clear widespread evidence for a suffix, as in the case of *ašk* 'tear', found in New Persian, Gilaki, and other dialects.

ETYMON	*u̯- > b	*u̯- > g	*u̯- > w/v	*rj́ > Vl	*rj́ > Vrz
*u̯ar-	.99	.005	.005		
*u̯rka-	.005	.005	.99		
*sprj́an-				.99	.01
*rj́ifia-				.97	.03

Table 2: Hypothetical sound change probabilities for a latent dialect component. Note that probabilities of outcomes for the relevant PIr sound(s) sum to one, and that distributions are SPARSE (with the majority of mass concentrated on one outcome).

problems for our inference procedure, and enumerating all possible combinations of categorical feature states is computationally unfeasible. The second pertains to the real world, namely, to account for irregularity within a component that cannot be explained (due to analogy, so-called "sporadic" change, or some other mechanism). However, it is still ideal to constrain these probability distributions such that they are SPARSE, with the majority of mass concentrated on one outcome, rather than SMOOTH (i.e., with mass distributed quasi-uniformly across outcomes). Ultimately, while we cannot constrain our model to enforce REGULAR sound change, we can employ priors that REGULARIZE sound change.

For the purposes of this study, we make no attempt to model intermediate stages in sound change. For instance, it is not entirely clear whether the *f-* in Sivandi *fin* 'blood' < **u̯ahuni-* comes from an intermediate **x(u̯)*, or directly from **u̯* (though the latter scenario is more likely, as such changes are better attested in Sivandi). Techniques have been proposed for reconstructing forms at intermediate nodes on fixed phylogenies (Bouchard-Côté et al. 2007, 2013), but not for situations like ours, where a form in a given language is generated by one of an unknown number of dialect components, rather than a single fixed ancestor. Our relatively abstract model of feature representation at least partly ensures that the sound changes dealt with by our model are meaningful. Our data set comprises 1160 sound change instances instantiating 190 unique sound change types in 32 West Iranian languages.

7 Inference

The generative process underlying the HDP and the technical details of inference can be found in the appendix. A non-technical description of the HDP follows. Each data point (i.e., the reflex of a Proto-Iranian sound in a particular etymon in a given language, e.g., PIr **u̯-* > NP *b-* in **u̯arma-*) is associated with a latent dialect component. The probability that a data point is associated with a given latent dialect component is dependent on a language-level probability distribution over dialect components $\boldsymbol{\theta}$, as well as a component-level distribution over sound changes ϕ. We do not know the values of these parameters, and must infer parameter values of high posterior probability (i.e., of high likelihood as well as high prior probability) from the data. Additionally, we do not know the true number of dialect components; this unknown must be learned by the model as well.

The HDP involves three hyperparameters: α is the concentration parameter of the symmetric Dirichlet prior over each dialect component's SOUND CHANGE distribution; the parameter γ controls the dispersion of data points across dialect components within a given language; δ controls the

number of components inferred (at the risk of oversimplifying). These hyperparameters can be fixed, or (as in the case of the parameters described in the previous paragraph) given a fully Bayesian treatment by estimating them from the data.

Parameter and hyperparameter values can be estimated in several ways, including Markov chain Monte Carlo (MCMC) approaches such as Gibbs Sampling (Geman and Geman 1984) or Variational Bayesian methods (Bishop 2006). In the former procedure, values for each parameter are sampled stochastically on the basis of current values of all other parameters; after many iterations, the Gibbs sampler is guaranteed to draw samples from the posterior distribution of each parameter. Variational methods can be either deterministic or stochastic, and unlike MCMC methods, they assume a parametric form of the posterior distribution of each variable, the parameters of which are iteratively updated. We use Automatic Differentiation Variation Inference (ADVI, Kucukelbir et al. 2017), as implemented in PyMC3 (Salvatier et al. 2016), estimating the MAP configuration of $\boldsymbol{\theta}$ and ϕ (as described in the appendix).

8 Results

As stated in the previous section, our inference procedure finds posterior probability distributions for two key parameters: $\boldsymbol{\theta}$, which gives each language's posterior distribution over dialect components; ϕ, which gives each dialect component's distribution over sound changes.

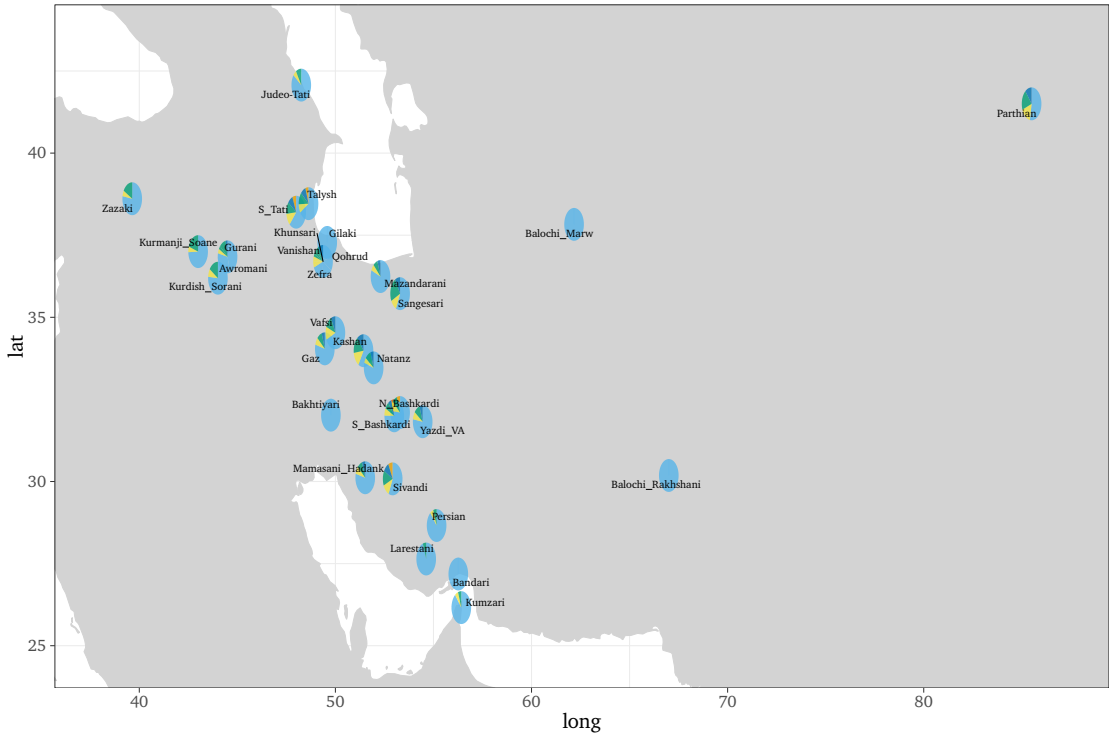

Figure 2: Language-level posterior distributions over latent dialect components

8.1 Language-level component distributions

As is clear from Figure 2, most languages in the sample show a relatively uniform profile in terms of their component makeup, favoring a small number of identical components, though the Middle Iranian language Parthian shows a different profile. This pattern dovetails with received wisdom regarding the widespread dominance of Persian over other West Iranian languages in the period following the Safavid empire roughly 500 years before the present day (Borjian 2009); this homogenization appears to have resulting in a more or less uniform profile for New West Iranian languages in terms of the sound changes reflected in their vocabularies (albeit with some degree of differentiation).

8.2 Posterior distributions over components for sound change instances

We use the MAP values of $\boldsymbol{\theta}, \phi$ to reconstruct the posterior probability that each individual sound change in each language in our data set is associated with a given dialect group, $P(z_i = k|\boldsymbol{\theta}, \boldsymbol{\phi})$. These probability distributions for each individual token — i.e., each sound change instance in each language — in the data set are given in the Appendix, as well as a table summarizing these values by averaging them across instances for each sound change type. These values allow us to address hypotheses about the provenance of certain sound changes (such as those discussed in §4.13). Many of these distributions exhibit high uncertainty; this is perhaps a consequence of the relatively small size of the data set used in this study. At first glance, this uncertainty may seem to make our results difficult to interpret, but on the contrary these results are quite interpretable in that this uncertainty is relatively informative. Consider the following posterior distributions, concerning reflexes of Proto-Iranian **bṛ́ant-* 'high' and **ć̯uaka-* 'dog', which show the posterior probability of a sound change type given a dialect component (we exclude components with probability mass under .05 for visual clarity).

etymon	sound	reflex	$p(k=1)$	$p(k=2)$	$p(k=3)$	$p(k=4)$	$p(k=5)$	$p(k=6)$
bṛ́ant	rí	l	0.58	0.11	0.15	0.08	0.04	0.02
bṛ́ant	rí	rí	1.00	0.00	0.00	0.00	0.00	0.00
ć̯uaka	ć̯u	s	0.01	0.23	0.34	0.16	0.09	0.10
ć̯uaka	ć̯u	sp	1.00	0.00	0.00	0.00	0.00	0.00

Tokens exhibiting the change **rí* > *rí* (our shorthand for forms such as *burz*, which do not undergo change to *l*) are associated strongly with a single latent dialect component, $k = 1$, as are token exhibiting the change *ć̯u* > *sp*. Tokens exhibiting the changes **rí* > *l* and **ć̯u* > *s* do not show a particularly strong affinity with any latent dialect component. What is critical here is that changes of the former type, usually associated with Northwest Iranian languages, show behavior that patterns much differently from changes usually associated with Southwest Iranian. This allows us to potentially classify individual change types according to whether the posterior distributions they exhibit are more in line with prototypical Northwest Iranian or Southwest Iranian sound changes.

On the basis of these distributions, we propose provisional solutions for the problems identified in §4.13. We find that Elfenbein's (1963) identification of Marw Balochi *burz* as a Southwest Iranian loan is indeed highly probable. Table 3 shows the component distributions of changes affecting PIr **u̯-* in *u̯aíra-* 'mace, club'. We see that change to *b-* shows a distribution similar to those of the prototypically Southwest Iranian sound changes discussed above, while change to *g-* and γ- shows Northwest Iranian behavior. Similarly, we find that change to *s* in **kaćiapa-* 'turtle/tortoise' patterns with canonically Northwest Iranian changes; hence, there is no strong reason to consider Balochi *kāsib/kasīp* a loan, as assumed by Korn (2005), since it patterns with many other typically Balochi features. Finally, changes concerning the etymon **u̯afra-* 'snow' suggest a Northwest Iranian origin

etymon	sound	reflex	$p(k=1)$	$p(k=2)$	$p(k=3)$	$p(k=4)$	$p(k=5)$	$p(k=6)$
u̯aǰra	u̯	b	0.08	0.17	0.39	0.17	0.07	0.02
u̯aǰra	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00
u̯aǰra	u̯	γ	1.00	0.00	0.00	0.00	0.00	0.00
kaćiapa	ći	š	0.03	0.24	0.28	0.15	0.06	0.18
kaćiapa	ći	s	1.00	0.00	0.00	0.00	0.00	0.00
u̯afra	u̯	b	0.70	0.06	0.12	0.04	0.03	0.03
u̯afra	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00
u̯afra	meta	meta	0.01	0.25	0.34	0.18	0.08	0.07
u̯afra	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00
u̯afra	meta	unclear	1.00	0.00	0.00	0.00	0.00	0.00

Table 3: Posterior component distributions for select sound changes (we exclude components with probability mass under .05 for visual clarity)

for the presence of *w-* and a Southwest Iranian origin for metathesis in the form; hence, Gazi *värf* is probably a genuine *Mischform*, *pace* Eilers (1978), stemming perhaps from a scenario where speakers in contact with a neighboring dialect exhibiting metathesis imposed this sound change on their inherited reflex of **u̯afra-*.

The results from our model are by no means the final word on these issues, and it is to be stressed that the conclusions drawn above are only tentative. It is likely that in many cases of idiosyncratic or unusual behavior, the paucity of data employed is the culprit. We have demonstrated however that this sort of methodology serves as a promising technique for teasing apart questions concerning dialectal admixture in Iranian and other dialect groups. We are confident that this method will produce increasingly realistic and reliable results as digital resources for Iranian languages grow, facilitating big data approaches to questions such as those addressed in this paper.

9 Discussion and future directions

In this paper, we outlined a series of unresolved problems in Iranian dialectology and developed a probabilistic methodology designed to address these problems. In doing this, for the most part, we sought proof of concept as to whether Bayesian applications to Iranian dialectology might yield results which shed light on outstanding problems in the field as well as those that jibe with received wisdom. We believe that this exercise was a success: we have shown that this model has great potential for resolving questions of the sort asked in this paper, but will greatly benefit from further refinement. Below, we identify future directions that will greatly improve this line of research:

9.1 Data

This paper made use of a relatively small data set compiled by hand from existing grammars. Sound changes were manually coded according to the behavior they displayed. Additionally, only sound changes thought to be of interest to West Iranian dialectology were included in the feature catalog. While we do not feel that this method of feature selection introduced any sort of pernicious bias that negatively affected our results — we were after all interested in the patterns displayed by sound changes thought to be probative for the purposes of Iranian dialect grouping across the vocabularies of West Iranian languages — it may be desirable to employ a more hands-off approach to feature selection and extraction, which will necessitate larger digitized etymological data sets. Additionally, this paper excluded East Iranian languages (including the languages Ormuri and Parachi), and shared

patterns across both East and West Iranian should not be neglected; again, fulfilling this desideratum requires bigger data. At least two tacks can be taken for the purpose of data expansion: the first would involve digitizing of existing etymological dictionaries (Cheung 2007, Rastorgueva and Ėdel'man 2003) and converting them into a computationally tractable data format; however, no complete Iranian etymological dictionary currently exists for all parts of the lexicon, though current efforts such as the *Atlas of the Languages of Iran* (Anonby et al. 2019), in its pilot phase at the time of writing, work towards filling this gap. The second approach involves applying semi-supervised cognate detection methods (List 2012, Rama 2016) to digitized Iranian word lists, which can potentially be coupled with semi-supervised methodologies for linguistic reconstruction (Meloni et al. 2019). While these methods still face many challenges, they can potentially save specialists a great deal of time and work in compiling large etymological resources. Whatever the approach employed, we believe that methods of the sort introduced in this paper will greatly benefit from the use of a larger data set. It is possible that the use of different data may yield different results from those reported in this paper.

9.2 Models

While this paper employed the HDP, several alternative types of nonparametric mixed-membership model exist. The HDP has certain properties that are undesirable for certain uses, possibly including the dialectological application explored in this paper: specifically, the proportion of a component across all data points is correlated with its proportion within languages. It may be the case that a certain component is very rare overall, but well represented within one or a small number of languages. Certain alternatives to the HDP deal explicitly with this issue (Williamson et al. 2010).

9.3 Representation of sound change

In designing our methodology, we made the radical decision to make no prior assumptions about the nature of the conditioning environments involved in the sound changes under study, instead treating entire etyma as conditioning environments. At first blush, this may seem like an implementation of the dictum that every word has its own history, attributed to dialect geographers such as Jules Gilliéron and Hugo Schuchardt. This is not the case: by linking the diachronic behavior of Proto-Iranian sounds in individual etyma to a finite number of dialect components exhibiting regularized sound change, we have learned patterns of sound change within components as well as patterns of admixture within languages; our model ultimately embodies the interpretation of the above problem posed by dialect geographers that was provided by Bloomfield (1933:360).

At the same time, it may be wrong to ignore the effect of phonetic similarity between conditioning environments on sound change. It may be the case that in a particular dialect component, *u̯͡- undergoes a particular type of change in similar-looking etyma like *u̯͡ahi̯͡a- and *u̯͡aći̯͡a-, but a different change in a more dissimilar etymon such as *u̯͡arka-. We have ignored this possibility; our goal was to let this systematicity fall out of the data in a bottom-up fashion. If desired, it is possible to employ a prior over sound change that can express covariance, such as the logistic normal distribution, which will encourage Proto-Iranian sounds to behave similarly in phonetically similar environments (which can potentially be operationalized via a smooth kernel function of the edit distance between the etyma containing these environments).

10 Conclusion

This paper introduced a new way of looking at Iranian dialectal relationships. The focus was on sound change in West Iranian, but this method can potentially be extended to linguistic groups of similar geographic spread and time depth. Our chief goal was to provide a means for relaxing assumptions regarding the operation of individual sound changes in individual languages, and allow regular patterns to fall out of the data. Much work remains to be done in order to understand the complex history of the Iranian languages. Larger data resources are needed, and cooperation among linguists is needed in order to design and refine the probabilistic models we use; as data analysts, we need to work together to characterize the stochastic processes that we believe to have generated the data we observe, formalized in probabilistic terms. There needs to be a willingness to simplify models (if particular models are intractable), and an effort to keep models flexible, so that they can be expanded. We believe that many of these goals are well within reach.

References

Anonby, E. and A. Asadi (2014). *Bakhtiari studies: phonology, text, lexicon*. Uppsala: Acta Universitatis Upsaliensis.

Anonby, E., M. Taheri-Ardali, and A. Hayes (2019). The *Atlas of the Languages of Iran* (ALI): A research overview. *Iranian Studies 52*, 199–230.

Asatrian, G. (2002). The Lord of Cattle in Gilan. *Iran and the Caucasus 6*, 75–85.

Asatrian, G. (2012). Marginal remarks on the history of some Persian words. *Iran and the Caucasus 16*, 105–116.

Authier, G. (2012). *Grammaire juhuri, ou judeo-tat, langue iranienne des Juifs du Caucase de l'est*. Beitrage Zur Iranistik. Dr. Ludwig Reichert Verlag.

Azami, C. A. and G. Windfuhr (1972). *A Dictionary of Sangesari with a Grammatical Outline*. Tehran: Franklin Book Programs.

Back, M. (1978). *Die sassanidischen Staatsinschriften*. Leiden: Brill.

Baghbidi, H. R. (1383 [2005]). Guyeš-i vidari. *Guyeš Šenāsi 2*(19), 18–26.

Bailey, H. W. (1933). Western Iranian dialects. *Transactions of the Philological Society*, 46–64.

Bailey, H. W. (1973). Mleccha, Balōč, and Gadrōsia. *Bulletin of the School of Oriental and African Studies 36*(3), 584–587.

Barker, M. A. (1969). *A Course in Baluchi*. Montreal: Institute of Islamic Studies, McGill University.

Bartholomae, C. (1883). *Handbuch der altiranischen Dialekte (Kurzgefasste vergleichende Grammatik, Lesestücke und Glossar)*. Leipzig: Breitkopf & Härtel.

Bartholomae, C. (1904). *Altiranisches Wörterbuch*. Strassburg [Strasbourg]: Karl J. Trübner.

Beekes, R. S. P. (1997). Historical phonology of Iranian. *Journal of Indo-European Studies 25*(1-2), 1–26.

Benedictsen, Å. M. and A. Christensen (1921). *Les dialectes d'Awromān et de Pāwä*, Volume 6 of *Historisk-filosofiske Meddelelser*. Copenhagen: Det Kgl. Danske Videnskabernes Selskab.

Benveniste, E. (1935). *Les infinitifs avestiques*. Paris: Adrien Maisonneuve.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Berlin & New York: Springer.

Blau, J. (1980). *Manuel de Kurde (dialecte Sorani)*. Paris: Klincksieck.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of Machine Learning research 3*, 993–1022.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Borjian, H. (2009). Median succumbs to Persian after three millennia of coexistence: Language shift in the Central Iranian Plateau. *Journal of Persianate Studies 2*, 62–87.

Bouchard-Côté, A., D. Hall, T. L. Griffiths, and D. Klein (2013). Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences 110*, 4224–4229.

Bouchard-Côté, A., P. Liang, T. Griffiths, and D. Klein (2007). A probabilistic approach to diachronic phonology. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, Prague, pp. 887–896. Association for Computational Linguistics.

Brust, M. (2018). *Historische Laut- und Formenlehre des Altpersischen: mit einem etymologischen Glossar*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Cantera, A. (2009). On the history of the Middle Persian nominal inflection. In W. Sundermann, A. Hintze, and F. de Blois (Eds.), *Exegisti Monumenta: Festschrift in Honour of Nicholas Sims-Williams*, Volume 17 of *Iranica*, pp. 17–30. Wiesbaden: Harrassowitz.

Cathcart, C. (2015). *Iranian dialectology and dialectometry*. Ph. D. thesis, University of California, Berkeley.

Cheung, J. (2007). *Etymological dictionary of the Iranian verb*. Leiden: Brill.

Davies, A. M. (1978). Analogy, segmentation and the early Neogrammarians. *Transactions of the Philological Society 76*, 36–60.

Durkin-Meisterernst, D. (2004). *Dictionary of Manichaean Texts III*. Turnhout: Brepols.

Durkin-Meisterernst, D. (2014). *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Efimov, V. A. (1986). *Jazyk ormuri: v sinxronnom i istoričeskom osveščenii*. Moscow: Nauka.

Eilers, W. (1976). *Westiranische Mundarten aus der Sammlung Wilhelm Eilers. Vol. 1: Die Mundart von Chunsar.* Wiesbaden: Steiner. With assistance from Ulrich Schapka.

Eilers, W. (1978). *Westiranische Mundarten aus der Sammlung Wilhelm Eilers. Vol. 2: Die Mundart von Gäz*. Wiesbaden: Steiner. With assistance from Ulrich Schapka.

Elfenbein, J. (1963). *A vocabulary of Marw Baluchi*. Naples: Istituto Universitario Orientale di Napoli.

Emmerick, R. E. (1992). Iranian. In J. Gvozdanović (Ed.), *Indo-European Numerals*, pp. 289–346. Berlin & New York: Mouton de Gruyter.

Geiger, W. (1901). Kleinere Dialekte und Dialektgruppen. In W. Geiger and E. Kuhn (Eds.), *Grundriss der iranischen Philologie*, Volume 1, Chapter 8, pp. 287–423. Strassburg [Strasbourg]: Karl J. Trübner.

Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 6*, 721–741.

Gershevitch, I. (1952). Ancient survivals in Ossetic. *Bulletin of the School of Oriental and African Studies 14*(3), 483–495.

Gershevitch, I. (1962a). Dialect variation in early Persian. *Transactions of the Philological Society 63*, 1–29.

Gershevitch, I. (1962b). Outdoor terms in Iranian. In W. B. Henning and E. Yarshater (Eds.), *A locust's leg: studies in honor of S.H. Taqizadeh*, pp. 76–84. London: Percy Lund, Humphries & Co.

Gharib, B. (1995). *Sogdian Dictionary: Sogdian-Persian-English*. Tehran: Farhangan Publications.

Grierson, G. A. (1918). The Ōrmuṛī or Bargistā language, an account of a little-known Eranian dialect. *Memoirs of the Asiatic Society of Bengal 7*, 1–101.

Hadank, K. (1930). *Kurdisch-persische Forschungen, Abt. 3 (Nordwestiranisch) Bd. 2, Mundarten der Gûrân besonders das Kändûläî, Auramânî und Bâdschälânî*. Berlin: Walter de Gruyter.

Hammarström, H., R. Forkel, and M. Haspelmath (2017). Glottolog 3.3. Max Planck Institute for the Science of Human History.

Henning, W. B. (1954). The ancient language of Azerbaijan. *Transactions of the philological society*, 157–177.

Henning, W. B. (1963). The Kurdish elm. *Asia Major 10*(1), 68–72.

Hoenigswald, H. M. (1965). *Language change and linguistic reconstruction*. Chicago: University of Chicago Press.

Hoffmann, K. (1976). Zur altpersischen Schrift. In J. Narten (Ed.), *Karl Hoffmann: Aufsätze zur Indoiranistik*, Volume 2, pp. 620–645. Wiesbaden: Ludwig Reichert Verlag.

Hoffmann, K. and B. Forssman (2004). *Avestische Laut- und Flexionslehre* (2nd ed.), Volume 84 of *Innsbrucker Beiträge zur Sprachwissenschaft*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Horn, P. (1893). *Grundriss der neupersischen Etymologie*. Strassburg [Strasbourg]: Karl J. Trübner.

Horn, P. (1901). Neupersische Schriftsprache. In W. Geiger and E. Kuhn (Eds.), *Grundriss der iranischen Philologie*, Volume 1, pp. 1–200. Strassburg [Strasbourg]: Karl J. Trübner.

Hübschmann, H. (1895). *Persische Studien*. Strassburg [Strasbourg]: Karl J. Trübner.

Ishwaran, H. and L. F. James (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association 96*(453), 161–173.

Ivanow, W. (1940). *The Gabri dialect spoken by the Zoroastrians of Persia*, Volume 16 of *Rivista degli Studi Orientali*. Rome: Scuola Orientale nella R. Università di Roma.

Kamioka, K. and M. Yamada (1979). *Lārestani Studies*, Volume 1. Tokyo: Institute for the Study of Cultures of Asia and Africa.

Kent, R. (1942). Vocalic *r* in Old Persian before *n*. *Language 18*(2), 79–82.

Kent, R. (1951). *Old Persian*, Volume 33 of *American Oriental Series*. New Haven: American Oriental Society.

Kieffer, C. (1989). Le parāčī, l'ōrmuṛī. In R. Schmitt (Ed.), *Compendium Linguarum Iranicarum*, pp. 445–455. Wiesbaden: Ludwig Reichert Verlag.

Kingma, D. P. and J. Ba (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.

Klingenschmitt, G. (2000). Mittelpersisch. In B. Forssman and R. Plath (Eds.), *Indoarisch, Iranisch und die Indogermanistik: Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*, pp. 191–229. Wiesbaden: Ludwig Reichert Verlag.

Korn, A. (2003). Balochi and the concept of North-Western Iranian. In C. Jahani and A. Korn (Eds.), *The Baloch and their neighbours: ethnic and linguistic contacts in Balochistan in historical and modern times*, pp. 49–60. Wiesbaden: Dr. Ludwig Reichert Verlag.

Korn, A. (2005). *Towards a Historical Grammar of Balochi*. Wiesbaden: Ludwig Reichert Verlag.

Korn, A. (2019). Isoglosses and subdivisions of iranian. *Journal of Historical Linguistics 9*(2), 239–281.

Krahnke, K. (1976). *Linguistic Relationships in Central Iran*. Ph. D. thesis, University of Michigan.

Kucukelbir, A., D. Tran, R. Ranganath, A. Gelman, and D. M. Blei (2017). Automatic differentiation variational inference. *The Journal of Machine Learning Research 18*(1), 430–474.

Kümmel, M. (2007). *Konsonantenwandel*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lecoq, P. (1979). *Le dialecte de Sivand*, Volume 10 of *Beiträge zur Iranistik*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lentz, W. (1926). Die nordiranischen Elemente in der neupersischen Literatursprache bei Firdosi. *Zeitschrift für Indologie und Iranistik 4*, 251–316.

Lipp, R. (2009). *Die indogermanischen und einzelsprachlichen Palatale im Indoiranischen*. Heidelberg: Carl Winter. 2 vols.

List, J.-M. (2012). Lexstat: Automatic detection of cognates in multilingual wordlists. In *Proceedings of the EACL 2012 Joint Workshop of LINGVIS & UNCLH*, pp. 117–125. Association for Computational Linguistics.

Lubotsky, A. (2001). The Indo-Iranian substratum. In C. Carpelan, A. Parpola, and P. Koskikallio (Eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at the Tvärminne Research Station of the University of Helsinki 8-10 January 1999*, Helsinki, pp. 301–317.

MacKenzie, D. N. (1961). The origins of Kurdish. *Transactions of the Philological Society*, 68–86.

MacKenzie, D. N. (1971). *A Concise Pahlavi Dictionary*. London: Oxford University Press.

MacKenzie, D. N. (2003). The missing link. In L. Paul (Ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*, pp. 103–110. Wiesbaden: Harrassowitz.

Mann, O. (1909). *Kurdisch-Persische Forschungen, Abt. 1: Die Tâjîk-Mundarten der Provinz Fârs*. Berlin: Reimer.

Mann, O. and K. Hadank (1906-1932). *Kurdisch-persische Forschungen*. Berlin: Walter de Gruyter.

Meloni, C., S. Ravfogel, and Y. Goldberg (2019). Ab antiquo: Proto-language reconstruction with RNNs. *arXiv preprint arXiv:1908.02477*.

Miller, V. F. (1892). *Materialy dlja izučenija evrejsko-tatskogo jazyka*. St. Petersburg: [Tip. Imp. Akademij Nauk].

Morgenstierne, G. (1926). *Report on a linguistic mission to Afghanistan*. Oslo: H. Aschehoug & Co.

Morgenstierne, G. (1929). *Parachi and Ormuri*, Volume 1 of *Indo-Iranian Frontier Languages*. Oslo: Instituttet for Sammenlignende Kulturforskning, H. Aschehoug & Co. (W. Nygaard).

Morgenstierne, G. (1932). Persian etymologies. *Norsk Tidsskrift for Sprogvidenskap 5*, 54–56.

Nawata, T. (1984). *Mazandarani*, Volume 17 of *Asian and African Grammatical Manual*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Oranskij, I. M. (1963 [1977]). *Les Langues Iraniennes*. Translated by Joyce Blau. Paris: Klincksieck.

Osthoff, H. and K. Brugmann (1879). *Morphologische Untersuchungen auf dem Gebiet der indogermanischen Sprachen*, Volume 2. Leipzig: Hirzel.

Paul, D. (2011). *A comparative dialectal description of Iranian Taleshi*. Ph. D. thesis, University of Manchester.

Paul, L. (1998a). The position of Zazaki among West Iranian languages. In N. Sims-Williams (Ed.), *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995. Part I: Old and Middle Iranian Studies*, Wiesbaden, pp. 163–177. European Conference of Iranian Studies: Dr. Ludwig Reichert Verlag.

Paul, L. (1998b). *Zazaki: Grammatik und Versuch einer Dialektologie*, Volume 18 of *Beiträge zur Iranistik*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Paul, L. (2005). The language of the *Šāhnāme* in historical and dialectal perspective. In D. Weber (Ed.), *Languages of Iran: Past and Present: Iranian Studies in Memoriam David Neil MacKenzie*, pp. 141–151. Wiesbaden: Dr. Ludwig Reichert Verlag.

Paul, L. (2013). *A Grammar of Early Judaeo-Persian*. Wiesbaden: Ludwig Reichert Verlag.

Peeters, P. (1910). S. Eleutherios-Guhištazad. *Analecta Bollandiana 29*, 151–156.

Pelevin, M. (2010). Materials on the Bandari dialect. *Iran and the Caucasus 14*, 57–78.

Phillips, B. S. (1984). Word frequency and the actuation of sound change. *Language 60*, 320–342.

Pisowicz, A. (1985). *Origins of the new and middle Persian phonological systems*. Kraków: Uniwersytet Jagielloński.

Pritchard, J. K., M. Stephens, and P. Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics 155*(2), 945–959.

Rama, T. (2016, December). Siamese convolutional networks for cognate identification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, pp. 1018–1027. The COLING 2016 Organizing Committee.

Rastorgueva, V. S. and D. I. Ėdel'man (2000-2003). *Ėtimologičeskij slovar' iranskix jazykov*. Moscow: Vostočnaja Literatura.

Rastorgueva, V. S., A. A. Kerimova, A. K. Mamedzade, L. A. Pireiko, and J. I. Edelman (2012). *The Gilaki Language*. Uppsala: Uppsala Universitet.

Reesink, G., R. Singer, and M. Dunn (2009). Explaining the linguistic diversity of Sahul using population models. *PLoS Biology 7*, e1000241.

Rix, H., M. Kümmel, T. Zehnder, R. Lipp, and B. S. (eds.) (2001). *Lexikon der indogermanischen Verben, die Wurzeln und ihre Primärstammbildungen* (2nd ed.). Wiesbaden: Ludwig Reichert Verlag.

Salemann, C. (1901). Mittelpersisch. In W. Geiger and E. Kuhn (Eds.), *Grundriss der iranischen Philologie*, Volume 1, Chapter 3, pp. 249–332. Strassburg [Strasbourg]: Karl J. Trübner.

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science 2*, e55.

Schapka, U. (1972). *Die persischen Vogelnamen*. Ph. D. thesis, University of Würzburg.

Schmitt, R. (1989). Altpersisch. In R. Schmitt (Ed.), *Compendium Linguarum Iranicarum*, pp. 56–85. Wiesbaden: Ludwig Reichert Verlag.

Schmitt, R. (2009). *Die altpersischen Inschriften der Achaimeniden*. Wiesbaden: Ludwig Reichert Verlag.

Schulze, W. (2000). *Northern Talysh*. Munich: Lincom.

Schwartz, M. (1970 [1971]). On the Khwarezmian version of the *muqaddimat al-adab* as edited by Johannes Benzing. *Zeitschrift der Deutschen Morgenländischen Gesellschaft 120*, 288–304.

Schwartz, M. (1982). "Blood" in Sogdian and Old Iranian. In *Monumentum Georg Morgenstierne II*, pp. 189–196. Leiden: Brill.

Schwartz, M. (2006). On Haoma, and its liturgy in the Gathas. In A. Panaino and A. Piras (Eds.), *Proceedings of the 5th Conference of the Societas Iranologica Europæa*, Volume 1, Milan, pp. 215–224. Mimesis.

Schwartz, M. (2008). Iranian *l, and some Persian and Zaza etymologies. *Iran and the Caucasus 12*, 281–287.

Schwarzschild, L. A. (1960). Review of *Un Editto Bilingue Greco-Aramaico di Aśoka* by G. Pugliese Carratelli and G. Levi Della Vida. *Journal of the American Oriental Society 80*, 155–157.

Sims-Williams, N. (1989). Eastern Middle Iranian. In R. Schmitt (Ed.), *Compendium Linguarum Iranicarum*, pp. 165–172. Wiesbaden: Dr. Ludwig Reichert Verlag.

Sims-Williams, N. (1996). Eastern Iranian languages. *Encyclopædia Iranica 7*, 649–652.

Skjærvø, P. O. (1988). Baškardi. *Encyclopædia Iranica 3*(8), 846–850.

Skjærvø, P. O. (1989). Pashto. In R. Schmitt (Ed.), *Compendium Linguarum Iranicarum*, pp. 384–410. Wiesbaden: Ludwig Reichert Verlag.

Skjærvø, P. O. (2009). Old Iranian. In G. Windfuhr (Ed.), *The Iranian languages*, pp. 43–195. London: Routledge.

Soane, E. B. (1913). *Grammar of the Kurmanji or Kurdish language*. London: Luzac.

Steingass, F. J. (1892). *A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*. London: Routledge & K. Paul.

Stilo, D. (1981). The Tati Language Group in the Sociolinguistic Context of Northwestern Iran and Transcaucasia. *Iranian Studies 14*(3/4), 137–187.

Stilo, D. (2004). *Vafsi Folk Tales*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Stilo, D. (2005). Iranian as a buffer zone between the universal typologies of turkic and semitic. In E. Csató, B. Isaksson, and C. Jahani (Eds.), *Linguistic Convergence and Areal Diffusion. Case Studies from Iranian, Semitic and Turkic*, pp. 35–63. Routledge.

Stilo, D. (2007). Isfahan xix. Jewish dialect. *Encyclopaedia Iranica 14*(1), 77–84.

Stilo, D. (2018). Numeral classifier systems in the praxes-iran linguistic area. In W. B. McGregor and S. Wichmann (Eds.), *The Diachrony of Classification Systems*, pp. 135–164. Amsterdam: Benjamins.

Stollenwerk, D. A. (1986). Word frequency and dialect borrowing. *Ohio State University Working Papers in Linguistics 34*, 133–141.

Syrjänen, K., T. Honkola, J. Lehtinen, A. Leino, and O. Vesakoski (2016). Applying population genetic approaches within languages: Finnish dialects as linguistic populations. *Language Dynamics and Change 6*, 235–283.

Tafazzoli, A. (1974). Pahlavica II. *Acta Orientalia 36*, 113–23.

Tavernier, J. (2007). *Iranica in the Achaemenid period (ca. 550-330 B.C.): lexicon of old Iranian proper names and loanwords*. Leuven: Peeters.

Tedesco, P. (1921). Dialektologie der westiranischen Turfantexte. *Le Monde Oriental 15*, 184–258.

Teh, Y. W., M. I. Jordan, M. J. Beal, and D. M. Blei (2005). Sharing clusters among related groups: Hierarchical Dirichlet processes. In *Advances in Neural Information Processing Systems*, pp. 1385–1392.

Teh, Y. W., M. I. Jordan, M. J. Beal, and D. M. Blei (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association 101*, 1566–1581.

Thackston, W. M. (n.d.). Kurmanji Kurdish. a reference grammar with selected readings.

Thomas, B. (1930). *The Kumzari dialect of the Shihuh tribe, Arabia and a vocabulary*. London: The Royal Asiatic society.

Vahman, F. and G. Asatrian (2002). *Notes on the Language and Ethnography of the Zoroastrians of Yazd*, Volume 85 of *Historisk-filosofiske Meddelelser*. Copenhagen: The Royal Danish Academy of Sciences and Letters.

van der Wal Anonby, C. (2015). *A grammar of Kumzari: A mixed Perso-Arabian language of Oman*. Ph. D. thesis, Rijksuniversiteit te Leiden.

Žukovskij, V. A. (1888-1922). *Materialy dlja izučenija persidskix narečij*. St. Petersburg: [n.s.] 3 vols.

Weiss, M. (2009). *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave Press.

Wells, J. (1982). *Accents of English 1: an introduction*. Cambridge: Cambridge University Press.

Wendtland, A. (2009). The position of the Pamir languages within East Iranian. *Orientalia Suecana 58*, 172–188.

Wieling, M., J. Nerbonne, and R. H. Baayen (2011). Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PloS one 6*(9), e23613.

Williamson, S., C. Wang, K. A. Heller, and D. M. Blei (2010). The IBP compound dirichlet process and its application to focused topic modeling. In *Proceedings of the 27th International Conference on Machine Learning*. Haifa, Israel.

Windfuhr, G. (2009). Dialectology and topics. In G. Windfuhr (Ed.), *The Iranian languages*, Chapter 2, pp. 5–42. New York; London: Routledge.

Yar-shater, E. (1969). *A Grammar of Southern Tati Dialects*. Number 1 in Median Dialect Studies. The Hague: Mouton.

Yarshater, E. (1962). The Tati dialects of Ramand. In W. B. Henning and E. Yarshater (Eds.), *A locust's leg: studies in honor of S.H. Taqizadeh*, pp. 240–245. London: Percy Lund, Humphries & Co.

Zehnder, T. (1999). *Atharvaveda-Paippalāda, Buch 2, Text, Übersetzung, Kommentar: eine Sammlung altindischer Zaubersprüche vom Beginn des 1. Jahrtausends v. Chr.* Idstein: Schulz-Kirchner.

Appendix

Model specification and inference

The generative process for the HDP involving the truncated stick-breaking construction (Ishwaran and James 2001) is given below. We set the truncation cutoff T at 10.

Draw hyperparameters α, δ, γ
$\boldsymbol{\beta} \sim \text{GEM}(\gamma)$ [draw atoms governing number of components learned]
$\boldsymbol{\theta}_\ell \sim \text{Dirichlet}(\delta\boldsymbol{\beta}) : \ell \in \{1, ..., \mathcal{L}\}$ [draw language-level distributions over component membership]
$\boldsymbol{\phi}_{t,s} \sim \text{Dirichlet}(\alpha) : t \in \{1, ..., T\}, s \in \{1, ..., S\}$ [draw sound change parameters for every environment s and every component t]
For $i \in \{1, ..., N\}$ [for each data point (i.e., sound change instance)]
  $z_i \sim \text{Categorical}(\boldsymbol{\theta}_{\ell_i})$ [draw a component label]
  $y_i | z_i = k \sim \text{Categorical}(\boldsymbol{\phi}_{k,x_i})$ [sample the observed reflex on the basis of parameters associated with the label drawn in the previous step]

GEM(γ) denotes the Griffiths-Engen-McCloskey distribution, which has the following function when parameterized by γ:

$$\beta'_t \sim \text{Beta}(1, \gamma) : t \in \{1, ..., T\}$$

$$\beta_t = \begin{cases} \beta'_t \prod_{i=1}^{t-1}(1 - \beta'_i) & \text{if } t \in 1, ..., T-1 \\ \prod_{i=1}^{t-1}(1 - \beta'_i) & \text{if } t = T \end{cases}$$

Under this process, each data point has the following likelihood:

$$P(y_i, x_i, z_i = k | \boldsymbol{\theta}, \boldsymbol{\phi}) = P(z_i = k | \boldsymbol{\theta}_{\ell_i}) P(y_i | \boldsymbol{\phi}_{k,x_i})$$

Marginalizing out the discrete variable $\boldsymbol{z}$ yields the following likelihood:

$$P(y_i, x_i | \boldsymbol{\theta}, \boldsymbol{\phi}) = \sum_{k=1}^{K} P(z_i = k | \boldsymbol{\theta}_{\ell_i}) P(y_i | \boldsymbol{\phi}_{k,x_i})$$

We place uninformative Gamma$(1, 1)$ priors over δ and γ, since we do not know *a priori* the degree to which data points within a given language should be dispersed across components, or how many components we should expect to find. We fix α, the concentration parameter of the symmetric Dirichlet prior over each dialect component's SOUND CHANGE distributions, at $.0001$ to encourage sparse sound change distributions. We carry out inference using ADVI in PyMC3 over 4 separate initializations of 100000 iterations each, monitoring the evidence lower bound (ELBO) for convergence. The learning rate and β_1 parameters of the Adam optimizer (Kingma and Ba 2015) are

set to $.01$ and $.8$, respectively. Posterior samples for each parameter are constructed by drawing 500 samples from the fitted variational posterior.[20]

Mixture models suffer from the so-called label switching problem, in which indices of identical components differ across initializations/chains. To address this problem, we relabel the components inferred across initializations 2–4 by permuting component labels and selecting the permutation which minimizes the Kullback-Leibler divergence from the parameters for initialization 1 to the permuted parameters for the initialization under consideration. This allows us to average parameters across initializations, giving us an approximation to the maximum a posteriori (MAP, e.g., of highest posterior probability) configuration over component assignments for each item in our data set. Aggregating over these assignments produces MAP language-level distributions over component makeup.

Posterior distributions over dialect components for sound change instances, averaged by type

The following tape gives $P(z_i|\boldsymbol{\theta}, \boldsymbol{\phi})$ for every sound change instance in our data set, averaged across sound change TYPES.

etymon	sound	reflex	$p(k=1)$	$p(k=2)$	$p(k=3)$	$p(k=4)$	$p(k=5)$	$p(k=6)$	$p(k=7)$	$p(k=8)$	$p(k=9)$	$p(k=10)$
(f)šupāna	š	š	0.32	0.15	0.26	0.10	0.06	0.06	0.02	0.01	0.01	0.01
(f)šupāna	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
jŗd	rd	l	0.68	0.08	0.11	0.06	0.02	0.02	0.01	0.01	0.01	0.00
jŗd	rd	rd	0.92	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00
ćiāųa	ći	š	0.01	0.23	0.27	0.14	0.08	0.21	0.02	0.02	0.02	0.01
ćiāųa	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ćųaka	ćų	s	0.01	0.23	0.34	0.16	0.09	0.10	0.02	0.02	0.02	0.01
ćųaka	ćų	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ćųiša	ćų	s	0.83	0.04	0.05	0.03	0.01	0.03	0.00	0.00	0.00	0.00
ćųiša	ćų	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųa)rda	ų	g	0.01	0.22	0.35	0.16	0.08	0.10	0.02	0.02	0.02	0.02
ųa)rda	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųa)rda	rd	l	0.72	0.06	0.10	0.05	0.02	0.03	0.01	0.00	0.00	0.00
ųa)rda	rd	rd	0.48	0.15	0.18	0.10	0.04	0.02	0.02	0.01	0.01	0.01
ųājí(jaka?)	ų	b	0.52	0.12	0.17	0.06	0.04	0.06	0.01	0.01	0.01	0.01
ųājí(jaka?)	ų	g	0.99	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
ųājí(jaka?)	ų	w	0.96	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
ųājí(n)	ų	b	0.10	0.22	0.34	0.15	0.08	0.05	0.02	0.02	0.01	0.01
ųājí(n)	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųāf	ų	b	0.47	0.14	0.16	0.10	0.03	0.06	0.01	0.01	0.01	0.01
ųāf	ų	g	0.99	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
ųāf	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųār	ų	b	0.73	0.05	0.11	0.04	0.02	0.03	0.01	0.00	0.00	0.00
ųār	ų	g	0.95	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųār	ų	w	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųāta	ų	b	0.63	0.08	0.14	0.05	0.03	0.04	0.01	0.01	0.01	0.00
ųāta	ų	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųāta	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųiāka	ų	g	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
ųiāka	ų	j	0.79	0.05	0.08	0.04	0.02	0.02	0.00	0.00	0.00	0.00
ųiāna	ų	g	0.42	0.14	0.22	0.10	0.05	0.03	0.02	0.01	0.01	0.01
ųiāna	ų	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųiāna	ų	y	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųŗća-ka-	ų	g	0.12	0.18	0.31	0.15	0.09	0.07	0.02	0.02	0.02	0.02
ųŗća-ka-	rć	l	0.24	0.16	0.27	0.13	0.08	0.06	0.02	0.02	0.01	0.02
ųršnaka	ų	b	0.01	0.18	0.49	0.10	0.10	0.07	0.01	0.01	0.01	0.01
ųršnaka	ų	f	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųršnaka	ų	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųršnaka	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųršnaka	rs	š	0.01	0.24	0.32	0.18	0.08	0.09	0.03	0.02	0.02	0.02
ųršnaka	rs	rs	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųŗka	ų	g	0.34	0.16	0.23	0.11	0.06	0.06	0.02	0.01	0.01	0.01
ųŗka	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųŗtka	ų	g	0.02	0.24	0.34	0.16	0.08	0.11	0.02	0.02	0.02	0.01
ųŗtka	ų	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųaŗr(a)ka	ų	b	0.48	0.13	0.17	0.10	0.05	0.04	0.01	0.01	0.01	0.01
ųaŗr(a)ka	ų	g	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ųaŗr(a)ka	ų	w	0.91	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[20] Code can be found at `https://github.com/chundrac/w_ir_layers` .

u̯aɟra	u̯	b	0.08	0.17	0.39	0.17	0.07	0.02	0.03	0.02	0.02	0.02
u̯aɟra	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aɟra	u̯	ɣ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aćia	ći̯	š	0.71	0.08	0.09	0.06	0.02	0.02	0.01	0.01	0.00	0.00
u̯aćia	ći̯	s	0.64	0.12	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.00
u̯aćia	u̯	b	0.02	0.27	0.30	0.17	0.07	0.09	0.02	0.02	0.01	0.01
u̯aćia	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aćia	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯afra	u̯	b	0.70	0.06	0.12	0.04	0.03	0.03	0.00	0.00	0.00	0.00
u̯afra	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯afra	meta	meta	0.01	0.25	0.34	0.18	0.08	0.07	0.02	0.02	0.02	0.01
u̯afra	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯afra	meta	unclear	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯agza	u̯	b	0.62	0.08	0.18	0.05	0.03	0.03	0.01	0.00	0.00	0.00
u̯agza	u̯	g	0.98	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
u̯agza	u̯	m	0.89	0.00	0.00	0.00	0.03	0.08	0.00	0.00	0.00	0.00
u̯agza	u̯	w	0.96	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.00
u̯ahāra	u̯	b	0.01	0.25	0.31	0.18	0.08	0.10	0.02	0.02	0.02	0.01
u̯ahāra	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahi̯a	u̯	b	0.35	0.18	0.21	0.12	0.05	0.04	0.01	0.01	0.01	0.01
u̯ahi̯a	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahi̯a	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahuni	u̯	f	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
u̯ahuni	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahuni	u̯	h	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahuni	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ahuni	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aiša	u̯	b	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
u̯aiša	u̯	w	0.72	0.07	0.09	0.05	0.02	0.02	0.01	0.01	0.01	0.00
u̯aicĩna?	u̯	b	0.03	0.20	0.39	0.11	0.09	0.13	0.01	0.01	0.01	0.01
u̯aicĩna?	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aicĩna?	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aina	u̯	b	0.02	0.24	0.33	0.14	0.07	0.14	0.02	0.01	0.01	0.01
u̯aina	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aina	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯aita	u̯	b	0.20	0.18	0.31	0.12	0.07	0.07	0.02	0.01	0.01	0.01
u̯aita	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ana	u̯	b	0.02	0.26	0.28	0.17	0.09	0.10	0.02	0.02	0.02	0.01
u̯ana	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯anjecaka	u̯	b	0.06	0.23	0.28	0.16	0.06	0.14	0.02	0.02	0.02	0.01
u̯anjecaka	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯anjecaka	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arāja	u̯	b	0.03	0.21	0.41	0.15	0.09	0.04	0.02	0.02	0.02	0.01
u̯arāja	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arāja	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ara	u̯	b	0.74	0.07	0.07	0.03	0.02	0.06	0.00	0.00	0.00	0.00
u̯ara	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ara	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arka	u̯	b	0.72	0.07	0.09	0.04	0.03	0.03	0.01	0.00	0.00	0.00
u̯arka	u̯	w	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arka	r	l	0.00	0.25	0.36	0.19	0.08	0.04	0.03	0.02	0.02	0.01
u̯arka	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arna-ka	u̯	b	0.70	0.07	0.10	0.05	0.03	0.03	0.01	0.01	0.00	0.00
u̯arna-ka	u̯	g	0.97	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arna-ka	u̯	w	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯arna-ka	rn	(r)r	0.01	0.24	0.34	0.17	0.09	0.07	0.03	0.02	0.02	0.02
u̯arna-ka	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯art-	u̯	g	0.02	0.25	0.30	0.18	0.07	0.10	0.02	0.02	0.02	0.02
u̯art-	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯asi	u̯	b	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
u̯asĩ	u̯	w	0.62	0.11	0.12	0.08	0.03	0.01	0.01	0.01	0.01	0.00
u̯at-caka	u̯	b	0.63	0.08	0.14	0.06	0.03	0.04	0.01	0.01	0.01	0.01
u̯at-caka	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ata	u̯	b	0.03	0.22	0.32	0.16	0.08	0.12	0.02	0.02	0.02	0.02
u̯ata	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ićati	u̯	b	0.01	0.23	0.34	0.17	0.08	0.09	0.02	0.02	0.02	0.02
u̯ićati	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ićati	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ïapāna	u̯	b	0.05	0.24	0.26	0.14	0.07	0.17	0.02	0.02	0.02	0.01
u̯ïapāna	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ïapāna	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯icāra	u̯	g	0.69	0.08	0.09	0.05	0.03	0.05	0.01	0.01	0.01	0.00
u̯idāna	u̯	b	0.57	0.10	0.10	0.08	0.03	0.08	0.01	0.01	0.01	0.01
u̯idāna	u̯	g	0.98	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯idāna	u̯	w	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯idau̯ā	u̯	b	0.02	0.24	0.33	0.17	0.09	0.08	0.02	0.02	0.02	0.01
u̯idau̯ā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯ināća	ć	h	0.74	0.06	0.09	0.04	0.02	0.03	0.01	0.00	0.00	0.00
u̯ināća	u̯	g	0.02	0.22	0.34	0.15	0.08	0.12	0.02	0.02	0.02	0.01
u̯ināća	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u̯rinji	u̯	b	0.72	0.07	0.09	0.05	0.02	0.02	0.01	0.00	0.00	0.00
u̯rinji	u̯	g	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
aćua	ću̯	s	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
aćua	ću̯	sp	0.71	0.07	0.10	0.05	0.02	0.02	0.01	0.01	0.00	0.00
aćru	ćr	rs	0.02	0.24	0.28	0.17	0.08	0.14	0.02	0.02	0.02	0.02

aćru	ćr	sr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aćru	pro	no pro	0.26	0.18	0.21	0.12	0.06	0.12	0.02	0.01	0.01	0.01
aćru	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
arša	pro	no pro	0.00	0.18	0.50	0.12	0.10	0.06	0.01	0.01	0.01	0.01
arša	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
arša	rs	š	0.67	0.10	0.06	0.08	0.05	0.01	0.01	0.01	0.01	0.01
arša	rs	rch	0.43	0.00	0.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00
arša	rs	rs	0.63	0.00	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00
bṛzant	rj́	l	0.58	0.11	0.15	0.08	0.04	0.02	0.01	0.01	0.01	0.01
bṛzant	rj́	rj́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
barda	rd	l	0.02	0.23	0.35	0.16	0.08	0.10	0.02	0.02	0.01	0.01
barda	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
barzād	rj́	l	0.68	0.08	0.12	0.05	0.03	0.02	0.01	0.01	0.00	0.00
barzād	rj́	rj́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
caθu̯ŕćat	ŕć	l	0.00	0.25	0.34	0.18	0.08	0.07	0.03	0.02	0.02	0.02
caθu̯ŕćat	ŕć	ŕć	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
daća	ć	h	0.71	0.07	0.10	0.05	0.02	0.02	0.01	0.01	0.00	0.00
daća	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
darn	rn	(r)n	0.03	0.20	0.41	0.11	0.09	0.12	0.01	0.01	0.01	0.01
darn	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
g(a)rna-ka	rn	(r)n	0.02	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
g(a)rna-ka	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
gau-ćuanta-	ću̯	s	0.82	0.05	0.05	0.04	0.01	0.00	0.00	0.00	0.00	0.00
gau-ćuanta-	ću̯	sp	0.76	0.06	0.08	0.04	0.02	0.02	0.01	0.00	0.00	0.00
hiǰu̯āna	ǰu̯	zb	0.01	0.25	0.32	0.18	0.06	0.10	0.03	0.02	0.02	0.02
hiǰu̯āna	ǰu̯	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
hiǰu̯āna	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
kṛmi	r	l	0.58	0.11	0.11	0.08	0.02	0.06	0.01	0.01	0.01	0.01
kṛmi	r	r	0.69	0.08	0.11	0.06	0.02	0.02	0.01	0.01	0.00	0.00
kaćiapa	ći	š	0.03	0.24	0.28	0.15	0.06	0.18	0.02	0.02	0.01	0.01
kaćiapa	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
laupāća	ć	h	0.64	0.09	0.14	0.06	0.03	0.02	0.01	0.01	0.01	0.01
laupāća	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
laupāća	l	l	0.59	0.11	0.14	0.07	0.04	0.02	0.01	0.01	0.01	0.01
laupāća	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mušti	št	št	0.53	0.12	0.17	0.08	0.04	0.03	0.01	0.01	0.01	0.01
mušti	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
pṛtu	rt	l	0.58	0.13	0.11	0.09	0.04	0.03	0.01	0.01	0.01	0.01
pṛtu	rt	rt	0.91	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
parna	rn	(r)n	0.01	0.22	0.29	0.16	0.12	0.13	0.02	0.02	0.02	0.02
parna	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
parna	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
puθra	tr	s	0.70	0.08	0.09	0.05	0.02	0.04	0.01	0.01	0.01	0.00
puθra	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
rāšta	št	št	0.00	0.25	0.38	0.18	0.10	0.02	0.03	0.02	0.02	0.01
rāšta	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
skar	r	l	0.09	0.20	0.24	0.16	0.12	0.11	0.02	0.02	0.02	0.02
skar	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
spṛzan	rj́	l	0.37	0.14	0.22	0.12	0.05	0.05	0.01	0.01	0.01	0.01
spṛzan	rj́	rj́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
xšu̯ifta-	t	r	0.62	0.11	0.13	0.08	0.03	0.01	0.01	0.01	0.01	0.00
xšu̯ifta-	t	t	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01

Posterior distributions over dialect components for sound change instances, raw values

The following tape gives $P(z_i|\boldsymbol{\theta}, \boldsymbol{\phi})$ for every sound change instance in our data set.

Gurani	aću̯a	asp	ću̯	sp	0.60	0.08	0.19	0.06	0.03	0.01	0.01	0.01	0.00	0.01
Gurani	arša	haš	rs	š	0.67	0.11	0.06	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Gurani	arša	haš	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	arša	xirs	rs	rs	0.41	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	arša	xirs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	barzād	bālā	rj́	l	0.54	0.09	0.22	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Gurani	bṛzant	barz	rj́	rj́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	caθu̯ŕćat	cil	ŕć	l	0.00	0.21	0.49	0.16	0.07	0.02	0.01	0.02	0.01	0.01
Gurani	ćiāu̯a	siyā(u)	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	daća	dah	ć	h	0.57	0.09	0.21	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Gurani	g(a)rna-ka	gallā	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	hiǰu̯āna	zūān	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	mušti	mišt	št	št	0.42	0.12	0.28	0.09	0.04	0.01	0.01	0.01	0.01	0.01
Gurani	parna	pal	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	pṛtu	pirdi	rt	rt	0.88	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯ahāra	vahār	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯(a)rda	wīlī	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯(a)rda	wīlī	rd	l	0.60	0.08	0.19	0.06	0.03	0.01	0.01	0.01	0.00	0.01
Gurani	u̯afra	varwa	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯afra	varwa	meta	meta	0.00	0.21	0.49	0.16	0.07	0.02	0.01	0.02	0.01	0.01
Gurani	u̯ahuni	winī	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯ahi̯a	waš	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯aicina?	vicina	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	u̯ār	vārān	u̯	w	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Gurani	ɰarna-ka	valala	ɰ	w	0.84	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰarna-ka	valala	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰarna-ka	varāla	ɰ	w	0.84	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰarna-ka	varāla	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰāta	vā	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰrka	warya	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gurani	ɰiāka	jā	ɰ	j	0.60	0.08	0.19	0.06	0.03	0.01	0.01	0.01	0.00	0.01
Gurani	ɰiāna	gyān	ɰ	g	0.36	0.13	0.31	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Gurani	ɾɽd	zil	rd	l	0.57	0.09	0.21	0.07	0.02	0.01	0.01	0.01	0.01	0.01
Khunsari	aćɰa	asm	ćɰ	sp	0.71	0.08	0.10	0.06	0.02	0.01	0.01	0.01	0.00	0.00
Khunsari	barda	bar	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	caθɰrćat	cil	rć	l	0.00	0.26	0.34	0.20	0.08	0.03	0.04	0.02	0.02	0.02
Khunsari	ćɰaka	isba	ćɰ	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	híjɰāna	zebūn	ɟɰ	zb	0.01	0.26	0.34	0.20	0.08	0.03	0.04	0.02	0.02	0.02
Khunsari	híjɰāna	zevūn	ɟɰ	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	parna	par	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	puθra	pir	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰafra	varf	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰafra	varf	meta	meta	0.01	0.26	0.34	0.20	0.08	0.03	0.04	0.02	0.02	0.02
Khunsari	ɰahuni	xin	ɰ	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰār	vārūn	ɰ	w	0.93	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰara	var	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰarka	valg	ɰ	w	0.92	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰarka	valg	r	l	0.00	0.26	0.34	0.20	0.08	0.03	0.04	0.02	0.02	0.02
Khunsari	ɰarna-ka	berra	ɰ	b	0.63	0.10	0.10	0.08	0.03	0.01	0.02	0.01	0.01	0.01
Khunsari	ɰarna-ka	berra	rn	(r)r	0.01	0.26	0.34	0.20	0.08	0.03	0.04	0.02	0.02	0.02
Khunsari	ɰart-	vardana	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰasi	vas	ɰ	w	0.71	0.08	0.10	0.06	0.02	0.01	0.01	0.01	0.00	0.00
Khunsari	ɰat-caka	vec(c)a	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰata	vad	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Khunsari	ɰrka	gurg	ɰ	g	0.26	0.19	0.25	0.15	0.06	0.02	0.03	0.01	0.01	0.01
Khunsari	ɰiāna	giyūn	ɰ	g	0.48	0.14	0.18	0.11	0.04	0.01	0.02	0.01	0.01	0.01
Khunsari	ɾɽd	dil	rd	l	0.68	0.08	0.11	0.07	0.02	0.01	0.01	0.01	0.01	0.01
Khunsari	gau-ćɰanta-	gūsband	ćɰ	sp	0.71	0.08	0.10	0.06	0.02	0.01	0.01	0.01	0.00	0.00
Khunsari	gau-ćɰanta-	gūsfand	ćɰ	sp	0.71	0.08	0.10	0.06	0.02	0.01	0.01	0.01	0.00	0.00
Gaz	aćɰa	asm	ćɰ	sp	0.76	0.07	0.08	0.05	0.02	0.00	0.01	0.00	0.00	0.00
Gaz	barda	bard	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	barda	bērci	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	brzant	biland	rʃ	l	0.66	0.10	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Gaz	caθɰrćat	cil	rć	l	0.00	0.29	0.33	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	g(a)rna-ka	gal(l)e	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	híjɰāna	uzūn	ɟɰ	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	laupāća	lū	ć	h	0.65	0.10	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Gaz	laupāća	lū	l	l	0.70	0.09	0.10	0.06	0.02	0.01	0.01	0.00	0.01	0.01
Gaz	mušti	mušt	št	št	0.60	0.12	0.13	0.08	0.03	0.01	0.01	0.01	0.01	0.01
Gaz	ɰićati	vīs	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰ(a)rda	gul	ɰ	g	0.01	0.28	0.32	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰ(a)rda	gul	rd	l	0.76	0.07	0.08	0.05	0.02	0.00	0.01	0.00	0.00	0.00
Gaz	ɰafra	vaf	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰafra	vaf	meta	unclear	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰafra	varf	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰafra	varf	meta	meta	0.01	0.29	0.33	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰahuni	xūn	ɰ	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰaiša	vēš	ɰ	w	0.76	0.07	0.08	0.05	0.02	0.00	0.01	0.00	0.00	0.00
Gaz	ɰaita	bid	ɰ	b	0.15	0.24	0.28	0.17	0.07	0.02	0.03	0.01	0.02	0.01
Gaz	ɰana	van	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰanjecaka	gunjišg	ɰ	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰār	vārūn	ɰ	w	0.94	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰara	var	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰarka	valg	ɰ	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰarka	valg	r	l	0.00	0.29	0.33	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰarna-ka	vere	ɰ	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰarna-ka	vere	rn	(r)r	0.01	0.28	0.32	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰarna-ka	vire	ɰ	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰarna-ka	vire	rn	(r)r	0.01	0.28	0.32	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰasi	vas	ɰ	w	0.76	0.07	0.08	0.05	0.02	0.00	0.01	0.00	0.00	0.00
Gaz	ɰat-caka	vac(c)e	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰata	vat	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰāta	vā	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰrka	gurg	ɰ	g	0.31	0.20	0.23	0.14	0.05	0.01	0.02	0.01	0.01	0.01
Gaz	ɰršnaka	vašše	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɰršnaka	vašše	rs	š	0.01	0.29	0.33	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰrtka	gurde	ɰ	g	0.01	0.28	0.32	0.20	0.08	0.02	0.03	0.02	0.02	0.02
Gaz	ɰiāna	jūn	ɰ	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gaz	ɾɽd	dil	rd	l	0.73	0.08	0.09	0.05	0.02	0.00	0.01	0.00	0.01	0.00
Gaz	gau-ćɰanta-	gūsfand	ćɰ	sp	0.76	0.07	0.08	0.05	0.02	0.00	0.01	0.00	0.00	0.00
Mazandarani	brzant	belan	rʃ	l	0.69	0.08	0.11	0.07	0.02	0.01	0.01	0.01	0.01	0.01
Mazandarani	caθɰrćat	cehel	rć	l	0.00	0.26	0.34	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	ćɰaka	sag	ćɰ	s	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	ćɰiša	‘espij	ćɰ	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	daća	da	ć	h	0.75	0.06	0.08	0.05	0.02	0.01	0.01	0.01	0.01	0.00
Mazandarani	híjɰāna	zabun	ɟɰ	zb	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	mušti	mis	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	ɰahāra	bahār	ɰ	b	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	ɰićati	bist	ɰ	b	0.01	0.26	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	ɰafra	varf	ɰ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	ɰafra	varf	meta	meta	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02

Mazandarani	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯ār	vāreš	u̯	w	0.95	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯arka	valg	u̯	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯arka	valg	r	l	0.00	0.26	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	u̯at-caka	vaca	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯āta	vā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯āǰi(n)	vāz	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯rka	varg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯ršnaka	vašnā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mazandarani	u̯ršnaka	vašnā	rs	š	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	u̯rtka	gārda	u̯	g	0.01	0.25	0.33	0.22	0.06	0.03	0.03	0.02	0.02	0.02
Mazandarani	ǰrd	del	rd	l	0.75	0.06	0.08	0.06	0.01	0.01	0.01	0.01	0.01	0.00
Mazandarani	gau-ć̯anta-	gesfan	ć̯u̯	sp	0.78	0.06	0.07	0.05	0.01	0.01	0.01	0.00	0.01	0.00
S Tati	(f)šupāna	cupun	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	aćru	arsa	ćr	rs	0.00	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	aćru	arsa	pro	no pro	0.12	0.23	0.23	0.17	0.05	0.12	0.03	0.01	0.02	0.01
S Tati	aćᵘa	asb	ćᵘ	sp	0.58	0.11	0.11	0.08	0.02	0.06	0.01	0.01	0.01	0.01
S Tati	aršа	xers	rs	rs	0.53	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	aršа	xers	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	barda	bila	rd	l	0.01	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	barda	bar(r)a	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	caθu̯ćat	cel	ć	l	0.00	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	ćᵘaka	esbā	ćᵘ	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	ćᵘišа	espeja	ćᵘ	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	ćᵘišа	esbeja	ćᵘ	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	ćiāu̯a	siā	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	daća	dah	ć	h	0.54	0.12	0.12	0.09	0.03	0.06	0.02	0.01	0.01	0.01
S Tati	g(a)rna-ka	galle	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	hiǰu̯āna	zōn	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	krmi	kelma	r	l	0.58	0.11	0.11	0.08	0.02	0.06	0.01	0.01	0.01	0.01
S Tati	laupāća	luās	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	laupāća	luās	l	l	0.50	0.13	0.13	0.10	0.03	0.07	0.02	0.01	0.01	0.01
S Tati	mušti	mošta	št	št	0.39	0.16	0.16	0.12	0.04	0.08	0.02	0.01	0.01	0.01
S Tati	parna	parr	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	puθra	fer	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	puθra	pur	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯ahāra	behār	u̯	b	0.00	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	u̯ićati	vist	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯(a)rda	vela	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯(a)rda	vela	rd	l	0.58	0.11	0.11	0.08	0.02	0.06	0.01	0.01	0.01	0.01
S Tati	u̯āf	bev	u̯	b	0.27	0.20	0.20	0.15	0.04	0.08	0.03	0.01	0.02	0.01
S Tati	u̯afra	vara	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯afra	vara	meta	unclear	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯ahuni	xemn	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯aina	vin	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯aita	via	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯anjecaka	viškinj	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯ār	vār	u̯	w	0.89	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯arka	varga	u̯	w	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯arka	varga	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯art-	gard	u̯	g	0.01	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	u̯āta	vā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯āǰi(i̯aka?)	vāz	u̯	w	0.95	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯idau̯ā	vivia	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯rka	varg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯ršnaka	gošna	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	u̯ršnaka	gošna	rs	š	0.00	0.26	0.26	0.19	0.06	0.14	0.03	0.01	0.02	0.02
S Tati	u̯iāna	yon	u̯	y	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Tati	ǰrd	del	rd	l	0.54	0.12	0.12	0.09	0.02	0.06	0.02	0.01	0.01	0.01
Awromani	aćᵘa	asp	ćᵘ	sp	0.83	0.04	0.06	0.03	0.01	0.00	0.01	0.00	0.01	0.00
Awromani	caθu̯ćat	cēl	rć	l	0.00	0.23	0.37	0.18	0.06	0.03	0.05	0.03	0.03	0.02
Awromani	daća	da	ć	h	0.81	0.04	0.07	0.04	0.01	0.01	0.01	0.00	0.01	0.00
Awromani	hiǰu̯āna	zawān	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	laupāća	rowās	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	laupāća	rowās	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	laupāća	rubā	ć	h	0.74	0.06	0.10	0.05	0.02	0.01	0.01	0.01	0.01	0.01
Awromani	laupāća	rubā	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	pṛtu	pord	rt	rt	0.97	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	rāšta	ras	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯ićati	bis(t)	u̯	b	0.01	0.22	0.37	0.18	0.06	0.03	0.05	0.03	0.03	0.02
Awromani	u̯ahuni	hūn	u̯	h	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯ahuni	wun	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯aiša	wēšā	u̯	w	0.83	0.04	0.06	0.03	0.01	0.00	0.01	0.00	0.01	0.00
Awromani	u̯ār	waran	u̯	w	0.97	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯arna-ka	wærā	u̯	w	0.96	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯arna-ka	wærā	rn	(r)r	0.01	0.22	0.37	0.18	0.06	0.03	0.05	0.03	0.03	0.02
Awromani	u̯rka	wærg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Awromani	u̯iāna	gyān	u̯	g	0.64	0.08	0.13	0.07	0.02	0.01	0.02	0.01	0.01	0.01
Bakhtiyari	u̯ahuni	hĩn	u̯	h	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯arna-ka	varra	u̯	w	0.98	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯arna-ka	varra	rn	(r)r	0.02	0.21	0.26	0.15	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	u̯idāna	bōūn	u̯	b	0.67	0.08	0.07	0.05	0.03	0.07	0.01	0.01	0.01	0.01
Mamasani Hadank	spṛzan	usbul	rʃ	l	0.43	0.16	0.18	0.11	0.04	0.02	0.02	0.02	0.01	0.01
Mamasani Hadank	u̯anjecaka	bangisht	u̯	b	0.05	0.26	0.30	0.19	0.07	0.03	0.03	0.03	0.02	0.02
Mamasani Hadank	u̯arāǰa	burāz	u̯	b	0.03	0.27	0.30	0.19	0.08	0.03	0.03	0.03	0.02	0.02
Mamasani Hadank	u̯arna-ka	varra	u̯	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mamasani Hadank	u̯arna-ka	varra	rn	(r)r	0.01	0.27	0.31	0.20	0.08	0.03	0.03	0.03	0.02	0.02

Larestani	aćru	xars	ćr	rs	0.02	0.27	0.30	0.18	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	aćru	xars	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	caθurćat	cel	rć	l	0.01	0.27	0.30	0.18	0.04	0.13	0.03	0.02	0.02	0.02
Larestani	ćuaka	sag	ću	s	0.02	0.27	0.30	0.18	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	ćuiša	‘eš	ću	s	0.81	0.05	0.06	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	ćiāua	siyā	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	daća	da	ć	h	0.83	0.05	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	hijuāna	zabu	ju	zb	0.01	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	krmi	kerm	r	r	0.85	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	mušti	mošt	št	št	0.73	0.07	0.08	0.05	0.01	0.03	0.01	0.01	0.00	0.00
Larestani	parna	fal	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	prtu	pol	rt	l	0.73	0.08	0.07	0.05	0.01	0.04	0.01	0.01	0.01	0.00
Larestani	puθra	pos	tr	s	0.71	0.08	0.09	0.05	0.01	0.04	0.01	0.01	0.01	0.00
Larestani	rāšta	rāss	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uahāra	ba’ār	u	b	0.02	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	uićati	biss	u	b	0.01	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	u(a)rda	gol	u	g	0.02	0.27	0.30	0.18	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	u(a)rda	gol	rd	l	0.85	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	uaćja	beš	ći	š	0.85	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	uaćja	beš	u	b	0.04	0.26	0.29	0.17	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	uāf	bāfta	u	b	0.60	0.11	0.13	0.07	0.02	0.04	0.01	0.01	0.01	0.01
Larestani	uafra	vafr	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uafra	vafr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uafra	barf	u	b	0.79	0.06	0.06	0.04	0.01	0.03	0.01	0.00	0.00	0.00
Larestani	uafra	barf	meta	meta	0.01	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	uahuni	xūn	u	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uaina	binada	u	b	0.03	0.26	0.30	0.17	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	uanjecaka	biji	u	b	0.08	0.25	0.28	0.16	0.03	0.12	0.02	0.02	0.02	0.01
Larestani	uār	bareda	u	b	0.79	0.05	0.07	0.04	0.01	0.03	0.01	0.00	0.00	0.00
Larestani	uarka	barg	u	b	0.79	0.06	0.05	0.04	0.01	0.03	0.01	0.00	0.00	0.00
Larestani	uarka	barg	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uarna-ka	barra	u	b	0.80	0.06	0.05	0.04	0.01	0.03	0.01	0.00	0.00	0.00
Larestani	uarna-ka	barra	rn	(r)r	0.02	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	uat-caka	bec	u	b	0.68	0.09	0.10	0.06	0.01	0.04	0.01	0.01	0.01	0.00
Larestani	uata	bad	u	b	0.03	0.26	0.30	0.17	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	uāta	bād	u	b	0.71	0.08	0.09	0.05	0.01	0.04	0.01	0.01	0.01	0.00
Larestani	uāji(iaka?)	bāzī	u	b	0.60	0.11	0.13	0.06	0.02	0.05	0.01	0.01	0.01	0.01
Larestani	urinji	berenj	u	b	0.85	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	uršnaka	gošna	u	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	uršnaka	gošna	rs	š	0.01	0.27	0.30	0.18	0.04	0.13	0.02	0.02	0.02	0.02
Larestani	urtka	gordakü	u	g	0.02	0.27	0.30	0.18	0.04	0.12	0.02	0.02	0.02	0.02
Larestani	uiāka	jā	u	j	0.85	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Larestani	uiāna	ju	u	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Larestani	jrd	del	rd	l	0.83	0.05	0.05	0.03	0.00	0.02	0.00	0.00	0.00	0.00
Sivandi	uāf	bāftan	u	b	0.24	0.19	0.24	0.15	0.04	0.09	0.01	0.01	0.01	0.01
Sivandi	barda	bīl	rd	l	0.01	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	brzant	boland	rʃ	l	0.43	0.14	0.17	0.11	0.03	0.09	0.01	0.01	0.01	0.01
Sivandi	uanjecaka	bonješk	u	b	0.02	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	jrd	del	rd	l	0.51	0.12	0.15	0.09	0.02	0.08	0.01	0.01	0.01	0.01
Sivandi	darn	derr	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	ćuaka	espe	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	sprzan	espel	rʃ	l	0.23	0.19	0.23	0.15	0.04	0.12	0.01	0.01	0.01	0.01
Sivandi	uršnaka	feše	u	f	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uršnaka	feše	rs	š	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	uahuni	fīn	u	f	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	u(a)rda	gol	u	g	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	u(a)rda	gol	rd	l	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
Sivandi	uināća	gonā	ć	h	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
Sivandi	uināća	gonā	u	g	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	ćuiša	peš	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	kaćiapa	kalapošt	ći	š	0.01	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	krmi	kerm	r	r	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
Sivandi	uagza	maze	u	m	0.89	0.00	0.00	0.00	0.03	0.08	0.00	0.00	0.00	0.00
Sivandi	prtu	perd	rt	rt	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	parna	parr	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	rāšta	rās	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	laupāća	rūbā	ć	h	0.42	0.14	0.18	0.11	0.03	0.09	0.01	0.01	0.01	0.01
Sivandi	laupāća	rūbā	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	ćiāua	siya	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	aćua	usūr	ću	s	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
Sivandi	uahāra	vāhār	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uarka	valk	u	w	0.86	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uarka	valk	r	l	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	uār	vārān	u	w	0.88	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uafra	varf	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uafra	varf	meta	meta	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	uarna-ka	vare	u	w	0.87	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uarna-ka	vare	rn	(r)r	0.00	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	uaita	vi	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uaćja	vīštar	ći	š	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01
Sivandi	uaćja	vīštar	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uāta	voy	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	hijuāna	zuwān	ju	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	g(a)rna-ka	galle	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	(f)šupāna	cūpān	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	uahia	va	u	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	urinji	berc	u	b	0.54	0.11	0.14	0.09	0.02	0.07	0.01	0.01	0.01	0.01

Sivandi	u̯rtka	velk	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	u̯art-	gir	u̯	g	0.01	0.24	0.30	0.19	0.05	0.15	0.02	0.02	0.02	0.02
Sivandi	u̯iāna	gyān	u̯	g	0.31	0.17	0.21	0.13	0.03	0.11	0.01	0.01	0.01	0.01
Sivandi	u̯aicīna?	višne	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sivandi	u̯ićati	vīs	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	aćru	hars	ćr	rs	0.03	0.21	0.26	0.14	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	aćru	hars	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	aćua	asp	ću	sp	0.88	0.03	0.03	0.02	0.01	0.02	0.00	0.00	0.00	0.00
Bakhtiyari	arša	xers	rs	rs	0.86	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	arša	xers	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	caθurćat	cel	rć	l	0.01	0.21	0.27	0.15	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	ćuaka	say	ću	s	0.02	0.21	0.26	0.14	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	ćuiša	šeš	ću	s	0.85	0.03	0.04	0.02	0.01	0.03	0.00	0.00	0.00	0.00
Bakhtiyari	ćiāua	šah	ći	š	0.02	0.21	0.27	0.15	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	daća	dah	ć	h	0.87	0.03	0.04	0.02	0.01	0.03	0.00	0.00	0.00	0.00
Bakhtiyari	darn	der	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	g(a)rna-ka	gale	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	híjuāna	zovn	ju	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	kaćiapa	kāse	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	laupāća	ruvā	ć	h	0.82	0.04	0.05	0.03	0.02	0.04	0.00	0.00	0.00	0.00
Bakhtiyari	laupāća	ruvā	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	mušti	most	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	parna	par	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	prtu	pol	rt	l	0.78	0.05	0.05	0.04	0.02	0.05	0.01	0.00	0.01	0.00
Bakhtiyari	rāšta	rāst	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯ahāra	bohār	u̯	b	0.02	0.21	0.26	0.14	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	u̯ićati	bist	u̯	b	0.02	0.21	0.27	0.15	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	u̯(a)rda	gol	u̯	g	0.03	0.21	0.26	0.14	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	u̯(a)rda	gol	rd	l	0.88	0.03	0.03	0.02	0.01	0.02	0.00	0.00	0.00	0.00
Bakhtiyari	u̯āf	baften	u̯	b	0.67	0.07	0.09	0.05	0.03	0.05	0.01	0.01	0.01	0.00
Bakhtiyari	skar	šekāl	r	l	0.12	0.19	0.24	0.13	0.08	0.18	0.02	0.02	0.02	0.01
Bakhtiyari	u̯agza	baq??	u̯	b	0.80	0.05	0.06	0.03	0.01	0.03	0.01	0.00	0.00	0.00
Bakhtiyari	u̯ahuni	hin	u̯	h	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯aita	beyd	u̯	b	0.30	0.15	0.19	0.10	0.06	0.14	0.02	0.01	0.01	0.01
Bakhtiyari	u̯anjecaka	bengešt	u̯	b	0.10	0.19	0.24	0.13	0.08	0.18	0.02	0.02	0.02	0.01
Bakhtiyari	u̯ār	bārovn	u̯	b	0.83	0.03	0.05	0.03	0.02	0.04	0.00	0.00	0.00	0.00
Bakhtiyari	u̯arāja	gorāz	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯arna-ka	bare	u̯	b	0.84	0.04	0.04	0.03	0.02	0.04	0.00	0.00	0.00	0.00
Bakhtiyari	u̯arna-ka	bare	rn	(r)r	0.02	0.21	0.26	0.15	0.09	0.20	0.03	0.02	0.02	0.01
Bakhtiyari	u̯at-caka	bace	u̯	b	0.74	0.06	0.07	0.04	0.02	0.05	0.01	0.00	0.01	0.00
Bakhtiyari	u̯āta	bād	u̯	b	0.76	0.05	0.07	0.04	0.02	0.05	0.01	0.00	0.01	0.00
Bakhtiyari	u̯ājí(ịaka?)	bāzi	u̯	b	0.66	0.08	0.10	0.04	0.03	0.07	0.01	0.01	0.01	0.00
Bakhtiyari	u̯ajra	gorz	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯icāra	gozar	u̯	g	0.73	0.06	0.07	0.04	0.02	0.05	0.01	0.00	0.01	0.00
Bakhtiyari	u̯iiapāna	biyāvovn	u̯	b	0.06	0.20	0.25	0.14	0.08	0.19	0.02	0.02	0.02	0.01
Bakhtiyari	u̯rinji	berenj	u̯	b	0.88	0.03	0.03	0.02	0.01	0.02	0.00	0.00	0.00	0.00
Bakhtiyari	u̯rka	gorg	u̯	g	0.51	0.10	0.13	0.07	0.04	0.10	0.01	0.01	0.01	0.01
Bakhtiyari	u̯ršnaka	gosne	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	u̯rtka	gordāle	u̯	g	0.03	0.21	0.26	0.14	0.09	0.19	0.03	0.02	0.02	0.01
Bakhtiyari	u̯iāka	jā	u̯	j	0.88	0.03	0.03	0.02	0.01	0.02	0.00	0.00	0.00	0.00
Bakhtiyari	u̯iāna	jovn	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bakhtiyari	jrd	del	rd	l	0.87	0.03	0.04	0.02	0.01	0.03	0.00	0.00	0.00	0.00
Zazaki	(f)šupāna	šiwāne	š	š	0.22	0.13	0.36	0.11	0.10	0.02	0.02	0.01	0.01	0.01
Zazaki	aćru	hesri	ćr	sr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	aćru	hesri	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	aćua	ospore	ću	sp	0.69	0.05	0.15	0.04	0.04	0.01	0.01	0.00	0.00	0.00
Zazaki	arša	heš	rs	š	0.74	0.06	0.05	0.05	0.05	0.01	0.01	0.01	0.01	0.01
Zazaki	arša	heš	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	caθurćat	cewres	rć	rć	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	ćuiša	ešpiį	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	ćiāua	siyā	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	darn	dirā	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	híjuāna	ziwān	ju	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	kaćiapa	kese	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	laupāća	lü	ć	h	0.57	0.07	0.20	0.06	0.06	0.01	0.01	0.01	0.01	0.01
Zazaki	laupāća	lü	l	l	0.62	0.06	0.18	0.05	0.05	0.01	0.01	0.01	0.00	0.01
Zazaki	parna	pur	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	prtu	pird	rt	rt	0.91	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	rāšta	rāšt	št	št	0.00	0.16	0.47	0.14	0.13	0.02	0.03	0.02	0.01	0.01
Zazaki	u̯ićati	vīst	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	daća	des	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯(a)rda	gul	u̯	g	0.01	0.16	0.47	0.14	0.13	0.02	0.03	0.02	0.01	0.01
Zazaki	u̯(a)rda	gul	rd	l	0.69	0.05	0.15	0.04	0.04	0.01	0.01	0.00	0.00	0.00
Zazaki	u̯afra	vewr	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯afra	vewr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯ahuni	gūnī	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯ār	vārān	u̯	w	0.95	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯ara	virār	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯arka	velg	u̯	w	0.88	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯arka	velg	r	l	0.00	0.16	0.47	0.14	0.13	0.02	0.03	0.02	0.01	0.01
Zazaki	u̯arna-ka	vorek	u̯	w	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯arna-ka	vorek	rn	(r)r	0.01	0.16	0.47	0.14	0.13	0.02	0.03	0.02	0.01	0.01
Zazaki	u̯āta	vā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯idau̯ā	viyā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯ršnaka	veyšān	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zazaki	u̯ršnaka	veyšān	rs	š	0.01	0.16	0.47	0.14	0.13	0.02	0.03	0.02	0.01	0.01
Zazaki	u̯iāka	jā	u̯	j	0.69	0.05	0.15	0.04	0.04	0.01	0.01	0.00	0.00	0.00

Zazaki	u̯iāna	gān	u̯	g	0.45	0.09	0.26	0.08	0.07	0.01	0.02	0.01	0.01	0.01
Zazaki	ɟ́rd	zeři	rd	rd	0.90	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00
Talysh	aćua	asp	ću	sp	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	caθu̯ŕćat	cül	rć	l	0.00	0.22	0.29	0.22	0.06	0.15	0.02	0.02	0.02	0.01
Talysh	ćuaka	səpa	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	rāšta	rost	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ićati	vist	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯aiša	viša	u̯	w	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	ɟ́rd	del	rd	l	0.59	0.09	0.12	0.09	0.02	0.06	0.01	0.01	0.01	0.00
Talysh	arša	xərs	rs	rs	0.55	0.00	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	arša	xərs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	barda	bur	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	ćuaka	əspa	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	ćuiša	əsbəj	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	ćiāua	siā	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	híjuāna	zabōn	ju̯	zb	0.00	0.22	0.29	0.22	0.06	0.14	0.02	0.02	0.02	0.01
Talysh	híjuāna	zuvun	ju̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	híjuāna	zymyn	ju̯	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	laupāća	rəvus	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	laupāća	rəvus	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	pṛtu	pard	rt	rt	0.93	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	rāšta	rust	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ahāra	bahār	u̯	b	0.00	0.22	0.29	0.22	0.06	0.14	0.02	0.02	0.02	0.01
Talysh	u̯(a)rda	vəl	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯(a)rda	vəl	rd	l	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	u̯(a)rda	gul	u̯	g	0.01	0.22	0.29	0.22	0.06	0.14	0.02	0.02	0.02	0.01
Talysh	u̯(a)rda	gul	rd	l	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	u̯aćia	vištatar	ći̯	š	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	u̯aćia	vištatar	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯āf	bof	u̯	b	0.31	0.16	0.21	0.16	0.04	0.08	0.02	0.01	0.01	0.01
Talysh	u̯afra	var	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯afra	var	meta	unclear	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯aina	vin	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯aita	vi	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ār	vār	u̯	w	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯arna-ka	vaṛā	u̯	w	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯art-	gard	u̯	g	0.01	0.22	0.29	0.22	0.06	0.14	0.02	0.02	0.02	0.01
Talysh	u̯āta	vu	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯āɟi(ia̯ka?)	bāzi	u̯	b	0.31	0.16	0.21	0.12	0.04	0.11	0.02	0.01	0.01	0.01
Talysh	u̯idau̯ā	viva	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯rinji	barz	u̯	b	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	u̯rka	varg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ršnaka	veši	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	u̯ršnaka	veši	rs	š	0.00	0.22	0.29	0.22	0.06	0.14	0.02	0.02	0.02	0.01
Talysh	u̯iāka	jā	u̯	j	0.62	0.08	0.11	0.08	0.02	0.05	0.01	0.01	0.01	0.00
Talysh	u̯iāna	jān	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Talysh	ɟ́rd	dəl	rd	l	0.59	0.09	0.12	0.09	0.02	0.06	0.01	0.01	0.01	0.00
Balochi Marw	(f)šupāna	šwānág	š	š	0.58	0.08	0.18	0.08	0.03	0.01	0.01	0.01	0.01	0.01
Balochi Marw	ćiāua	siyā	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	aćua	asp	ću̯	sp	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	arša	xirs	rs	rs	0.84	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	arša	xirs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	brzant	burz	rɟ́	rɟ́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	barda	bēl	rd	l	0.06	0.17	0.40	0.18	0.07	0.02	0.03	0.02	0.02	0.02
Balochi Marw	barzād	balad	rɟ́	l	0.89	0.02	0.05	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	caθu̯ŕćat	cil	rć	l	0.01	0.18	0.42	0.19	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	daća	da	ć	h	0.90	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	darn	dir	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	g(a)rna-ka	gal	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	híjuāna	zubān	ju̯	zb	0.03	0.18	0.42	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	krmi	kirm	r	r	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	laupāća	rūbā	ć	h	0.86	0.03	0.06	0.03	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	laupāća	rūbā	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	parna	parr	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	rāšta	rāstam	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯ićati	bīst	u̯	b	0.02	0.18	0.42	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	u̯ićati	gīst	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯(a)rda	gul	u̯	g	0.04	0.18	0.41	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	u̯(a)rda	gul	rd	l	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯āta	gwāt	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯āɟi(n)	bāz	u̯	b	0.28	0.13	0.31	0.13	0.06	0.02	0.02	0.02	0.01	0.02
Balochi Marw	u̯āɟi(ia̯ka?)	gwāzī	u̯	g	0.99	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯ršnaka	gušnag	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯ršnaka	gušnag	rs	š	0.02	0.18	0.42	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	u̯rka	gurk	u̯	g	0.60	0.08	0.17	0.08	0.03	0.01	0.01	0.01	0.01	0.01
Balochi Marw	u̯aćia	geštir	ći̯	š	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯aćia	geštir	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯aɟra	burz	u̯	b	0.08	0.17	0.39	0.17	0.07	0.02	0.03	0.02	0.02	0.02
Balochi Marw	u̯aɟra	γurz ‘club, spear’	u̯	γ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯ahi̯a	battir	u̯	b	0.61	0.07	0.17	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Balochi Marw	u̯ahi̯a	getir ‘better’	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯aina	gind	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯aita	bed	u̯	b	0.37	0.12	0.27	0.12	0.05	0.02	0.02	0.02	0.01	0.01
Balochi Marw	u̯arka	balg	u̯	b	0.88	0.03	0.04	0.03	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	u̯arka	balg	r	l	0.01	0.18	0.42	0.19	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	u̯arna-ka	gwarag	u̯	g	0.97	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Balochi Marw	ɻarna-ka	gwarag	rn	(r)r	0.03	0.18	0.42	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	ɻart-	gardišt	ɻ	g	0.05	0.18	0.41	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	ɻat-caka	bacak	ɻ	b	0.80	0.04	0.09	0.04	0.02	0.00	0.01	0.00	0.00	0.00
Balochi Marw	ɻata	bad	ɻ	b	0.05	0.18	0.41	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	ɻidāna	gidan	ɻ	g	0.99	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	ɻināća	gunā	ć	h	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	ɻināća	gunā	ɻ	g	0.04	0.18	0.41	0.18	0.08	0.02	0.03	0.02	0.02	0.02
Balochi Marw	ɻiiapāna	giābān	ɻ	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Marw	ɻiāka	jā	ɻ	j	0.91	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Balochi Marw	ɻrća-ka-	gulg	ɻ	g	0.15	0.16	0.36	0.16	0.07	0.02	0.02	0.02	0.02	0.02
Balochi Marw	ɻrća-ka-	gulg	rć	l	0.29	0.13	0.30	0.13	0.06	0.02	0.02	0.02	0.01	0.02
Balochi Marw	ɻrd	dil	rd	l	0.90	0.02	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	(f)šupāna	shūān	š	š	0.21	0.12	0.43	0.08	0.09	0.05	0.01	0.01	0.01	0.01
Kurmanji Soane	ćuaka	seh	ću	s	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ćuaka	seg	ću	s	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ać̣ua	hasp	ću	sp	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	arša	hirch	rs	rch	0.41	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	arša	hirch	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	arša	wirch	rs	rch	0.41	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	arša	wirch	pro	no pro	0.00	0.15	0.54	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	arša	birch	rs	rch	0.41	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	arša	birch	pro	no pro	0.00	0.15	0.54	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	brzant	barz	rʃ	rʃ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	brzant	bilin	rʃ	l	0.56	0.07	0.24	0.04	0.05	0.03	0.00	0.00	0.00	0.00
Kurmanji Soane	barda	bīał	rd	l	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	barda	bair	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	barzād	berzī	rʃ	rʃ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	barzād	belinī	rʃ	l	0.61	0.06	0.21	0.04	0.05	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	caθurćat	chil	rć	l	0.00	0.15	0.54	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	daća	dah	ć	h	0.64	0.06	0.20	0.04	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	darn	derīān	rn	(r)n	0.02	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	g(a)rna-ka	gārān	rn	(r)n	0.02	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	híjuāna	zwān	ʃu	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	híjuāna	hizwān	ʃu	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	híjuāna	zemān	ʃu	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	híjuāna	azmān	ʃu	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	krmi	kerma	r	r	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	kać̣apa	kīsal	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	laupāća	riwī	ć	h	0.55	0.07	0.24	0.04	0.05	0.03	0.00	0.00	0.00	0.00
Kurmanji Soane	laupāća	riwī	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	mušti	must	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	mušti	misht	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	pṛtu	prt	rt	rt	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	pṛtu	prr	rt	rt	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	parna	par	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	rāšta	rāst	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻahāra	bahār	ɻ	b	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻićati	bīs	ɻ	b	0.00	0.15	0.54	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻ(a)rda	gul	ɻ	g	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻ(a)rda	gul	rd	l	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻār	bārān	ɻ	b	0.56	0.05	0.24	0.05	0.05	0.03	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻāta	bā	ɻ	b	0.46	0.08	0.29	0.05	0.06	0.03	0.00	0.01	0.01	0.00
Kurmanji Soane	ɻāǰi(n)	bāz	ɻ	b	0.07	0.14	0.50	0.09	0.11	0.05	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻāǰi(ɻaka?)	bāzīn	ɻ	b	0.34	0.10	0.36	0.05	0.08	0.04	0.00	0.01	0.01	0.01
Kurmanji Soane	ɻršnaka	birsī	ɻ	b	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻršnaka	birsī	rs	rs	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻrka	gurg	ɻ	g	0.22	0.12	0.42	0.08	0.09	0.04	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻrtka	gurchī	ɻ	g	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻafra	bafr	ɻ	b	0.57	0.07	0.23	0.04	0.05	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻafra	bafr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻafra	wafr	ɻ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻafra	wafr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻagza	boq	ɻ	b	0.51	0.08	0.27	0.05	0.04	0.02	0.00	0.01	0.01	0.00
Kurmanji Soane	ɻahuni	khūn	ɻ	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻaicīna?	bezhenk	ɻ	b	0.02	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻaita	bī	ɻ	b	0.10	0.14	0.48	0.09	0.10	0.05	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻarāǰa	burāz	ɻ	b	0.02	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻarāǰa	vurāz	ɻ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻarka	barak	ɻ	b	0.58	0.07	0.19	0.05	0.06	0.03	0.00	0.01	0.00	0.00
Kurmanji Soane	ɻarka	barak	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻarna-ka	berkh	ɻ	b	0.60	0.07	0.19	0.05	0.05	0.03	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻarna-ka	berkh	rn	(r)r	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻart-	war garrīān	ɻ	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻidaɻā	bī	ɻ	b	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻināća	gunā	ć	h	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻināća	gunā	ɻ	g	0.01	0.15	0.53	0.10	0.11	0.06	0.01	0.01	0.01	0.01
Kurmanji Soane	ɻiāka	gah	ɻ	g	0.67	0.05	0.18	0.03	0.04	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻiāna	jān	ɻ	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻrd	dił	rd	l	0.64	0.06	0.20	0.04	0.03	0.02	0.00	0.00	0.00	0.00
Kurmanji Soane	ɻrd	zer	rd	rd	0.91	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00
Vafsi	(f)šupāna	cupan	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	ćuaka	æsbæ	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	ćiāua	sia(h)	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	brzant	bolænd	rʃ	l	0.52	0.16	0.15	0.09	0.05	0.01	0.01	0.01	0.01	0.01
Vafsi	daća	dæ(h)	ć	h	0.61	0.13	0.12	0.07	0.04	0.01	0.01	0.01	0.01	0.00
Vafsi	híjuāna	zuan	ʃu	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Vafsi	laupāća	luas	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	laupāća	luas	l	l	0.56	0.14	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vafsi	mušti	méšdæ	št	št	0.45	0.18	0.17	0.10	0.05	0.01	0.01	0.01	0.01	0.01
Vafsi	parna	pær(r)	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	rāšta	rasd	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯ahāra	vaar	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯ār	varan	u̯	w	0.89	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯āta	va	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯ršnaka	gosnæ	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯rka	værg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯aća	bisdær	ći̯	s	0.64	0.12	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.00
Vafsi	u̯aća	bisdær	u̯	b	0.01	0.32	0.31	0.18	0.09	0.02	0.02	0.02	0.02	0.01
Vafsi	u̯ajr(a)ka	bozorgevar	u̯	b	0.33	0.18	0.23	0.13	0.07	0.01	0.01	0.02	0.01	0.01
Vafsi	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯aiša	væši	u̯	w	0.64	0.12	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.00
Vafsi	u̯aina	vin	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯art-	gerd	u̯	g	0.01	0.32	0.31	0.18	0.09	0.02	0.02	0.02	0.02	0.01
Vafsi	u̯idau̯ā	víæ	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	u̯jāna	gan	u̯	g	0.40	0.20	0.19	0.11	0.06	0.01	0.01	0.01	0.01	0.01
Vafsi	u̯jāna	jan	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vafsi	ǰṛd	del	rd	l	0.61	0.13	0.12	0.07	0.03	0.01	0.01	0.01	0.01	0.00
Balochi Rakhshani	(f)šupāna	šwpank	š	š	0.43	0.13	0.16	0.09	0.07	0.08	0.01	0.01	0.01	0.01
Balochi Rakhshani	ćuaka	səg	ću̯	s	0.02	0.22	0.28	0.16	0.11	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	ćiāu̯a	sya[h]	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	aćru	ərs	ćr	rs	0.02	0.22	0.28	0.16	0.11	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	aćru	ərs	pro	no pro	0.37	0.14	0.18	0.10	0.07	0.08	0.01	0.01	0.01	0.01
Balochi Rakhshani	aću̯a	əsp	ću̯	sp	0.85	0.03	0.04	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	brzant	bwrz	rǰ	rǰ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	barzād	balad	rǰ	l	0.82	0.04	0.05	0.03	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	caθu̯rćat	čyll	rć	l	0.01	0.22	0.29	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	daća	də	ć	h	0.84	0.04	0.05	0.03	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	darn	dyrrəg	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	g(a)rna-ka	gəl	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	híju̯āna	zwban	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	laupāća	roba	ć	h	0.78	0.05	0.06	0.04	0.03	0.03	0.00	0.00	0.00	0.00
Balochi Rakhshani	laupāća	roba	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	prtu	pol	rt	l	0.74	0.06	0.06	0.05	0.03	0.04	0.01	0.01	0.01	0.01
Balochi Rakhshani	parna	pənn	rn	(r)n	0.01	0.22	0.29	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	rāšta	rast	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ahāra	bə[h]ar	u̯	b	0.02	0.22	0.28	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯ićati	bist	u̯	b	0.01	0.22	0.28	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯āf	gvəpəg	u̯	g	0.99	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ār	gvaran	u̯	g	0.98	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯āta	gvat	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯āǰi(ia̯ka?)	gvazi	u̯	g	0.99	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ća-ka-	gvaləg	u̯	g	0.09	0.20	0.26	0.15	0.11	0.12	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯ća-ka-	gvaləg	rć	l	0.19	0.18	0.23	0.13	0.09	0.11	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯ršnaka	gwšnəg	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯rka	gwrk	u̯	g	0.45	0.12	0.16	0.09	0.06	0.07	0.01	0.01	0.01	0.01
Balochi Rakhshani	u̯aća	geš	ći̯	š	0.85	0.03	0.04	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯aća	geš	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯afra	bərp	u̯	b	0.80	0.04	0.06	0.03	0.02	0.03	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯afra	bərp	meta	meta	0.01	0.22	0.28	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯ahuni	hon	u̯	h	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ahia	vəšš	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯aina	gyndəg	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ara	gvər	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯arna-ka	gvərəg	u̯	g	0.97	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯arna-ka	gvərəg	rn	(r)r	0.02	0.22	0.28	0.16	0.12	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯ata	bəd	u̯	b	0.03	0.22	0.28	0.16	0.11	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯icāra	gvəzəg	u̯	g	0.68	0.07	0.09	0.05	0.04	0.04	0.01	0.01	0.01	0.01
Balochi Rakhshani	u̯idāna	gydan	u̯	g	0.98	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ināća	gwna[h]	ć	h	0.85	0.03	0.04	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯ināća	gwna[h]	u̯	g	0.02	0.22	0.28	0.16	0.11	0.13	0.02	0.02	0.02	0.02
Balochi Rakhshani	u̯rinji	byrynj	u̯	b	0.85	0.03	0.04	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	u̯jāka	ja	u̯	j	0.85	0.03	0.04	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	ǰṛd	dyl	rd	l	0.84	0.04	0.05	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Balochi Rakhshani	ǰṛd	zyrd	rd	rd	0.97	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	(f)šupāna	çupon	š	š	0.29	0.15	0.22	0.12	0.05	0.11	0.02	0.02	0.01	0.01
Judeo-Tati	ćuaka	seg	ću̯	s	0.01	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	ćiāu̯a	siyeh	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	aćru	ʕärs	ćr	rs	0.01	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	aćru	ʕärs	pro	no pro	0.24	0.16	0.24	0.13	0.05	0.11	0.02	0.02	0.01	0.01
Judeo-Tati	aću̯a	ʕäsb	ću̯	sp	0.76	0.05	0.08	0.04	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	arša	xürs	rs	rs	0.69	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	arša	xürs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	barda	bil	rd	l	0.02	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	barzād	bala	rǰ	l	0.71	0.06	0.09	0.05	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	caθu̯rćat	çül	rć	l	0.00	0.22	0.32	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	daća	deh	ć	h	0.74	0.06	0.08	0.05	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	darn	durra	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	g(a)rna-ka	gele	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	gau-ću̯anta-	gusbend	ću̯	sp	0.76	0.05	0.08	0.04	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	híju̯āna	zuhun	ǰu̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	krmi	kürm	r	r	0.76	0.05	0.08	0.04	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	mušti	mušd	št	št	0.60	0.09	0.13	0.07	0.03	0.06	0.01	0.01	0.01	0.01
Judeo-Tati	rāšta	rasd	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯ahāra	vasal	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Judeo-Tati	u̯ićati	bisd	u̯	b	0.01	0.21	0.32	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯(a)rda	gül	u̯	g	0.01	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯(a)rda	gül	rd	l	0.76	0.05	0.08	0.04	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	u̯āf	bofde	u̯	b	0.46	0.12	0.18	0.10	0.04	0.06	0.01	0.01	0.01	0.01
Judeo-Tati	u̯ār	voruș	u̯	w	0.96	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯āta	vor	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯āǰí(i̯aka?)	vozire	u̯	w	0.98	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯uršnaka	gisne	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯urka	gürg	u̯	g	0.31	0.15	0.22	0.12	0.04	0.10	0.02	0.02	0.01	0.01
Judeo-Tati	u̯afra	verf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯afra	verf	meta	meta	0.01	0.21	0.32	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯aiša	vișe	u̯	w	0.76	0.05	0.08	0.04	0.02	0.04	0.01	0.01	0.00	0.00
Judeo-Tati	u̯aina	vini	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯anjecaka	bäʕäveçe	u̯	b	0.05	0.21	0.30	0.17	0.06	0.14	0.02	0.02	0.02	0.01
Judeo-Tati	u̯arka	velg	u̯	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯arka	velg	r	l	0.00	0.22	0.32	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯arna-ka	vere	u̯	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judeo-Tati	u̯arna-ka	vere	rn	(r)r	0.01	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯ata	bad	u̯	b	0.01	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	u̯idāna	bine	u̯	b	0.47	0.12	0.14	0.10	0.04	0.09	0.01	0.01	0.01	0.01
Judeo-Tati	u̯idaụā	biy	u̯	b	0.02	0.21	0.31	0.17	0.06	0.15	0.02	0.02	0.02	0.01
Judeo-Tati	ǰṛd	dül	rd	l	0.74	0.06	0.09	0.05	0.01	0.04	0.01	0.01	0.00	0.00
Sangesari	(f)šupāna	cuppwn	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	ć̣uaka	ēspē	ć̣u	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	ć̣uiša	ēspaz	ć̣u	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	ć̣i̯ā̯ụa	so	ć̣i̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	aćụa	asm	ć̣u	sp	0.53	0.10	0.22	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Sangesari	arša	xērs	rs	rs	0.34	0.00	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	arša	xērs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	arša	aš	rs	š	0.59	0.13	0.07	0.10	0.05	0.01	0.01	0.01	0.01	0.01
Sangesari	arša	aš	pro	no pro	0.00	0.21	0.46	0.17	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	bṛzant	bēlēnd	rǰ	l	0.41	0.12	0.27	0.10	0.05	0.01	0.01	0.01	0.01	0.01
Sangesari	kṛmi	kērm	r	r	0.53	0.10	0.22	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Sangesari	laupāća	rēva	ć	h	0.40	0.13	0.28	0.10	0.05	0.01	0.01	0.01	0.01	0.01
Sangesari	laupāća	rēva	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	mušti	mošt	št	št	0.34	0.14	0.31	0.11	0.05	0.01	0.01	0.01	0.01	0.01
Sangesari	parna	par	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	puθra	pur	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	spṛzan	ēspol	rǰ	l	0.22	0.16	0.36	0.13	0.06	0.01	0.01	0.01	0.01	0.01
Sangesari	u̯ahāra	vihər	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯(a)rda	gol	u̯	g	0.00	0.21	0.46	0.17	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯(a)rda	gol	rd	l	0.53	0.10	0.22	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Sangesari	u̯āf	vō	u̯	w	0.99	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Sangesari	u̯ār	vər	u̯	w	0.89	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯āta	ve	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯āǰí(n)	bəz	u̯	b	0.04	0.20	0.45	0.16	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯uršnaka	vēššwn	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯urka	vark	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯urtka	vakku	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯aćia	vēšter	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯afra	varf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯afra	varf	meta	meta	0.00	0.21	0.46	0.17	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯ahi̯a	veytar	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯aita	bid	u̯	b	0.06	0.20	0.44	0.16	0.07	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯arāǰa	vērəz	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯arka	valg	u̯	w	0.79	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯arka	valg	r	l	0.00	0.21	0.46	0.17	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯arna-ka	varē	u̯	w	0.80	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	u̯arna-ka	varē	rn	(r)r	0.00	0.21	0.46	0.17	0.08	0.02	0.02	0.02	0.01	0.01
Sangesari	u̯at-caka	bad	u̯	b	0.30	0.15	0.33	0.12	0.06	0.01	0.01	0.01	0.01	0.01
Sangesari	u̯iāna	jwn	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sangesari	ǰṛd	dal	rd	l	0.49	0.11	0.24	0.09	0.03	0.01	0.01	0.01	0.01	0.01
S Bashkardi	ć̣uaka	sax	ć̣u	s	0.01	0.25	0.27	0.18	0.17	0.03	0.03	0.03	0.02	0.02
S Bashkardi	arša	hors	rs	rs	0.68	0.00	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	arša	hors	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	barda	bāhr	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	u̯āf	vahv	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	u̯urka	verx	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	u̯arna-ka	vark	u̯	w	0.94	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Bashkardi	u̯arna-ka	vark	rn	(r)r	0.01	0.25	0.27	0.18	0.17	0.03	0.03	0.03	0.02	0.02
S Bashkardi	u̯rinji	berenz	u̯	b	0.73	0.07	0.07	0.05	0.05	0.01	0.01	0.01	0.01	0.00
S Bashkardi	ǰṛd	der	rd	rd	0.89	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00
S Bashkardi	skar	šekāl	r	l	0.05	0.24	0.26	0.17	0.17	0.03	0.03	0.02	0.02	0.02
N Bashkardi	u̯ār	gwaron	u̯	g	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N Bashkardi	u̯agza	gwäk	u̯	g	0.98	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Yazdi VA	ć̣uaka	sva	ć̣u	s	0.01	0.25	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	aćụa	asb	ć̣u	sp	0.72	0.07	0.09	0.05	0.03	0.01	0.01	0.01	0.01	0.01
Yazdi VA	bṛzant	bilend	rǰ	l	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Yazdi VA	barda	bard(a)	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	caθu̯ṛćat	cil	rć	l	0.00	0.26	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	gau-ć̣uanta-	güsfand	ć̣u	sp	0.72	0.07	0.09	0.05	0.03	0.01	0.01	0.01	0.01	0.01
Yazdi VA	laupāća	rüwōs	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	laupāća	rüwōs	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	mušti	mušt	št	št	0.55	0.12	0.15	0.09	0.04	0.02	0.01	0.01	0.01	0.01
Yazdi VA	spṛzan	svarz	rǰ	rǰ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯(a)rda	gol	u̯	g	0.01	0.25	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02

Yazdi VA	u̯(a)rda	gol	rd	l	0.72	0.07	0.09	0.05	0.03	0.01	0.01	0.01	0.01	0.01
Yazdi VA	u̯ār	wōrōn	u̯	w	0.94	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯āta	wōd	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯r̆šnaka	wōšna	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯r̆šnaka	wōšna	rs	š	0.01	0.26	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	u̯afra	varf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯afra	varf	meta	meta	0.01	0.26	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	u̯ahi̯a	vahtèr	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯aita	vid	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯arna-ka	varra	u̯	w	0.93	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯arna-ka	varra	rn	(r)r	0.01	0.25	0.33	0.20	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	u̯art-	gèrtvūn	u̯	g	0.01	0.25	0.33	0.19	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	u̯at-caka	vacca	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯idau̯ā	vedva	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	u̯idau̯ā	bi̅va	u̯	b	0.02	0.25	0.33	0.19	0.10	0.04	0.02	0.02	0.02	0.02
Yazdi VA	u̯iāna	jün	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yazdi VA	ǰr̆d	dil	rd	l	0.70	0.08	0.10	0.06	0.02	0.01	0.01	0.01	0.01	0.01
Kumzari	ć̣uaka	sağ	ć̣u	s	0.01	0.30	0.31	0.21	0.07	0.02	0.03	0.02	0.02	0.02
Kumzari	ć̣iāu̯a	siya	ć̣i	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	ać̌ru	xars	ć̌r	rs	0.02	0.30	0.31	0.21	0.07	0.02	0.03	0.02	0.02	0.02
Kumzari	ać̣ru	xars	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	ać̣ua	asp	ć̣u	sp	0.82	0.05	0.05	0.04	0.01	0.00	0.00	0.00	0.00	0.00
Kumzari	br̆zant	blind	rǰ	l	0.75	0.08	0.08	0.05	0.02	0.01	0.01	0.01	0.00	0.00
Kumzari	caθu̯ŕ̆at	cil	rć	l	0.00	0.30	0.31	0.21	0.07	0.02	0.03	0.02	0.02	0.02
Kumzari	daća	da	ć	h	0.80	0.06	0.06	0.04	0.01	0.00	0.01	0.00	0.00	0.00
Kumzari	gau-ć̣uanta-	gŏsin	ć̣u	s	0.82	0.05	0.05	0.04	0.01	0.00	0.00	0.00	0.00	0.00
Kumzari	hijų̆āna	ẓwan	jų	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	kr̆mi	kurm	r	r	0.82	0.05	0.05	0.04	0.01	0.00	0.00	0.00	0.00	0.00
Kumzari	laupāća	rewa	ć	h	0.74	0.08	0.08	0.05	0.02	0.01	0.01	0.01	0.00	0.00
Kumzari	laupāća	rewa	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	parna	par	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	puθra	pis	tr	s	0.67	0.10	0.10	0.07	0.02	0.01	0.01	0.01	0.01	0.01
Kumzari	rāšta	rasti	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	u̯ićati	bis	u̯	b	0.01	0.30	0.31	0.21	0.07	0.02	0.03	0.02	0.02	0.02
Kumzari	u̯ār	baram	u̯	b	0.76	0.06	0.08	0.05	0.02	0.01	0.01	0.01	0.00	0.00
Kumzari	u̯ājí(i̯aka?)	baẓ	u̯	b	0.56	0.14	0.15	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Kumzari	u̯r̆šnaka	gišnag	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	u̯ahuni	xwem	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	u̯arka	barg	u̯	b	0.76	0.08	0.06	0.06	0.02	0.01	0.01	0.01	0.00	0.00
Kumzari	u̯arka	barg	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kumzari	u̯rinji	brinẓ	u̯	b	0.82	0.05	0.05	0.04	0.01	0.00	0.00	0.00	0.00	0.00
Kumzari	ǰr̆d	dil	rd	l	0.80	0.06	0.06	0.04	0.01	0.00	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯arna-ka	berx	u̯	b	0.61	0.09	0.15	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯arna-ka	berx	rn	(r)r	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯r̆šnaka	birsêtî	u̯	b	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯r̆šnaka	birsêtî	rs	rs	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	ać̣ua	esp	ć̣u	sp	0.68	0.07	0.14	0.04	0.03	0.03	0.01	0.00	0.00	0.00
Kurdish Sorani	ć̣uaka	seg	ć̣u	s	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	ǰr̆d	dił	rd	l	0.65	0.08	0.16	0.04	0.02	0.03	0.01	0.00	0.00	0.00
Kurdish Sorani	darn	diṛîn	rn	(r)n	0.02	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	daća	de	ć	h	0.65	0.07	0.16	0.04	0.03	0.03	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯arka	berg	u̯	b	0.59	0.10	0.15	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯arka	berg	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	u̯(a)rda	gul	u̯	g	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯(a)rda	gul	rd	l	0.68	0.07	0.14	0.04	0.03	0.03	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯agza	boq	u̯	b	0.53	0.11	0.22	0.05	0.03	0.03	0.01	0.01	0.00	0.00
Kurdish Sorani	br̆zant	berz	rǰ	rǰ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	br̆zant	bi̅ind	rǰ	l	0.57	0.09	0.19	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	barzād	berzī	rǰ	rǰ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	hijų̆āna	ziman	jų	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	u̯r̆ka	gurg	u̯	g	0.23	0.16	0.34	0.08	0.07	0.07	0.01	0.01	0.01	0.01
Kurdish Sorani	u̯ata	bed	u̯	b	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯afra	befr	u̯	b	0.59	0.09	0.18	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯afra	befr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	ć̣iāu̯a	siya	ć̣i	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	barda	bêł	rd	l	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯ār	baran	u̯	b	0.58	0.07	0.20	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	mušti	mist	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	pr̆tu	pird	rt	rt	0.91	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	ar̆ša	wurch	rs	rch	0.47	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	ar̆ša	wurch	pro	no pro	0.00	0.21	0.45	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	caθu̯ŕ̆at	chil	rć	l	0.00	0.21	0.45	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	laupāća	r̆êwî	ć	h	0.56	0.09	0.20	0.05	0.04	0.04	0.01	0.00	0.00	0.00
Kurdish Sorani	laupāća	r̆êwî	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	u̯rinji	birinc	u̯	b	0.68	0.07	0.14	0.04	0.03	0.03	0.01	0.00	0.00	0.00
Kurdish Sorani	u̯ahuni	xön	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	g(a)rna-ka	gel	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	u̯āta	ba	u̯	b	0.48	0.11	0.23	0.06	0.05	0.05	0.01	0.01	0.00	0.01
Kurdish Sorani	u̯iāna	giyan	u̯	g	0.44	0.12	0.25	0.06	0.05	0.05	0.01	0.01	0.00	0.01
Kurdish Sorani	u̯iāna	jîn	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kurdish Sorani	u̯ićati	bîst	u̯	b	0.00	0.21	0.45	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Kurdish Sorani	u̯aina	bîn-	u̯	b	0.01	0.21	0.44	0.11	0.09	0.09	0.02	0.01	0.01	0.01
Vanishan	ać̣ua	asp	ć̣u	sp	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯āta	vōd	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯ār	vōrůn	u̯	w	0.90	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯ājí(n)	bōz	u̯	b	0.06	0.25	0.32	0.20	0.09	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯ājí(i̯aka?)	vōzî	u̯	w	0.96	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00

Vanishan	u̯at-caka	vecá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯ara	ver	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯afra	varf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯afra	varf	meta	meta	0.00	0.26	0.34	0.22	0.10	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯arka	valg	u̯	w	0.88	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯arka	valg	r	l	0.00	0.26	0.34	0.22	0.10	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯rinji	berénj	u̯	b	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯arna-ka	verrá	u̯	w	0.89	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯arna-ka	verrá	rn	(r)r	0.00	0.26	0.34	0.22	0.09	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯asī	ves	u̯	w	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	bṛzant	buländ	rʃ	l	0.50	0.13	0.17	0.11	0.05	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯ahāra	vōr	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯ahi̯a	veidér	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯aita	vīd	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯ići̯ati	bíst	u̯	b	0.00	0.26	0.34	0.22	0.10	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯ać̣i̯a	bīštár	ć̣i̯	š	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯ać̣i̯a	bīštár	u̯	b	0.01	0.26	0.34	0.21	0.09	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯ać̣i̯a	vīštér	ć̣i̯	š	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯ać̣i̯a	vīštér	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	barda	bār	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯idau̯ā	bívá	u̯	b	0.01	0.26	0.34	0.21	0.09	0.02	0.02	0.02	0.01	0.01
Vanishan	parna	palṓ'ī	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	puθra	pār	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	pṛtu	pūl	rt	l	0.44	0.16	0.16	0.13	0.06	0.01	0.01	0.01	0.01	0.01
Vanishan	(f)šupāna	cepún	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	caθu̯ṛ́at	cihél	ṛć	l	0.00	0.26	0.34	0.22	0.10	0.02	0.02	0.02	0.01	0.01
Vanishan	u̯ahuni	xin	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	ʃṛd	dil	rd	l	0.59	0.11	0.14	0.09	0.03	0.01	0.01	0.01	0.01	0.01
Vanishan	daća	dei	ć	h	0.59	0.11	0.14	0.09	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	rāšta	rōss	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	laupāća	rubṓ	ć	h	0.49	0.13	0.17	0.11	0.05	0.01	0.01	0.01	0.01	0.01
Vanishan	laupāća	rubṓ	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	hijúu̯āna	zevún	ju̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	ću̯aka	esbá	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	ći̯āu̯a	siyó	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	ću̯išša	ešpéš	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	skar	šekṓr	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	xšu̯ifta-	šir	t	r	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	krmi	kirm	r	r	0.62	0.10	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Vanishan	u̯ršnaka	gusná	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vanishan	u̯rka	gurg	u̯	g	0.19	0.21	0.28	0.18	0.08	0.01	0.02	0.02	0.01	0.01
Qohrud	ać̣u̯a	asp	ću̯	sp	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Qohrud	u̯āta	vōd	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯ār	vōrún	u̯	w	0.87	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯ājí(n)	bōz	u̯	b	0.05	0.29	0.31	0.18	0.09	0.02	0.03	0.01	0.01	0.01
Qohrud	barzād	bōlṓ	rʃ	l	0.50	0.15	0.16	0.09	0.05	0.01	0.01	0.01	0.01	0.01
Qohrud	u̯at-caka	vacá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯ara	ver	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯afra	vafr	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯afra	vafr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯arka	valg	u̯	w	0.87	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯arka	valg	r	l	0.00	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯rinji	berénj	u̯	b	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Qohrud	u̯arna-ka	verá	u̯	w	0.87	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯arna-ka	verá	rn	(r)r	0.00	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯aʃr(a)ka	gurd	u̯	g	0.93	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯asī	ves	u̯	w	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Qohrud	bṛzant	bulén	rʃ	l	0.45	0.17	0.18	0.10	0.05	0.01	0.02	0.01	0.01	0.01
Qohrud	u̯ana	benā	u̯	b	0.01	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯ahāra	bohōr	u̯	b	0.00	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯aita	vēt	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯ići̯ati	vīs	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	barda	bōl	rd	l	0.01	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯idau̯ā	vīgá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	parna	pári	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	puθra	pār	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	pṛtu	pāl	rt	l	0.39	0.20	0.16	0.12	0.06	0.01	0.02	0.01	0.01	0.01
Qohrud	(f)šupāna	cūpún	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	caθu̯ṛ́at	cil	ṛć	l	0.00	0.30	0.32	0.19	0.10	0.02	0.03	0.01	0.01	0.01
Qohrud	u̯ahuni	xān	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	ʃṛd	dil	rd	l	0.54	0.14	0.15	0.09	0.03	0.01	0.01	0.01	0.01	0.01
Qohrud	daća	di	ć	h	0.53	0.14	0.15	0.09	0.05	0.01	0.01	0.01	0.01	0.01
Qohrud	rāšta	rōst	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	laupāća	rubṓ	ć	h	0.44	0.17	0.18	0.10	0.05	0.01	0.02	0.01	0.01	0.01
Qohrud	laupāća	rubṓ	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	hijúu̯āna	īlzún	ju̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	spṛzan	esbārz	rʃ	rʃ	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	ću̯aka	espá	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	ći̯āu̯a	sigṓ	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	ću̯išša	üšpüš	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	skar	eškṓr	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	xšu̯ifta-	šit	t	t	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Qohrud	krmi	kirm	r	r	0.57	0.13	0.14	0.08	0.04	0.01	0.01	0.01	0.01	0.01
Qohrud	u̯ršnaka	veš	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Qohrud	u̯rka	var	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	ać̣u̯a	asm	ću̯	sp	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01

Kashan	u̯āta	vōi	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯ār	vōrůn	u̯	w	0.86	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	barzād	bōlő	rʃ	l	0.47	0.15	0.17	0.11	0.05	0.01	0.01	0.01	0.01	0.01
Kashan	u̯at-caka	vacá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯ara	ver	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯afra	varf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯afra	varf	meta	meta	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	u̯arka	valg	u̯	w	0.85	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯arka	valg	r	l	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	u̯rinji	beréńj	u̯	b	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	u̯arna-ka	verá	u̯	w	0.86	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯arna-ka	verá	rn	(r)r	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	u̯aʃr(a)ka	gurd	u̯	g	0.92	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯asī	ves	u̯	w	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	bṛzant	beláṇd	rʃ	l	0.42	0.17	0.19	0.12	0.05	0.01	0.01	0.01	0.01	0.01
Kashan	u̯ahāra	bohōr	u̯	b	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	u̯ahi̯a	bahtár	u̯	b	0.15	0.25	0.28	0.18	0.07	0.01	0.02	0.01	0.01	0.01
Kashan	u̯aita	vē	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯ićati	víssā	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯aći̯a	vīštár	ći̯	š	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	u̯aći̯a	vĩštár	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	barda	bāl	rd	l	0.01	0.29	0.33	0.21	0.08	0.02	0.02	0.02	0.02	0.01
Kashan	u̯idau̯ā	bĩvá	u̯	b	0.01	0.29	0.33	0.21	0.08	0.02	0.02	0.02	0.02	0.01
Kashan	parna	péri	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	puθra	pắr	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	prtu	pōl	rt	l	0.36	0.20	0.17	0.15	0.06	0.01	0.02	0.01	0.01	0.01
Kashan	(f)šupāna	cūpún	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	caθu̯ṛća	cáltō	ṛć	l	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	u̯ahuni	xắn	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	ʃṛd	dil	ṛd	l	0.51	0.15	0.17	0.11	0.03	0.01	0.01	0.01	0.01	0.01
Kashan	daća	dah	ć	h	0.51	0.14	0.16	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	rāšta	rōšt	št	št	0.00	0.29	0.33	0.21	0.09	0.02	0.02	0.02	0.02	0.01
Kashan	hiju̯āna	uzmůn	ju̯	zm	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	ći̯āu̯a	siyő	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	ćuišа	ešpíš	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	skar	eškőr	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	xšu̯ifta-	šir	t	r	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	krmi	karm	r	r	0.54	0.13	0.15	0.10	0.04	0.01	0.01	0.01	0.01	0.01
Kashan	u̯ršnaka	vešá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kashan	u̯rka	varg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	aćua	asm	ću̯	sp	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	u̯āta	vō	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯ār	vōrō	u̯	w	0.91	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯āji(n)	bōz	u̯	b	0.08	0.29	0.29	0.18	0.08	0.02	0.02	0.02	0.02	0.01
Zefra	barzād	bōlő	rʃ	l	0.63	0.12	0.12	0.07	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	u̯at-caka	vecá	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯ara	ver	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯afra	varf	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯afra	varf	meta	meta	0.01	0.31	0.31	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯arka	velg	u̯	w	0.92	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯arka	velg	r	l	0.00	0.31	0.32	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯rinji	beréńj	u̯	b	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	u̯arna-ka	veré	u̯	w	0.92	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯arna-ka	veré	rn	(r)r	0.01	0.31	0.31	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯aʃr(a)ka	bür	u̯	b	0.38	0.16	0.21	0.13	0.06	0.01	0.01	0.01	0.01	0.01
Zefra	u̯asī	ves	u̯	w	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	bṛzant	belén	rʃ	l	0.58	0.13	0.13	0.08	0.04	0.01	0.01	0.01	0.01	0.00
Zefra	u̯ana	bené	u̯	b	0.02	0.31	0.31	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯ahāra	bohōr	u̯	b	0.01	0.31	0.31	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯ahi̯a	bijtér	u̯	b	0.25	0.23	0.24	0.15	0.06	0.01	0.02	0.01	0.01	0.01
Zefra	u̯aita	vē	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯ićati	vĩss	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯aći̯a	vĩštér	ći̯	š	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	u̯aći̯a	vĩštér	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	barda	berd	rd	rd	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯idau̯ā	bĩvé	u̯	b	0.01	0.31	0.31	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	parna	per	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	puθra	pūr	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	prtu	pōl	rt	l	0.52	0.16	0.12	0.10	0.05	0.01	0.01	0.01	0.01	0.01
Zefra	(f)šupāna	capő	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	caθu̯ṛća	cáltō	ṛć	l	0.00	0.31	0.32	0.19	0.09	0.02	0.02	0.02	0.02	0.01
Zefra	u̯ahuni	xắn	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	ʃṛd	dil	ṛd	l	0.66	0.11	0.11	0.07	0.02	0.01	0.01	0.01	0.01	0.00
Zefra	daća	dei	ć	h	0.66	0.11	0.11	0.07	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	rāšta	rōss	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	hiju̯āna	oző	ju̯	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	ći̯āu̯a	siyő	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	ćuišа	išpíš	ću̯	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	skar	šegőr	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	xšu̯ifta-	šir	t	r	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	krmi	kerm	r	r	0.69	0.10	0.10	0.06	0.03	0.01	0.01	0.01	0.01	0.00
Zefra	u̯ršnaka	vešé	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zefra	u̯rka	varg	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Natanz	aćua	asm	ću̯	sp	0.76	0.05	0.09	0.05	0.02	0.01	0.01	0.01	0.01	0.00
Natanz	u̯āta	bōd	u̯	b	0.57	0.10	0.16	0.08	0.04	0.01	0.02	0.01	0.01	0.01
Natanz	u̯ār	vōrůn	u̯	w	0.95	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Natanz	u̯ara	ver	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Natanz	u̯arna-ka	barά	u̯	b	0.69	0.08	0.09	0.07	0.03	0.01	0.01	0.01	0.01	0.01
Natanz	u̯arna-ka	barά	rn	(r)r	0.01	0.22	0.36	0.19	0.09	0.02	0.04	0.02	0.02	0.02
Parthian	(f)šupāna	šubān	š	š	0.11	0.25	0.31	0.17	0.07	0.03	0.03	0.01	0.01	0.01
Parthian	ćuaka	ispag	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	ćiāu̯a	syāw	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	aćua	asp	ću	sp	0.48	0.15	0.18	0.10	0.04	0.02	0.02	0.01	0.01	0.01
Parthian	caθurćat	cafrast	rć	rć	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	daća	das	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	hijuāna	iz,βān	ju	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	laupāća	rōbās	ć	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	laupāća	rōbās	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	prtu	purt	rt	rt	0.86	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	puθra	puhr	tr	tr	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	rāšta	rāšt	št	št	0.00	0.28	0.34	0.19	0.07	0.03	0.04	0.01	0.02	0.01
Parthian	u̯ićati	wīst	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯(a)rda	wār	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯(a)rda	wār	rd	rd	0.48	0.15	0.18	0.10	0.04	0.02	0.02	0.01	0.01	0.01
Parthian	u̯āta	wād	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯uršnaka	wišāy	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯uršnaka	wišāy	rs	š	0.00	0.28	0.34	0.19	0.07	0.03	0.04	0.01	0.02	0.01
Parthian	u̯urka	gurg	u̯	g	0.12	0.25	0.30	0.17	0.06	0.03	0.03	0.01	0.01	0.01
Parthian	u̯aȷr(a)ka	wuzurg	u̯	w	0.91	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯afra	wafr	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯afra	wafr	meta	no meta	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯ahuni	guxn	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯aina	wēn	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯arka	wargar	u̯	w	0.81	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯arka	wargar	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯arna-ka	warrag	u̯	w	0.82	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯arna-ka	warrag	rn	(r)r	0.00	0.28	0.34	0.19	0.07	0.03	0.04	0.01	0.02	0.01
Parthian	u̯art-	ward	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯asī	was	u̯	w	0.48	0.15	0.18	0.10	0.04	0.02	0.02	0.01	0.01	0.01
Parthian	u̯ata	wad	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯idāna	wiδān	u̯	w	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯ināća	wināh	ć	h	0.48	0.15	0.18	0.10	0.04	0.02	0.02	0.01	0.01	0.01
Parthian	u̯ināća	wināh	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯iyapāna	wiyābān	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Parthian	u̯iāna	gyān	u̯	g	0.26	0.21	0.26	0.14	0.05	0.02	0.03	0.01	0.01	0.01
Parthian	ȷrd	dil	rd	l	0.45	0.16	0.19	0.11	0.03	0.02	0.02	0.01	0.01	0.01
Persian	(f)šupāna	šubān	š	š	0.36	0.16	0.17	0.08	0.05	0.14	0.01	0.01	0.01	0.01
Persian	(f)šupāna	cubān	š	c	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	ćuaka	sag	ću	s	0.01	0.25	0.27	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	ćuiša	uspus	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	ćuiša	ušpus	ću	sp	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	ćiāu̯a	siyāh	ći	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	ćiāu̯a	šāh	ći	š	0.01	0.25	0.27	0.13	0.08	0.22	0.02	0.01	0.01	0.01
Persian	aćru	ars	ćr	rs	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	aćru	ars	pro	no pro	0.30	0.18	0.19	0.09	0.05	0.15	0.01	0.01	0.01	0.01
Persian	aćua	asb	ću	sp	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	arša	xirs	rs	rs	0.78	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	arša	xirs	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	brzant	buland	rȷ́	l	0.73	0.07	0.07	0.03	0.02	0.06	0.00	0.00	0.00	0.00
Persian	barda	bel	rd	l	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	barzād	bālā	rȷ́	l	0.77	0.06	0.06	0.03	0.02	0.05	0.00	0.00	0.00	0.00
Persian	caθurćat	cihil	rć	l	0.00	0.26	0.27	0.13	0.08	0.22	0.02	0.01	0.01	0.01
Persian	daća	da	ć	h	0.79	0.05	0.06	0.03	0.02	0.05	0.00	0.00	0.00	0.00
Persian	darn	darrīdan	rn	(r)n	0.04	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	g(a)rna-ka	galla	rn	l	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	gau-ćuanta-	gospand	ću	sp	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	gau-ćuanta-	gosfand	ću	sp	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	hijuāna	zabān	ju	zb	0.01	0.25	0.27	0.12	0.08	0.22	0.02	0.01	0.01	0.01
Persian	hijuāna	zawān	ju	zw	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	krmi	kirm	r	r	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	kaćiapa	kašaf	ći	š	0.04	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	laupāća	rūbā	ć	h	0.72	0.07	0.08	0.04	0.02	0.06	0.00	0.00	0.00	0.00
Persian	laupāća	rūbā	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	mušti	mušt	št	št	0.67	0.09	0.09	0.04	0.03	0.07	0.01	0.00	0.00	0.00
Persian	prtu	pūl	rt	l	0.67	0.09	0.07	0.04	0.03	0.08	0.01	0.01	0.00	0.00
Persian	parna	parr	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	puθra	pus	tr	s	0.65	0.09	0.10	0.04	0.03	0.08	0.01	0.01	0.00	0.00
Persian	rāšta	rāst	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	sprzan	uspurz	rȷ́	rȷ́	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯ahāra	bahār	u̯	b	0.01	0.25	0.27	0.12	0.08	0.22	0.02	0.01	0.01	0.01
Persian	u̯ićati	bīst	u̯	b	0.01	0.25	0.27	0.13	0.08	0.22	0.02	0.01	0.01	0.01
Persian	u̯(a)rda	gul	u̯	g	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯(a)rda	gul	rd	l	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯āf	bāf	u̯	b	0.54	0.13	0.13	0.06	0.04	0.08	0.01	0.01	0.01	0.00
Persian	u̯ār	bār	u̯	b	0.74	0.05	0.08	0.04	0.02	0.06	0.00	0.00	0.00	0.00
Persian	u̯āta	bād	u̯	b	0.64	0.09	0.10	0.05	0.03	0.08	0.01	0.01	0.00	0.00
Persian	u̯āȷi(n)	bāz	u̯	b	0.14	0.22	0.23	0.11	0.07	0.19	0.01	0.01	0.01	0.01
Persian	u̯āȷi(iaka?)	bāzī	u̯	b	0.53	0.13	0.13	0.05	0.04	0.11	0.01	0.01	0.01	0.00
Persian	u̯uršnaka	gurusna	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯uršnaka	gušna	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯uršnaka	gušna	rs	š	0.01	0.25	0.27	0.13	0.08	0.22	0.02	0.01	0.01	0.01
Persian	u̯urka	gurg	u̯	g	0.38	0.16	0.17	0.08	0.05	0.14	0.01	0.01	0.01	0.01
Persian	u̯urtka	gurda	u̯	g	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯aćȷa	beš	ći	š	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00

Persian	u̯aćia	beš	u̯	b	0.03	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	u̯aȷ́r(a)ka	buzurg	u̯	b	0.54	0.10	0.13	0.06	0.04	0.11	0.01	0.01	0.01	0.00
Persian	u̯aȷ́ra	gurz	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯afra	barf	u̯	b	0.74	0.07	0.07	0.03	0.02	0.06	0.00	0.00	0.00	0.00
Persian	u̯afra	barf	meta	meta	0.01	0.25	0.27	0.13	0.08	0.22	0.02	0.01	0.01	0.01
Persian	u̯agza	wazaγ	u̯	w	0.96	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.00
Persian	u̯ahuni	xūn	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯ahia	bih	u̯	b	0.39	0.16	0.16	0.08	0.05	0.13	0.01	0.01	0.01	0.01
Persian	u̯aišа	bīša	u̯	b	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯aicīna?	bēz	u̯	b	0.04	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	u̯aina	bīn	u̯	b	0.03	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯aita	bīd	u̯	b	0.20	0.21	0.22	0.10	0.06	0.18	0.01	0.01	0.01	0.01
Persian	u̯ana	bun	u̯	b	0.03	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	u̯anjecaka	binjišk	u̯	b	0.06	0.24	0.25	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	u̯anjecaka	gunjišk	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯arāȷ́a	gurāz	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯arāȷ́a	wurāz	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯ara	bar	u̯	b	0.74	0.07	0.07	0.03	0.02	0.06	0.00	0.00	0.00	0.00
Persian	u̯arka	barg	u̯	b	0.73	0.07	0.06	0.04	0.02	0.06	0.00	0.00	0.00	0.00
Persian	u̯arka	barg	r	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯arna-ka	barra	u̯	b	0.74	0.07	0.06	0.03	0.02	0.06	0.00	0.00	0.00	0.00
Persian	u̯arna-ka	barra	rn	(r)r	0.01	0.25	0.27	0.12	0.08	0.22	0.02	0.01	0.01	0.01
Persian	u̯art-	gard	u̯	g	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯asī	bisyār	u̯	b	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯at-caka	bacca	u̯	b	0.62	0.10	0.10	0.05	0.03	0.08	0.01	0.01	0.00	0.00
Persian	u̯ata	bad	u̯	b	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯icāra	guzār	u̯	g	0.61	0.10	0.11	0.05	0.03	0.09	0.01	0.01	0.00	0.00
Persian	u̯idāna	giyān	u̯	g	0.98	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	u̯idau̯ā	bewa	u̯	b	0.03	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯ināća	gunāh	ć	h	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯ināća	gunāh	u̯	g	0.02	0.25	0.26	0.12	0.07	0.22	0.02	0.01	0.01	0.01
Persian	u̯iiapāna	biyābān	u̯	b	0.04	0.25	0.26	0.12	0.07	0.21	0.02	0.01	0.01	0.01
Persian	u̯rinji	birinj	u̯	b	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯rinji	gurinj	u̯	g	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯iāka	jā	u̯	j	0.81	0.05	0.05	0.02	0.01	0.04	0.00	0.00	0.00	0.00
Persian	u̯iāna	jān	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Persian	ȷ́rd	dil	rd	l	0.79	0.06	0.06	0.03	0.01	0.05	0.00	0.00	0.00	0.00
Bandari	aćru	xars	ćr	rs	0.02	0.18	0.25	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	aćru	xars	pro	pro	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	aću̯a	asp	ću̯	sp	0.86	0.03	0.04	0.03	0.02	0.02	0.00	0.00	0.00	0.00
Bandari	kaći̯apa	kāsapošt	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	sp̄rzan	espül	rʃ	l	0.60	0.07	0.10	0.08	0.05	0.05	0.01	0.01	0.01	0.01
Bandari	u̯ahāra	buhār	u̯	b	0.02	0.18	0.26	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	u̯(a)rda	gol	u̯	g	0.02	0.18	0.25	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	u̯(a)rda	gol	rd	l	0.86	0.03	0.04	0.03	0.02	0.02	0.00	0.00	0.00	0.00
Bandari	u̯āȷ́i(i̯aka?)	gāzi	u̯	g	0.99	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	u̯rs̈naka	gušna	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	u̯rs̈naka	gušna	rs	š	0.01	0.18	0.26	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	u̯agza	gwak	u̯	g	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	u̯aicīna?	gīz	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	u̯aina	gin	u̯	g	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bandari	u̯ana	bun	u̯	b	0.04	0.18	0.25	0.18	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	u̯ata	bad	u̯	b	0.03	0.18	0.25	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	u̯idau̯ā	biva	u̯	b	0.04	0.18	0.25	0.19	0.12	0.13	0.03	0.03	0.02	0.02
Bandari	ȷ́rd	del	rd	l	0.84	0.03	0.04	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Bandari	skar	eškāl	r	l	0.10	0.17	0.23	0.17	0.11	0.12	0.03	0.03	0.02	0.02
Gilaki	ću̯aka	səg	ću̯	s	0.02	0.29	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	ći̯āu̯a	siyā	ći̯	s	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	aću̯a	asp	ću̯	sp	0.88	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	br̄zant	buləɴd	rʃ	l	0.83	0.05	0.05	0.03	0.01	0.02	0.00	0.01	0.00	0.00
Gilaki	barzād	bålå	rʃ	l	0.85	0.04	0.04	0.03	0.01	0.02	0.00	0.00	0.00	0.00
Gilaki	caθu̯rćat	cel	rć	l	0.01	0.29	0.28	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	daća	da	ć	h	0.87	0.04	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	hiȷ́u̯āna	zəbån	ȷ́u̯	zb	0.02	0.29	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	laupāća	rubå	ć	h	0.82	0.05	0.05	0.03	0.01	0.02	0.00	0.01	0.00	0.00
Gilaki	laupāća	rubå	l	r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	pr̄tu	purd	rt	rt	0.98	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	parna	pər	rn	(r)r	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	puθra	pəsər	tr	s	0.77	0.07	0.07	0.04	0.01	0.03	0.01	0.01	0.01	0.00
Gilaki	rāšta	råst	št	st	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	u̯ahāra	bəhår	u̯	b	0.02	0.29	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	u̯ićati	bist	u̯	b	0.02	0.29	0.28	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	u̯(a)rda	gul	u̯	g	0.03	0.28	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	u̯(a)rda	gul	rd	l	0.88	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	u̯āf	båf	u̯	b	0.67	0.10	0.10	0.06	0.02	0.03	0.01	0.01	0.01	0.01
Gilaki	u̯ār	bårəš	u̯	b	0.83	0.04	0.05	0.03	0.01	0.02	0.00	0.01	0.00	0.00
Gilaki	u̯āta	båd	u̯	b	0.76	0.07	0.07	0.04	0.01	0.03	0.01	0.01	0.01	0.00
Gilaki	u̯āȷ́i(i̯aka?)	båzi	u̯	b	0.67	0.10	0.10	0.04	0.02	0.04	0.01	0.01	0.01	0.01
Gilaki	u̯rs̈naka	višta	u̯	w	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	u̯rka	gurg	u̯	g	0.52	0.14	0.13	0.08	0.02	0.05	0.01	0.01	0.01	0.01
Gilaki	u̯aȷ́r(a)ka	buzurg	u̯	b	0.68	0.08	0.10	0.06	0.02	0.04	0.01	0.01	0.01	0.01
Gilaki	u̯ahuni	xun	u̯	x	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	u̯art-	gərd-	u̯	g	0.03	0.28	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	u̯ata	bəd	u̯	b	0.03	0.28	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02
Gilaki	u̯icāra	guzər	u̯	g	0.73	0.08	0.07	0.05	0.01	0.03	0.01	0.01	0.01	0.00
Gilaki	u̯ināća	gunå	ć	h	0.88	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	u̯ināća	gunå	u̯	g	0.03	0.28	0.27	0.17	0.05	0.11	0.03	0.03	0.02	0.02

Gilaki	u̯iiapāna	biyåbån	u̯	b	0.06	0.27	0.26	0.16	0.04	0.10	0.03	0.03	0.02	0.02
Gilaki	u̯rinji	bərənǰ	u̯	b	0.88	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	u̯iāka	ǰåy	u̯	j	0.88	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
Gilaki	u̯iāna	ǰån	u̯	j	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gilaki	ǰṛd	dil	rd	l	0.87	0.04	0.04	0.02	0.00	0.01	0.00	0.00	0.00	0.00