The Caspian Language of Šahmirzād

Habib Borjian
Columbia University

Located in the Semnān area (midway between Tehran and Khorasan), the town of Šahmirzād and its neighboring villages are home to speakers of Šahmirzādi, a vernacular sharply differing from the other language types spoken in the Semnān area but closely related to the Mazandarani language spoken across the Alborz range to the north, along the Caspian coast. This article studies Šahmirzādi phonology, grammar, and vocabulary, with a look at cross-linguistic influence in the situation of language contact. The article concludes with a discussion of the possible status of Šahmirzādi as a separate language within the Caspian family.

1. INTRODUCTION

The township of Šahmirzād (locally Šâmerzâ) sprawls along the southern slopes of the Alborz range, 15 miles north of Semnān, at 35.8° north latitude, 53.3° east longitude, and 2,000 meters above sea level. Two parallel mountain ridges separate Šahmirzād, which is in the Semnān district, from Dodānga and Savādkuh districts of the Mazandaran (Māzandarān) province. Downslope from Šahmirzād lie Sangesar and Semnān, each with its distinct Iranian language, forming a Sprachbund with the nearby Sorxa'i, Lāsgerdi, and Aftari, all of which are crowded into the district of Semnān. The permanent population of the township of Šahmirzād, recorded as 7,273 individuals in the 2006 and 8,882 in 2011 censuses, swells significantly in summers upon the return of the residents who work elsewhere. The villages located in the valleys on the north of Šahmirzād, including Deh Ṣufiān, Āserān, Jāšm, and Garm Čašma, speak varieties of Šahmirzādi. See Figure 1.[1] The local residents estimate the total number of speakers of Šahmirzādi to be around 5,000 individuals.

1.1. *Documentation*

Documentation of Šahmirzādi was begun in the 1880s by Valentin Žukovskij (henceforth Žuk.), who elicited from Persian short texts and anecdotes and a list of words and verb paradigms (Žukovskij 1922: 5–8 and glossary). This was followed by compilation of basic morphology along with two short texts by Arthur Christensen (henceforth Chris.; 1935: 142–78),[2] brief notes by Georg Morgenstierne (henceforth Morgen.; 1960: 108–9), a comparative lexicon with other languages of Semnān by Manučehr Sotuda (henceforth Sotu.; 1963), an elicited text by Irān Kalbāsi (henceforth Kalb.; 2009: 531–33), and a typological study by Mohammad Dabir-Moqaddam (henceforth Dabir.; 2014: vol. II, 1034–90). The unpublished materials consulted in this study include a list of words and sentences from the 1960s collected by Hušang Purkarim (henceforth Pur.), kindly shared by a colleague of his

The author would like to thank Erik Anonby for his meticulous review of the section on phonology, and Daniel Kaufman and Daniel A. Barry for their insights on the phonetics of the recorded data. I am also indebted to the anonymous reviewer for many comments and insights.

1. To the southwest of Semnān, in the Garmsār area, the village of Farvān has an outlier Mazandarani dialect of Šahmirzādi type. See Borjian 2013b.

2. Examples are cited from Christensen (pp. 171–75), unless otherwise noted.

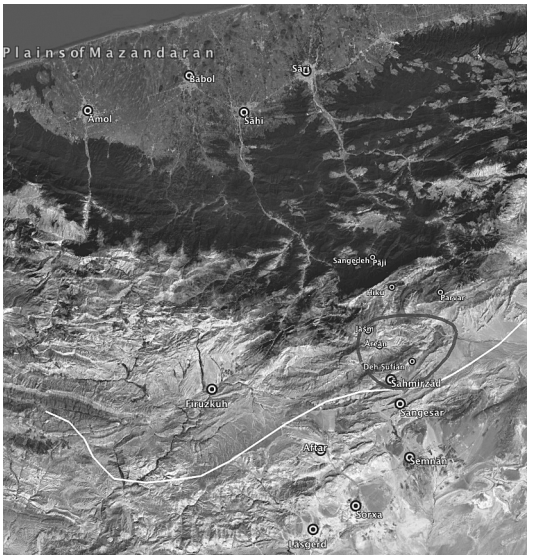

Fig. 1. Map of Semnān and central-eastern Mazandaran. The white line delineates the Caspian linguistic zone on the north. The dark ring marks the extent of Šahmirzādi. The satellite map is from Google Earth.

who preferred to stay anonymous; audio recordings in 2003 of a 227-item basic vocabulary, shared generously by Alexander Kolbitsch; and the present author's unpublished documentation (since 2009), which is left unmarked in the citations that follow.

In this study I have chosen to draw upon all these credible sources rather than eliminating past knowledge in favor of my own field notes. In view of the incessant weakening of the local idioms under the pressure of Persian, I believe that no reasonable documentation should be ignored with the pretext that the methodologies used in the past are incompatible with current practice. Multiplicity of sources surely adds to the richness of language description even if the job becomes more complicated.

The materials at hand, being from multiple sources, show considerable variance in transcription; for instance, the gloss 'eye' is elicited by Žukovskij as *čašm*, by Christensen as *čäš*, *čaš*, *čašm*, by Sotuda as *čaš*; and 'wind' has the outputs Žuk. *vō*, Sotu. *vâ*, Chris. *bå̄d*, and my field-notes [vɑ·].[3] In Christensen's documentation, it is evident that the speech of his chief informant is strongly persianized, marking phonology (*barf*, *barg*, *bīd*, *bīst*, which are *v*-initial words in other sources), morphology (verb personal endings, §4.4), and syntax (adpositions, §3.4); these discrepancies could partly be due to infected idiolects of seasonal workers who winter in Mazandaran and elsewhere. Note that the phoneme /â/ (§2.2.2) is rendered as variously as Žuk. <ō>,[4] Chris., Morgen. <å>, Sotu., Kalb. <â>, Pur., Dabir. <a>, while /a/ (§2.2.7) is rendered in the latter two sources as <æ>. The original symbols are

3. For more comparisons, see Rastorgueva and Edel'man 1982: 475–80.

4. Žukovskij tends to transcribe back low vowels with symbols that suggest mid rather than low place of articulation.

Table 1. Vowel Phonemes and Allophones

	Front	Back
High	/i/ [i]	/ü/ [u, ʉ, ʏ]
Mid	/e/ [e, ɪ]	/ö/ [o, ʊ, ɵ]
	/ɛ/ [ɛ, ɜ, ə]	
Low	/a/ [a, æ, ɐ]	/â/ [ɑ, ɒ]

retained in the citations below. I decided to stay loyal to transcription symbols from the original sources, although it may introduce an additional level of complexity.

2. PHONOLOGY

2.1. *Consonants.* The consonantal inventory of Šahmirzādi is /p b t d k g č j f v s z š ž x γ h m n r l y/.

2.1.1. The phonemic value of *ž* [ʒ] is to be verified; its presence in my notes is limited to *každöm* [kæʒ̈dø·m] 'scorpion' and *hižda* [hiʒ̈dæ] 'eighteen'. As in Mazandarani proper, the existence of *ž* therefore appears to be limited to regressive assimilation of the phoneme /j/ [dʒ] before /d/.

2.1.2. Dorsals. The articulations of /k/ and /g/ are not exclusively velar as is the case in Mazandarani; younger speakers tend to follow the Persian model of palatalizing before front vowels; thus, *gal* 'throat' is heard as both [gæl] and [ɟæl]. Likewise, the prevalent Persian way of switching between back velar/uvular fricative and stop ([ʁ] and [ɢ] in most speakers) is also the norm in Šahmirzādi, contrary to Mazandarani, in which these sounds have collapsed to voiced velar fricative [ʁ] in all positions. This study employs both *γ* and *q*, e.g., *qöšâr* [qɵ̈ʃɑ·r] 'pressure' and *marγona* [mæɾʁʊ̈næ] 'egg'; the distributional relationship between these allophones can only be established with sufficient data.

2.2. *Vowels*. This section proposes the Šahmirzādi vowel inventory as /â a ɛ e i ü ö/. These phonemes are arranged together with their phonetic realizations in Table 1. The choice to include the fronted allophones as basic for the phonemes /ü/ and /ö/ is made due to their higher occurrences (§2.2.8), although this choice leads to an asymmetrical phonemic inventory. The sections that follow present a detailed analysis of each vowel sound and its considerable variation among speakers. Variations are mostly due to different speakers, but they may also be the perception of different listeners to the same audiotapes.

2.2.1. Vowel length. Although Middle Iranian length distinction has not reached Šahmirzādi systematically, /â u i e/ are perceivably articulated long when compared to other vowel sounds. Besides, as in Persian, slight lengthening precedes consonant clusters in coda, as in *abr* [æ·br] 'cloud', *valg* [væ·lg] 'leaf'. The lengthening in *rasεn* [ræ·ˈsɛn] 'rope' and *ašün* [æ·ʃʉ̈n] 'last night' in my notes should be phonetic, due to pragmatic factors. In fact, vowel lengthening in initial syllables is common in many parts of Iran.

2.2.2. /â/ is realized as [ɑ], sometimes partially rounded to [ɒ], and is usually articulated half long: [ɑ·ˈtʒ̈ʃ] 'fire', [tʃɑ·ˈkæ] 'river', [mɑ·r] 'mother', [dømdæˈrɑ·z] 'snake', [vɑ·] 'wind', [tʃitɛ̈ˈkɑ·] 'chick', [xɒˈxɛr] 'sister'. It was heard as slightly raised in [dɒ̝·r] 'tree', [ʃɒ̝x] 'horn', [mɒ̝h] 'moon', [nɒ̝·f] 'navel', among other words. In [væʃnɑ̈ˈi] 'hunger', /â/ is noticeably short, compared with [væʃˈnɑ·] 'hungry'. It is therefore hard to judge whether the position in stressed vs. unstressed syllables imposes a condition in this phonetic variation. More data is needed to establish free vs. conditional variation for this phoneme.

2.2.3. /i/ is half long or long, especially in final position: [diː] ‘smoke’, [taʹliː] ‘rock’, [miːʃ] ‘mouse’, [diːʹmæ] ‘hill’, [mədʒiːʹlæ] ‘ant’, [tim] ~ [tiːm] ‘seed’, [riːʃæ] ‘root’. As is evident in these examples, historical *ī and *ū have merged into /i/; for example, *miš* < mūš ‘mouse’, *riša* < rīša ‘root’.

2.2.4. /e/ is half long and is principally derived from Middle West Iranian *ē, as in *jer* [d͡ʒeːr] ‘below’, *reg* [reːg] ‘pebble’, *tej* ‘sharp’, *meš* ‘ewe’, *xeš* ‘kin’, *aspe* [æsʹpęː] ‘white’ (<Mid. West Ir. *spēd), *ke* ‘who’ (cf. New Pers. kǐ, Mid. Pers. kē). Note the minimal pair between /e/ and /i/, in *der* ‘late’ (< dēr < *dayr < *dagr) versus *dir* ‘far’ (< dūr); cf. Pers. *dir* and *dur* respectively. Note also *šir* [ʃiːr] ‘lion’ (< šēr < *šayr < *šagr), which has become homonymous with *šir* ‘milk’ (< šīr); therefore, both words are likely borrowings from Persian.

There are a few cases in which /e/ has other origins: *per* [per] ~ [pęr] ‘father’ (probably, as other varieties of Caspian imply, with the protoform *pɪər, from Mid. West Ir. *pidar), *šeš* [ʃeʃ] ‘six’, and the “inverse-ežāfa” marker (§3.2), which is used to form possessive pronouns (*me*, *te*, *e*, etc.; Table 2).

There are instances where [e] overlaps with [ɪ], as in *aspez* [asʹpez] ~ [æsʹpɪz] ~ [ɛsʹpɛz] ‘louse’ (< OIr. *spiš-).

2.2.5. /ɛ/. The space between the phonemes /e/ and /a/ is filled with a range of open-mid vowels that includes [ɛ, ə, ɜ]. These sounds at times overlap with /a/, less likely with /e/, and therefore could reasonably be placed in the domain of either of these principal neighboring phonemes. On the other hand, the high frequency of the range [ɛ ə ɜ] affords a distinct phoneme, designated in this study as /ɛ/, which by and large agrees with the symbol ə used by Dabir-Moqaddam. I could not identify any minimal pairs in the data between either /ɛ/ and /e/ or /ɛ/ and /a/. However, in support of a phonemic status for /ɛ/ it should be added that, notwithstanding its frequent allophonic intersection with /a/, the Šahmirzādi speakers commonly perceive it as Persian *kasra* far more than *fatḥa*. The following examples from the data are meant to show the allophonic range of the sound: *yɛk* [yɛk] ~ [yək] ‘one’, *šɛn* [ʃɛn] ~ [ʃæn] ‘sand’, *lɛng* [lɛŋ] ~ [ləŋ] ~ [læŋ] ‘leg’, *vɛni* [vəʹniː] ~ [væʹniː] ‘nose’, *aškɛm* [əʃʹkɛm] ~ [æʃʹkæm] ‘belly’, *âtɛš* [ɑːʹtʂʃ] ~ [ɑːʹtəʃ] ‘fire’, *ruɣɛn* [ruʹʁɛn] ~ [ruːʹʁən] ‘ghee’, *bɛrâr* [bɜʹrɑr] ~ [bɛʹrɒr] ‘brother’, *sɛtâra* [səʈɑʹræ] ~ [sʂʈɑʹræ] ~ [sætɑʹræ] ‘star’, *assɛɣon* [æsːəʹʁõ] ~ [ɛsːəʹʁõ] ‘bone’, *ɛsm* [ʂs(m)] ~ [æs(m)] ‘name’.

2.2.6. In final position a front-mid vowel is heard in the following words of obscure etymon: *zâyɛ* ‘grandchild’ (<? zādak), *xalɛ* [xælęː] ‘many’ (cf. Tajik *xale*), *lavɛ* [læːʹvę] ‘pot’, *sâzɛ* [sɑːʹzɛ] ‘broom’. Nevertheless, the most common outcome of Middle West Iranian *-ak is /a/, such as in *barma* [bærʹmæ/ə] ‘weep, cry’, *šarša* [ʃærʹʃæ] ‘rock’, *vara* ‘lamb’, *šâxa* ‘branch’, and the unstressed third person ending (Table 5). More work is needed to determine if word-final /a/ and /ɛ/ really contrast or the variance is caused by a wavering pronunciation of one underlying form.

2.2.7. /a/ covers a wide range in the front open quarter of the vowel space, that is, open [a], near-open [æ], and central-unrounded [ɐ], as well as less open articulations in the speech of some speakers. It is therefore not always clear how to distinguish this phoneme from its neighbor /ɛ/ (see §2.2.5), as in [dæl] ~ [dʂl] ~ [dɐl] ‘heart’ and [kɜrk] ~ [kɐrk] ‘hen’. Apart from the strong tendency to centralize, both allophones [æ] and [a] can be heard in the speech of the same speaker in the data: [sækʰ] ‘dog’, [næʃt] ‘dirty’ vs. [gal] ‘throat’, [das] ‘hand’, [mard] ‘man’.

2.2.8. Round back vowels are /o/ and /u/ and their front rounded allophones /ö/ and /ü/, which appear to have higher frequencies than their back counterparts. Historical centralization or fronting of back rounded vowels, atypical of the Caspian languages (cf. §2.2.11), is in

all likelihood a local development of Šahmirzādi, and if so, the historical process of fronting is not yet complete. On the other hand, one may reasonably assume a regression from central ö and ü toward o and u under the more recent influence of Persian, considering the rarity of ö and ü in the speech of younger speakers.

In this study I have elected the fronted pair as principal phonemes, on the grounds of higher frequency of ö and ü and their rarity in the speech of younger speakers. Erik Anonby (pers. comm.) views o and u as the underlying phonemes, for reasons of descriptive symmetry in the structure of the vowel system and the fact that the typologically more common vowels o and u are more likely to have fronted allophones, than front rounded vowels ö and ü are to have backed allophones.

2.2.9. /ö/ is commonly close-mid central [ɵ] and can be realized as [o, ʊ] as well. Examples in the data are: [bɵz] 'goat', [nɵ(h)] 'nine', [pɵʃt] 'back', [dɵm] ~ [dʊm] 'tail', [kɵl] 'short (in length)', [gɵl ~ gʊl] 'flower', [pɵr] ~ [pʊr] 'full', [sɵrx] ~ [sʊrx] 'red', [ʃɵˈtɵr] ~ [ʃˈotɐr] 'camel', [bɵˈmo̜rdæ] 'he died', [bɵˈxo̜tæ] 'he slept', [bɵˈgo̜tæ] 'he said; he grasped', [bəpɵrˈsiæ] ~ [bɛpo̜rˈsiæ] 'he asked'.

As the examples above reveal, /ö/ is generally derived from historical short *u. Additionally, the sound may have developed via a transitional *u (see §2.2.10), as in [dɵ] ~ [dʊˑ] 'two' (< *dō), [tɵ] ~ [tʊ] 'you' (< *tō), [kʊˈtɐ] ~ [kɵˈtaˑ] 'child' (cf. Pers. kōdak), [mɵdʒiˑˈlæ] 'ant' (probably cognate with Pers. mōr-ča, with metathesis), [mɵ] ~ [mo̜] ~[mʊ] 'I' (< *mun < *man). There are also words that do not follow such derivation: *sö* [sɵ] 'three' (< *sē),[5] [bïjo̜rda] 'he brought' (< *ār- < *āwar-), and these words of obscure etymon: [qɵˈʃɑˑr] 'pressure', [kʊˈlɛʃ] 'cough'.[6]

2.2.10. /ü/ stands for the allophones [ʉ] and [u], of which the latter is usually heard somewhat longer. Morgenstierne's (1960: 94) hypothesis that this Šahmirzādi vowel (ũ in his transcription) is a development of the old *majhul* vowel[7] ō is supported by my data: *ü* [ʉ] ~ [uˑ] 'he, she', *güš* [gʉˑʃ] 'ear', *güšt* [gʉʃt] 'meat', *püss* [pʉˑss] 'skin', *rüγɛn* [rʉˈʁɛn] ~ [ruˑˈʁən] 'ghee', *süt-* 'burn' (in [bæˈsʉtˑæ] 'it burned'), etc. There are however numerous words in which /ü/ stems from the *maʿruf* vowel ū, e.g., *bâzü* [bɑˑˈzʉ] 'arm' (< bāzūk), *zânü* [zɒˑˈnʉ] 'knee' (<zānūk), *sülâx* 'hole' (< sūrāx), *ma:lüm* 'apparent' (< maʿlūm). Moreover, the multiple sources of /ü/ are evident from *süzi* 'grass' (< sabz), [guˑ] 'cow' (< *gāw), [zuˑn] 'tongue' (cf. Mid. Pers. uzwān, Parth. izβān), among other words. All these sound changes suggest a merger of ō and ū, first into ŭ, then fronted but with an uncompleted outcome (see §2.2.8). Note also [kuˑl] 'bark of trees' (also in Mazandarani) and *lülü/lüli* (Žuk. *lölü*) 'dress, clothes', whose etymons are unknown to me.

2.2.10.1. Near-close near-front rounded vowel [ʏ] may be considered another allophone of /ü/, as it is heard in *süzi* [sʏˈziˑ] ~ [sʉˈziˑ] 'grass', *sülâx* [sʏˈlɑx] ~ [sʉˈlɑx] 'hole', *nü* [nʏ] ~ [nʉ] 'new'. Note also *nö(h)* [nøh] ~ [nʏh] 'nine'.

2.2.11. Diachrony. Chart 1 exhibits a general hypothesis for vowel diachrony in Šahmirzādi. The upper row is the hypothetical vowel inventory of Middle West Iranian, inferred from Middle Persian and Parthian, the only two known West Iranian languages of the period.[8] Middle West Iranian is taken here as a historical frame of reference in the absence of any known ancestor to the Caspian language group. Only major developments are shown;

5. The vowel in 'three' can be ascribed to analogy when counting (*do*, *so* . . .; *dö*, *sö* . . .). This sound change is parallel to h arising on *hašt* 'eight' by analogy with *haft* 'seven' (pers. comm. Erik Anonby).

6. See also §2.2.10.1.

7. On *majhul* and *maʿruf* vowels, see Perry 1996.

8. See Skjærvø 2009: 200.

so, details such as *āN > uN, as in *juma* ‘shirt’, are ignored for simplicity. Accordingly, Šahmirzādi differs from Persian essentially in keeping the front *majhul* *ē and in fronting of round back vowels. Noteworthy is the double step in the fronting of *ū, that is, halfway to /ü/ in some words (§2.2.10) and fully to /i/ in others (§2.2.3). In this respect, the distribution of Šahmirzādi words carrying /ü/ and /i/ from *ū accords well with those of Mazandarani proper, for which outcomes are /u/ and /i/; this can be seen in Šahm. *güš* and *di* versus Maz. *guš* and *did* for ‘ear’ and ‘smoke’.

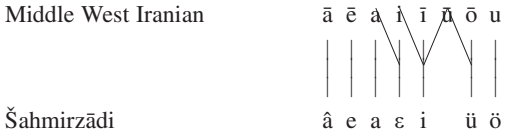

Chart 1. Vowel Diachrony

2.3. *Stress*. The stress is word final in nouns, with just a few exceptions following the norm in Persian, in adverbs such as *xálɛ* ‘very’. Verb forms appear to follow the stress patterns of Mazandarani proper (see Borjian 2005): the stress is on the verbal prefix in the imperative (*báxor* ‘eat!’) and the subjunctive (*báxori* ‘that you eat’); penultimate in the present-future (*xorémma* ‘I eat’); on the final syllable of the past stem (*ba-rɛkkí-a* ‘he scratched’, *da-pát-a* ‘he threw’); and on the nominal component (*vá dakɛta* [lit. wind (n.) it-fell] ‘it swelled’, *sâzé bɛzia* ‘he swept’). The negation marker (§4.7) always takes the primary stress.

3. NOUN PHRASE

3.1. *Number*. The common plural marker is *-(h)á*, as in *siuhâ* ‘apples’ and *karg(h)â* ‘hens’, while *-ón* is attested in *zanon* ‘women’. In Purkarim’s data only, the plural marker is predominantly *-un*, e.g., *miun* ‘hairs’, *čašun* ‘eyes’, *dârun* ‘trees’.

It should be noted however that plurality is seldom specified overtly in Šahmirzādi, as is the case in the Mazandarani variety of Dodānga. Singular forms are habitually used for the plural especially when the context takes care of quantity. A case in point is illustrated in the following Šahmirzādi sentence, followed by the Persian one it was evoked from:

Šahm.	*angir-∅*	*hama*	*širin=ena*	
	grape-SG	all	sweet=COP.3PL	
Pers.	*angur-hâ*		*širin*	*ast*
	grape-PL		sweet	COP.3SG

“The grapes are sweet.”

Noteworthy in the above sentence is the association between the verb and the subject in terms of number, contrasting with the Persian model, where the third singular copula serves the plural inanimate subject. Nevertheless, Šahmirzādi allows Persian style copulative sentences:

in	*kuš-â*	*mé=a*
this	shoe-PL	my=COP.3SG

“These shoes are mine.”

3.2. *Modifiers*. Adjectives and noun modifiers precede the head noun, separated by the marker *-e* (with free variant *-ə*). This noun-phrase binder is designated by Donald Stilo

Table 2. Personal Pronouns

	Subject	Object	Possessive
Sg. 1	*mö*	*már(a)*	*me*
2	*tö*	*tár(a)*	*te*
3	*ü*	*ür(a)*	*e*
Pl. 1	*(h)amá*	*(h)amára*	*áme*, *hámi*
2	*šamá*	*šamára*	*šáme*, *šámi*
3	*un(h)á*, *öšön*	*unára*, *öšönra*	*öšön(e)*

(2001) as "inverse-eżāfa," on the grounds that the noun phrase structures in Caspian and Persian are the inverse of each other. Examples: possessives: *böz-e sar* "goat's head," *bözâ-ye sar* "goats' heads"; adjectives: *sörx-e göl* "red flower," *sörx-e gölâ* "red flowers." For objects of postpositions, see §3.4.

3.3. *Pronouns*. Personal pronouns are declined in the ternary system typical of Caspian languages (Stilo 2001), that is, subject, object, and possessive cases. Table 2 demonstrates the most typical of the forms documented.[9] It is worth mentioning that, following the Caspian pattern, there are no pronominal clitics in Šahmirzādi.

The object set is formed by suffixing the object marker *-ra* (§3.5) to the subject set; the vowel sound of the object marker is liable to elision, while *-r* is always present. Example:

mö	*üra*	*tar*	*dɛmma*
1SG.SBJ	3SG.OBJ	2SG.OBJ	I.gave

"I will give it to you."

The possessive set *me*, *te*, *e*, etc.,[10] precedes the head noun when used as adjective (e.g., *me detar* "my daughter," *e ɛsm* "his name") and precedes the copular verb when used as pronoun, as in *in kušâ mé-a* "these shoes are (lit. is) mine," *in ti-a?* "is this yours?"; cf. the interrogative *in kušâ káni-a?* "whose shoes are these?"

3.3.1. Demonstrative pronouns are proximate *in* (object *ine*/*inna*, plural *inhâ*) and distal *un* (object *une*/*unna*, plural *unâ*). Demonstrative adjectives have no plural.

3.3.2. The reflexive is impersonal *xošten*[11] with the variant *xoš* in the third person singular. These words may function as (1) reflexive pronouns: *xošténn-a makelâšte* "he scratched himself," Dabir. *mö xöšten-æ ayne-dælæ bæ-di-mæ* (I self-OBJ mirror-in PFV-see.PST-1SG) "I saw myself in the mirror"; (2) possessives: Žuk. *xūš zanrō bogóta* "he told his wife," Pur. *mö ü-ræ hagetɛmæ xoštɛn-e düs-jɛn* "I took it from my friend"; Kalb. *xoš duš-e dim vešta* "he put it on his shoulder," Kalb. *xoštan-e fekr-e dele* "in his thoughts."

3.3.3. Relative and interrogative pronouns and adverbs are *ke*/*ki* 'who', *köja*/*köjâ* 'where', *čači*/*čeči* 'what', *čokâ* 'why', *čati* 'how', *čan* 'how many'. Demonstrative adverbs are *inja* 'here' and *unja* 'there' (see also §3.3.1).

9. For variation in sources, see Rastorgueva and Edel'man 1982: 509.

10. The possessive forms are likely to have emerged from the collision of the subject forms and the marker *-e* (Borjian 2005), for which see §3.2.

11. See §5.5.

3.4. *Adpositions*. Postpositions are the norm in Šahmirzādi[12] as in other Caspian varieties.[13] Common postpositions are *v(e)ra*, *ev* 'for', *dɛla*, *danim* 'in', *jɛn* 'from', *hamrâ*, *bâ* '(along) with' (comitative), *vari*, *bâ* 'with, by' (instrumental), *dim*, *sar* 'on', *jer*, *beni* 'under', *vâri* 'like'. This inventory of postpositions is in general agreement with that of Mazandarani proper, save for *bâ*, which is rare in Mazandarani. Also uncommon is *danim* (<? dar-miān). Note also the single occurrence *vista* in Chris. *jæk kotă̆ki-vistä* "about a child." Objects of postpositions normally receive the marker -e (cf. §3.2) when they are not vowel-final. Examples:

Kalb. *zan me-vra kota urana* "the wife will bring me children"
Dabir. *kətab-ev jæld bæxær* "buy a cover for the book"
Pur. *værf aftab-e delæ ū bünæ* "snow melts in the sun"
Dabir. *mö xöšten-æ ayne-dælæ bædimæ* "I saw myself in the mirror"
Kalb. *mö masjed-e danim davema* "I was in the mosque"
Kalb. *e das goma jan belaγziya* "the jar slipped off his hand"[14]
Pur. *köje jɛn ani* "where are you coming from?"
Pur. *ü bɛn(n)eke jɛn jēr ketæ* "he has fallen down off the roof"
Dabir. *me-jæn boləntær* "taller than me"
Kalb. *me hamrâ* "along with me"
Pur. *e-ba čekar hakuni?* "what will you do with it?"
Chris. *bäqqă̆l übå šüxĭ́ håkắrda* "the grocer joked with him"
Dabir. *u dær-ræ kæli-væri va hakərdæ* "he opened the door with the key"
Dabir. *mö divar-e-dim tekye hædamæ* "I leaned against the wall"
Kalb. *zamin-e dim* "on the ground"
Pur. *ta-ræ masürɛ-i dim depitɛn* "to spin yarn on (around) a spindle"
zamin-e sar "on the floor"
Kalb. *karg-e jer* "under the hen"
Pur. *mö benīštemæ yɛk dare bun* "I sat under a tree"
Chris. *te sár me sàre vorĭ́ kal bävéa* "your head is bald like my head"
Pur. *sölp-e vari sɛŋgin-e* "it is as heavy as lead"

3.5. The object marker for the accusative and dative is likely to be derived from the Persian morpheme *-rā*[15] in the forms *-ra*, *-a*, *-r*.[16] Examples of usage:

un miš-ra hišin "look at that mouse!"
Kalb. *mard-a bogota* "he told the man"
Dabir. *æma-r xævər dannæ* "they will let us know"
Pur. *mær xu giren(n)æ* (lit. for me the sleep grasps) "I am (becoming) sleepy"
mar sard-a (for Pers. *sard-am ast*) "I am cold"

12. Notwithstanding the preponderance of prepositions in Christensen's list (1935: 170–71), e.g., *äz ság tarsámmam* "I fear dogs."
13. See, i.a., Yoshie 1996: 22 for the Mazandarani dialect of Sāri; Stilo 2001 for Gilaki.
14. Note that *jan* should logically follow *e das*.
15. For the emergence of -rā in Persian, see Lazard 1970.
16. See also §3.3 and Table 2.

4. VERB PHRASE

4.1. *Verb stems*. Many Šahmirzādi past stems show irregular synchronic accord with their present pairs; examples are (present stem : past stem) *vin-* : *di-* 'see', *â-* : *ame-* 'come',[17] *ruš-* : *rut-* 'sell', *el-* : *ešt-* 'put, allow'. Regular past stems are synchronically derivable by adding a past formant to the present stem, as in *vâf-* : *vâf-t-* 'weave', *xor-* : *xor-d-* 'eat', *kan-* : *kan-i-* 'dig', *mun-* : *mun-ess-* 'stay'.

4.1.1. Among the aforementioned past stem formants, *-i-* (<*-īd-) is productive in the causative construction: causative present and past stems are formed by adding to the intransitive present stem the causative morphemes *-ɛn-* (< *-ānd-) and *-ɛnni-* (< *-ānd-īd-) respectively, as in *xos-* : *xot-* 'sleep' vs. *xosɛn-* : *xosɛnni-* 'put to bed'; *xor-* : *xord-* 'eat' vs. *xorɛn-* : *xorɛnni-* 'feed'.

4.2. *Preverbs*. Lexicalized preverbs are *(h)â-*, *da-*, *ve-*. They occur in *hâ-kon-* : *-kard-* 'do', *hâ-gir-* : *-gat-* 'seize'; *da-van-* : *-vass-* 'close; get firm', *da-pič-* : *-pit-* 'twist, fold', *da-kon-* : *-kard-* 'spill, shed; dress', *da-pač-* : *-patɛnni*[18] 'throw'; *ve-gir-* : *-gat-* 'pick up', *v-el-* : *-ešt-* 'put', etc. It is likely that *hi-šin-* : *-iši-* 'look at' carry a preverb similar to *hâ-* in Sāravi Mazandarani present stem *hā-r-eš-* 'id.' (with epenthetic -r-). Note also *de-pât-* 'pour' vs. *va-pât-* 'winnow', which may correspond to Lasg. *bepoton* 'scatter', Maz. of Espivard *he-pâj-* : *-pât-* 'sift (the threshed harvest)'. Žukovskij documented *de-pož-* : *-pot-* for Pers. *rixtan*, Russian *prolivat'sja* 'fill, pour'; its present stem might be either *pâj-* or *pâš-* (on account of absence of /ž/ in Šahmirzādi), the latter of which is comparable with Pers. *pāš-* 'scatter' (<Iranian root *parš; Cheung 2007: 298; Ḥasandust 2010: 611).

4.2.1. Preverbs preclude the modal-aspectual prefix *ba-* (§4.3). They normally accompany the stem in all tenses, including the present indicative, e.g., *hâkúmma* "I do" (cf. Sāravi *kəmbə* 'id.'), a feature that separates Šahmirzādi from Mazandarani proper. See more in §4.3.1.

4.3. *Modal and aspectual affixes*. The modal-aspectual prefix *ba-* (with varying vowel) marks the imperative, subjunctive present, preterite, past participle, and infinitive. The imperfective aspect is asymmetric in the present and the past, as discussed in the following sections.

4.3.1. The durative past is marked by *ma-* (with varying vowel), prefixed to the stem (or to the negative maker; see §4.7 below): *mo-got-a* "he used to say, he would say," *hâ-ma-kard-a* "he used to do," Žuk. *de-mī-pī́t-uma* "I would turn," Chris. *har váqt ke bäqā́l xunā́ mäšéa*, *un tūtī́ däkún-rä mäpā́sä*[19] "whenever the grocer went home, the parrot would look after the store." This past imperfective marker, apparently originating in Persian[20] and passed on to the Semnān area, marks a significant isogloss between Šahmirzādi and Mazandarani.[21]

4.3.2. The present indicative employs a nasal morpheme, *-(V)n-* or *-(V)m-*, wedged between the stem and the ending, as in *dâr-ɛn-a* 'she gives/will give birth to', *xar-ɛm-ma* 'I (will) buy', *de-m-ma* 'I (will) give'. Should this infix be a remnant of Old Iranian present

17. These two stems are historically suppletive.

18. The past stem is causative in form: *pat-ɛnni-*. See §4.1.1. Synonymous verbs for 'to throw' are *döm badâən* and *jer-ɛngiən.*

19. The conjugations in this sentence are ambiguous. The second verb has the stem *pâs-*, corresponding to Pers. *pādan*, *pāyidan* 'to guard'.

20. See Lazard 1963: 297.

21. The imperfective prefix me- is used for both the present and the imperfect in the Perso-Tabaric dialects that surround Mazandaran. See Borjian 2013b, §3.5.1.

Table 3. Verb Endings

	Indicative	Subjunctive
Sg. 1	*-ma*	*-(a)m*
2	*-i*	*-i*
3	*-a*	*-a*
Pl. 1	*-mi*	*-im*
2	*-ni*	*-in*
3	*-na*	*-an*

participle marker *-ant-,[22] the original *n* sound of the morpheme has been assimilated to the *m* sound of the first person ending (Table 3). From a synchronic vantage point, this morpheme may be taken as a part of present-tense personal endings, to form the distinctive set *-(ɛm)ma*, *-(ɛ)ni*, *-(ɛ)na*, *-(ɛ)mmi*, *-(ɛ)nni*, *-(ɛ)nna* (cf. Christensen 1935: 149; Morgenstierne 1960: 108; Borjian 2005). Alternatively, Dabir-Moqaddam (2014: 1045) regards the infix as an integral part of the present stems, e.g., *šün-*/*šüm-* 'go'. See also §4.4.1.

4.4. *Person markers*. Verb personal endings, shown in Table 3, are of two types: (1) the indicative set applicable to the present, the preterite, and the imperfect, and (2) the subjunctive set specific to the subjunctive present.[23] An epenthetic vowel (*ɛ* or *ə* or *a*) may stand between the stem and ending. Christensen (1935: 149) has an alternative, fuller set of indicative endings (first person singular *-mam*, etc.), apparently a result of mingling with Persian verb endings.

4.4.1. In the present indicative, the first person singular ending may contract or drop entirely, as in *xor-ém-ma* → *xoréma*, *xorém* 'I eat'; yet, the outcome stays distinct from any other person or tense. Likewise, the first person plural forms optionally conjugate without nasal germination, as in *xor-ém-mi* → *xorémi* 'we eat'. Moreover, stems ending in /n/ lose it in the second and third person conjugations to avoid gemination of the nasal in the singular, hence keeping the singular and plural distinct; for instance, present stem *zen-* 'hit' yields third singular *ze-n-a* versus third plural *ze-n-na.* Stems ending in /r/ normally lose or assimilate it depending on the person marker, as does the stem *dâr-*, yielding *dâ-n-a* 'he has', *dâ-n-na* 'they have', *dâ-m(-ma)* 'I have'. However, this rule does not hold under the condition that a merger with another verb would be possible; for instance, by resisting contraction *šur-ɛm-mi* 'we wash' stands distinct from *šu-m-mi* 'we go', as does *xor-ɛn-na* 'they eat' from *xon-ɛn-na* 'they read'.[24] See also §4.5.

4.4.2. The imperative takes no ending for the singular: *bɛ-ruš* 'sell!' *bi-rij* 'flee!' *ba-paj* 'cook!' *bu-šu* 'go!' *bɛu* (← bɛ-gu) 'say!' *hây* (← hâ-gir) 'seize!'; note the irregular form *buru* 'come!' (with suppletive stems *â-* : *ame-*). The plural imperative takes the ending *-in*.

4.5. *Tense*, *aspect*, *mood*. Simple verb forms consist of the imperative, present indicative and subjunctive, preterite (simple past), and imperfect (Table 4). There is no present perfect in Šahmirzādi; preterite is used in its place.

The two inflective past forms share the stem (past) (§4.1) and the ending (indicative) (Table 3); the distinction between the simple and continuous past is made by aspect markers (§4.3): perfective *ba-* and imperfective *ma-*. See Table 4 for paradigms.

22. See Azami and Windfuhr 1972, p. 198.
23. See also Table 6 for copulas.
24. Cf. Žuk., Chris. *xun-en-na*.

Table 4. Conjugation of "eat"

	Present		Past	
	Indicative	Subjunctive	Simple	Continuous
Sg. 1	*xor-ém-ma*	*bá-xor-am*	*bo-xórd-ɛma*	*mo-xórd-ɛma*
2	*xor-ɛn-i*	*ba-xor-i*	*bo-xord-i*	*mo-xord-i*
3	*xor-ɛn-a*	*ba-xor-a*	*bo-xord-a*	*mo-xord-a*
Pl. 1	*xor-ɛm-mi*	*ba-xor-im*	*bo-xord-ɛmi*	*mo-xord-ɛmi*
2	*xor-ɛn-ni*	*ba-xor-in*	*bo-xord-ɛni*	*mo-xord-ɛni*
3	*xor-ɛn-na*	*ba-xor-an*	*bo-xord-ɛna*	*mo-xord-ɛna*

Table 5. Verb Forms
(for the 3rd person singular)

Present	*xorɛ́na*
Subjunctive	*baxora*
Imperative (2 sg.)	*báxor!*
Preterite	*boxórda*
Imperfect	*moxórda*
Pres. progressive	*dar(a) xorɛna*
Past progressive	*dave moxorda*
Pluperfect	*boxórd be*
Perfect subjunctive	*boxórd bö*

The present-tense forms are divided by the mood. The indicative covers habitual as well as future. It is formally distinguished from the subjunctive by two morphemes: the ending set (Table 3) and the mood-aspect markers (§4.3), that is, the subjunctive prefix *ba-* and the durative infix *-en-*. This redundancy permits a considerable degree of leniency in the conjugation of the present indicative tense, as explained under Person markers above (§4.4.1). See Table 5.

4.6. *Periphrastic forms*. These include two perfective and two progressive tenses which are constructed analytically with substantive and locative verbs (§4.8) as auxiliary.

4.6.1. The pluperfect and the perfect subjunctive employ the past participle (§4.10.1) with the past and subjunctive of 'be' (Table 6) respectively, as in *bevőrd bena* "they had taken away," *vagét bâm*"'I may have picked up," *mö ke baresima ü bašé be* "he had gone when I arrived," *age ü bašé bö ma:lüm büne* "should he be gone, it will be known."

4.6.2. The progressives receive the imperfective conjugation preceded by the present or past of the locative verb (§4.8); the present locative optionally remains unconjugated. Examples: Pur. *jü dær res(s)ene* "the barley is ripening," Dabir. *dær(-enæ) næsihət ha-ku-n-næ* "they are giving advice," *dæv-enæ næsihət ha-me-kərd-enæ* "they were giving advice," Kalb. *davema kota-ra mogotam*[a] "I was telling the child."

4.7. *Negation*. The negative marker *ná-* precludes the modal prefix *ba-* and the preverb *hâ-* and succeeds that imperfect marker *ma-*. Examples: *güš me-né-kerd-a* "he wouldn't listen," Žuk. *me-ná-xōrd-ī* "you wouldn't eat," Chris. *jävv̊b mänádå* "he wouldn't answer" (p. 174). Prohibition is marked by either *na-* or *ma-*, as in *nákon* "don't do!" *mékelâšta* "don't scratch [yourself]!" The second form is apparently older, corresponding to classical Persian ma-.

Table 6. Copulas

	Present	Subjunctive	Past
Sg. 1	*-em(a)*	*bâm*	*béma*
2	*-i*	*bi*	*bi*
3	*-e*/*-a*	*bö*/*bü*	*be(a)*/*bia*
Pl. 1	*-em(i)*	*bem*/*bim*	*bemi*
2	*-en(i)*	*ben*/*bin*	*beni*
3	*-en(a)*	*ben*/*ban*	*bena*

4.8. *Copulative and locative verbs*. 'Be' is expressed by two verbs: copulas and locative-existential. The present copula equals the indicative set of person markers (Table 3), optionally on the base *(h)as(s)-*. The subjunctive and the past are built on the stem *b(e)-* (Table 6). The locative-existential verb is formed with the stems *dar-* (present) and *dav-* (past). These two verbs function as auxiliary in periphrastic verb forms (§4.6).

Examples: Copula: *šu târik-a* "night is dark," *mo Ali-e zumâ hass-em(a)* "I am the son-in-law of Ali," Morgen. *ü mahin-ase* "he is big," (past) Chris. *širin-zäbún béa* "he was sweet-tongued," (subjunctive) *har jâ dara*, *tandoros bü* "wherever he is, may he be in good health." Locative: *me inja dár-ema* "I am here," Pur. *pirezän yɛke därä* "the old woman lives alone," *kojâ dav-i?* "where were you?" Kalb. *e dele mâst davəa* "there was yogurt in it," *hamin yag šu-ra inja daven* "be ye here this very night," *dumma har-jâ davü* (Pers. *bāšad*), *tandoros hassa* "I know that wherever he may be, he is in good health."

4.8.1. 'Become' is conjugated regularly on the present stem *bü-* (medially, *-vü-*) and the past stem *v-*. 'Be' and 'become' merge in the subjunctive. Examples: *xub büma*/*navüma* "I will become/not become well," *be ommid-e xodâ xub bünni* "by God's grace you (pl.) will get well," *nâxoš bavéma* "I became sick," *battar me-v-e* "he used to get worse," Kalb. *marγona čitikâ büna . . . e čitikâhâ mahin bünna* "the egg becomes chick . . . her chicks will grow big." Counterfactual: Pur. *æge ū bæba müxærdemæ* "if there were water, I would have drunk [it]." More data are needed to arrive at a full paradigm for 'become'.

4.9. *Modal verbs*. The impersonal modals *vân(ə)* (pres.) and *vâs(ə)* (past) convey both meanings of 'must/should' and 'want'. They are succeeded by the present subjunctive of the main verb, as in *vân*/*vâs baxori* "you should eat/you should have eaten," Žuk. *vőna gusl hōkunĭ́m* "we must perform ablutions," Žuk. *vőssa ur azőb hőkunam* "I must castigate him," *in böz-ra vân davassan* "this goat must be tied" (with the infinitive); Dabir. *un mærdi van bē* "that man wants to come," Žuk. *tŭ́flo* (*tüfl*?) *mer névōna* (for Pers. *ṭefl-rā nemix*^{w}*āham*) "I don't want the child" (lit. for me the child is not a must), Pur. *ū̄ nævanæ dævi* "he doesn't want to stay (lit. to be in)," Pur. *ægær tær vānæ e-ba čekar hakuni?* "if you want it (lit. if it is desired by you), what will you do with it?" Pur. *vas(s)æ börim bæše* "he wanted to go out."

'Want' may alternatively be expressed by the verb *xâstan*, e.g., Chris. *ĭ́n kotắkrā̆ náxåmmam* "I don't want this child" (p. 174), Žuk. *ensőf bexőstena* "they requested judgment."

4.10. *Verbal nouns*. The infinitive is formed on the past stem (§4.1), prefixed by *bV-* or a preverb (§4.2) and suffixed by *-(a)n*, as in *bo-xord-an* 'to eat', *hâ-kard-an* 'to do', *ha-dâ-n* 'to give', *ba-zə-n* 'to hit', *bö-pörsi-n* 'to ask'.

4.10.1. The past participle is formed of the past stem and prefix *bV-*. It normally appears without a suffix when employed in perfect tenses (Table 5 and §4.6.1), but is suffixed by *-e*/*-a* when used as a participial adjective: Žuk. *besū̆té nun* "burned bread," Chris. *bäpetá* "cooked," Pur. *zæng-beze(y)æ* "rusted". The vowel *-e*/*-a* is normally elided when the stem

Table 7. Divergent Cognates

Gloss	Šahmirzādi	Semnāni	Old Iranian stem
big	*mahin*	*masin*	*masya-
son	*pesar*	*pir*	*puθra-
three	*sö*	*he(y)ra*	*θraya-
dog	*sak*	*esba*	*spa-ka-
yesterday	*adi*	*izi*	*zya-ka-
door	*dar*	*bar*	*dwar-
woman	*zan*	*jini*	*jani-
tomorrow	*fardâ*	*haren*	*fra-tanaka-

ends in a vowel, e.g., Pur. *jü bɛresi*$_1$-*a*$_2$ "the barley is$_2$ ripe(ned)$_1$." The boundary between the categories of periphrastic verb and copular clause tend to be untidy for the stems ending in vowels; for instance, the sentence *e dim juš baze bea* "his face was pimpled" (with the stems *zan-* : *zə-* 'hit, stricken') may contain either the compound adjective *juš-bazé* or the pluperfect *júš baze bea*, depending on the position of the stress.

4.10.2. The present participle can be the present stem only, usually suffixed to a noun, as in *kâr-tɛn* 'weaver', *aškɛm-dâr* (lit. having a belly) 'pregnant', *sarmâ-gir* (Pers. *sarmâ'i*; cf. French frileux) 'sensitive to the cold; one who is always cold'. Alternatively, the present participle may take the suffix *-ana* (< -anda), as in *gazana-mâz* 'biting wasp', *tarsɛna* (for Pers. *tarsu*) 'one who fears habitually, coward, sheepish'. Note also *angir-bačin*, for Pers. *angur-čini* 'grape harvesting'.

5. LEXIS

The lexical inventory of Šahmirzādi is in close agreement with that of Mazandarani proper, but not without a certain degree of convergence towards its neighbors in the Semnān area,[25] as well as its own idiosyncrasies.

5.1. *Caspian pedigree.* As a Caspian variety, Šahmirzādi is genetically distinct from all its neighbors, including Semnāni, Sorxa'i, Lāsgerdi, Aftari, and Sangesari. As such, cognate words may differ considerably among these languages. Table 7 illustrates Šahmirzādi and Semnāni words that share an etymon but diverge in form beyond recognition for an average speaker. Other languages of the region generally follow Semnāni in phonological development, while Sangesari shows deviation, as, for example, in the outcome *še* for the isogloss 'three'.

5.2. *Mazandarani inventory*. In order to arrive at an estimation of the lexical agreement between Šahmirzādi and Mazandarani proper, I used a list of basic vocabulary collected by Kolbitsch (2008). It is a Swadesh-type list of 227 items based on the modified list prepared by Anonby (2003) for the languages of Iran. Of the 227 Šahmirzādi items of the list, only 33 items (15 percent) show significant disagreement with Mazandarani proper. These are listed in Table 8. We may further break down the latter into the words with major phonological deviation and the words of different roots. The latter reckon 18 in the table, constituting 8 percent of the list.

25. Which in turn share a great deal of their vocabularies with the Mazandarani language.

Table 8. Šahmirzādi vs. Mazandarani Wordlist

Gloss	Šahmirzādi	Mazandarani	Notes
tongue	*zun*	*zəvun*	syncopated
throat	*gali*	*gal*	Šahm.=Semn.
finger	*angöšt*	*angus*	Šahm.=Semn.
knee	*karb*, *zânü*	*zânu*	
bone	*assɛγon*	*essekâ*	Šahm.=Semn.
man	*mard*	*mardi*	
woman	*zan*	*zənâ*	See Tab. 7
child	*kota(k)*	*vačə*	See §5.5
chick	*čitekâ*	*čindekâ*	See §5.5
snake	*döm-darâz*	*mar*	See §5.6
ant	*möjila*	*məlijə*	metathesis
spider	*haštpâ*	*ban(d)*	
leaf	*valg*	*gelâm*, *valg*	
cloud	*abr*	*miâ*	Šahm.=Semn.=Pers.
hill	*tappa*, *dima*	*tappə*, *din*	
dew	*hasâr*	*šebrə*	
river	*čâka*	*čukə*	See §5.3
house	*xuna*	*səre*	persianism
oil, ghee	*ruγɛn*	*râγun*	
one	*yɛk(ka)*	*əttə*	persianism
three	*sö*	*se*	See §2.2.9 and Tab. 7
good	*xub*	*xâr*	persianism
short [height]	*köl-γad*	*γad-pass*	
short [length]	*köl*	*kətâ*	
full	*pör*	*mašt*, *pər*	
dirty	*našt*	*lačər*	Šahm.=Semn.
swim	*šɛno*	*hasno*	
to give birth	*dârtan*	*bazâən*	See §5.5
to swell	*vâ dakɛtan*	*čaft dakətən*	
to suck	*bösoftan*, *sök bazɛan*	*bačefessən*	
to scrape, scratch	*barɛkkian*	*barəkessən*, *bakossən*	for Pers. *xārāndan*
to throw	*jerɛngian*, *döm badâan*	*dəm bədâən*	
yesterday	*adi*	*diruz*	See Tab. 7

5.3. The following words exemplify Šahmirzādi words of Mazandarani origins vs. their counterparts in Semnāni and other neighbors.[26]

bâmši ‘cat’ vs. Semn., Lasg., Sang. *RUVÂ*, Aft., Lasg., Sorx. *NÂZU*.
juma ‘shirt’ vs. Semn. *šavi*.
bad ‘bad’ vs. Semn. *pis*.

26. Small capital letters signify the commonest pronunciation within and across languages.

xurd 'small' vs. Semn., etc. *kasin*.
qül 'deaf' vs. Sang. *gič*, Semn. *kar.*
šir 'milk' vs. Semn., etc. *šat*.
dezɛnnun 'tripod' vs. Semn., etc. *SEPĀYA*, Sang. *šepâya.*
dovenni 'shoe' vs. Sang., Aft. *PEYJÂR*, Semn., Sorx., Lasg. *LAL(A)KA*.
lɛsek 'snail' vs. Sang. *gal-karm*, Semn. *gāzâlâ.*
küp 'straw mat' vs. Sang. *sis*; all others use Persian *hasir.*
rik 'gum' vs. Lasg. *zē(h)*, Semn., Aft., Sang. *GUŠTI DANDONI.*
teli 'thorn' vs. Semn., etc. *ask.*
čâka, *čâkepe* 'river' (Maz. *čukə*) vs. *RO(D)XÂNA* elsewhere in the region.

5.4. Aftari and Sangesari share with Šahmirzādi a pool of Mazandarani words that other neighboring languages do not. The Šahmirzādi–Sangesari–Aftari lexical union can be demonstrated by these words:

sölla 'gutter' agrees with Sang. *sölley*, vis-à-vis Semn. *nâlazóna* and similar words in the other dialects of the region.
huz 'basin' is shared by Sangesari and Aftari, and contrasts with *xut* in other dialects.
sâzɛ 'broom' is shared by Sangesari and Aftari, vs. *rīhuna* in Semn. and *JÂRU* in other dialects.
tannir 'tandoor, oven' has similar forms in Sangesari and Aftari, vs. *KELA* in other dialects.
kalin 'ashes' is similar to Sang., Aft. *KALIM*, vs. Semn., etc. *XAKESTAR.*
kâp, *kâb* 'heel' is shared by Sangesari and Aftari, vs. Semn., etc. *PÂŠNE.*
vak 'frog' is shared by Sangesari and Aftari, vs. Semn., etc. *VAZAQ.*
jöma 'Friday', also in Aftari; Semn., etc. *uyna* (<ādīna).
katâr 'chin', also in Sangesari, vs. *ČA(KU)NE*, etc. in other dialects.
vasni 'co-wife', also in Aftari vs. Semn. *dohamin jan(n)ia* and the likes in most other dialects.
qöšâr 'pressure'; cf. Sangesari and Firuzkuhi *qešâr*, as opposed to Sorx. *fešâr* and Lasg. *šexâr*.

5.5. *Areal features*. Šahmirzādi's adaptation to the neighboring languages of Semnān zone can be illustrated in the following words.

kota(k) 'child' agrees with *k^{u}ota* in Sangesari and contrasts with *vačə* in Mazandarani and the rest of the Semnān area.
lülü/lüli (Žuk. *lölü*) 'dress, clothes' has similar forms in Sangesari and Aftari, vs. Semn., Sorx., Lasg. *HALA*.
dâr- : *dârt-/dâšt-* 'give birth to', also in Sangesari.
miš 'mouse'; cf. Semn. *müš* and the likes in other dialects vs. Maz. *gal.*
marγija 'sparrow', similar to Aft., Semn., Sorx., Sang. *MAR(GU)JA* vs. Maz. *MIČKÂ.*
xuz 'walnut', corresponds to Sorx. *hüž*, Sang. *yuz*, vs. Maz. *âγuz.*
vašir 'unsalted, tasteless',[27] with similar forms in Semn., Lasg., Sorx., Sang. vs. Maz. *be-nəmək.*

27. Sotuda 1963 glosses 'unsalted' only.

huften "weave" (along with *vâftən*), cf. Semn. *hönattion*, Sorx. *hâvâftɔn*, Lasg. *ōveton*, Aft. *hɔvaton*, Sang. *hövetan*, vs. Maz. *boftən.*

xošten 'self', also used in Semnāni proper, as *xošton*, and its varieties throughout the region: Semn. (also) *huštara*, Sorx. *hušterâ*, Lasg. *ušterâ*, Sang. *eštere*. Note the diffusion of this type into southern and eastern frontiers of Mazandarani, as *xaštən*, contrasting with common Mazandarani *še*.

An instance of hybridization across language zones is Šahmirzādi *čitikâ* 'chicken', standing halfway between Maz. *čin(n)ekâ* and Semn. *čuta*, Aftari *čute*.

5.6. *Idiosyncrasy*. I have come across several words in Šahmirzādi that are neither Mazandarani nor Semnāni. Most noticeable is *mahin* 'big', which contrasts with both Maz. *gat* and Semn., Sorx., Aft., Sang. *masin* (see Table 7). The Šahmirzādi form occurs also in archaic New Persian, thus could have spread out westward from Khorasan, although my attempts to find this word in the current Khorasani dialects of Persian were in vain. The only occurrence of *mahin* I could find in Mazandarani documents is in a verse attributed to Amir Pāzvāri, collected orally in the nineteenth century: *man botparastun-râ hama mahin bum* "I would be the biggest of idol worshippers" (Dorn 1860: 95); this incidence however seems to be mere archaism.

Other possible characteristic Šahmirzādi words are *šarša* 'rock'; *zököt* 'elbow', vs. Maz. *aleskin*, Semn. etc. *MARAKA*. Note also the circumlocution *döm-darâz* 'snake', lit. 'long-tail', which is an avoidance euphemism based on the belief that the subject can be invoked when its name is uttered.

6. LINGUISTIC POSITION

What warrants a detailed areal study of Šahmirzādi, as a Caspian variety, is its geographic position in a fairly small linguistic zone that hosts other Iranian linguistic types genetically distant from the Caspian language family. Besides Šahmirzādi, there are at least three language types spoken in Semnān district. While Semnāni proper and Sangesari can be classified as isolates within the Northwest Iranian family, the status of Sorxaʾi, Lāsgerdi, and Aftari as distinct languages or dialects of the same language is not yet established due to paucity of data (cf. Morgenstierne 1960, Lecoq 1989). Different language types notwithstanding, there can be seen a certain degree of convergence in the area that justifies the areal designation Semnān Sprachbund. Another designation I have used for this areal union is Komisenian, or Kumeši (Borjian 2008), after the historical name of Semnān province, to avoid confusion between the Sprachbund and Semnāni proper.

The striking similarities of Šahmirzādi with Mazandarani proper, as shown in this study, leave little doubt that Šahmirzādi has branched out from the bulk of Mazandarani, but a chronology is difficult to establish, given the lack of an historical record on Šahmirzād, let alone its vernacular. It is also not reasonable to arrive at a date of separation based on comparative linguistic analyses, for although Šahmirzādi is a linguistic outlier, it has always been in contact with Mazandarani, across the Alborz chain, through trade and seasonal migrations, a condition that must have decelerated the process of divergence.

Now let us consider the question, whether Šahmirzādi is a variety of Mazandarani or else has drifted away far enough from it to be classified a separate language directly under the Caspian family of languages? Existing opinions differ on this matter. While Šahmirzādi has been viewed as a distinct variety of the Mazandarani language by certain scholars (Rastorgueva and Edel'man 1982: 450; Lecoq 1989: 490), *Ethnologue* (2017) classifies Šahmirzādi (with language identifier srz) as a subgroup of the Caspian Language family, distinct from both Gilaki and Mazandarani, while *Glottolog* (2017, which treats all language varieties as "Lan-

guoids") classifies Caspian into Gilaki-Rudbari and Mazanderani-Shahmirzadi, and assigns shah1253 as the code for Šahmirzādi. Donald Stilo, an authority on the typology of Iranian languages, holds the view that there might be enough features that justify a separate classification for Šahmirzādi within the Caspian family (pers. comm.; see also Stilo 1981). Moreover, in a recent typological study of the language varieties surrounding Mazandarani proper, which incorporated ten grammatical features, Šahmirzādi differed from Mazandarani proper only in one feature, the imperfect (Borjian 2013b, Table 7; see also Borjian 2013a).

The imperfect is indeed the most salient morphological split of Šahmirzādi from its original speech community. The contrast, as shown under §4.3.1, leads to physiognomic differences such as Šahm. *mavâfta* vs. Maz. *boftə* 'he would weave', in which Šahmirzādi incorporates the durative prefix *mV-*. This marker, which results in structural asymmetry between the past and present in Šahmirzādi, could have been borrowed from either Semnāni or Persian; the marker in all likelihood originated in early New Persian through grammaticalization of the adverb *hamē*.[28] The functional item *mV-*, formidable as it looks, has proved to be highly contagious as it has permeated, in a span of less than a millennium, not only Semnāni but also Tāti, another Northwest Iranian language group genetically remote from Persian. As for Šahmirzādi, the marker may have been borrowed from Persian, the lingua franca of Semnān district.

Aside from the imperfect marker, another remarkable morphological feature that sets Šahmirzādi apart from Mazandarani is the preverb on the present (§4.2.1). In terms of phonology, the fronting of the round back vowels (§2.2.8), characteristic of the Semnān area but not Mazandarani proper, plays a far less important role in separating Šahmirzādi from its kin. Subsequently, phonological and morphological changes on the whole have not been profound enough to trigger syntactic changes.

On the lexis front, a broad comparative study of Šahmirzādi vocabulary is yet to emerge. The comparative basic vocabulary analysis above (§5.2) reveals that the lexical inventory of Šahmirzādi is at minimum in 85 percent agreement with Mazandarani, while some lexical leveling within the Semnān region is at work (§5.5), and there are striking idiosyncrasies (§5.6) to be explained by spreading out our areal net.

Returning to the question of the language or dialect status, we may draw a comparison between Šahmirzādi and the vernacular of Kalārestāq, which is spoken on the west of the Čālus river, and is classified by Stilo (2001) as a Central Caspian language different from both Gilaki and Mazandarani. In morphosyntax, Kalārestāqi shows more contrastive features with Mazandarani proper (Borjian 2010) than Šahmirzādi does, but certainly not in vocabulary, as far as it can be judged impressionistically in this preliminary survey. A detailed study of the Caspian family will make things clearer. Also wanting is a study of mutual intelligibility between Šahmirzādi and Mazandarani proper.

A classification can be defined and supported by linguistic data, but the status of a language variety as a "language" vs. "dialect" is relative and is almost always ultimately socially rather than linguistically defined, even if linguistic factors are important as well. Extralinguistic factors, such as ethnic and linguistic identities, are important in this regard. Šahmirzādis refer to themselves as *gelak* and to their language as *gelaki* "Caspian," as do the people of Mazandaran and Gilan. Many speakers of Šahmirzādi believe that their vernacular is merely a Mazandarani variety (close to the varieties spoken in Šāhi and Bābol but not as close to those of Firuzkuh and Sāri) and their culture is far closer to Mazandaran than their immediate neighbors. Recently Šahmirzādis have filed an official appeal on the grounds

28. For the imperfective prefix in Šāhrudi Persian, etc., see Borjian 2013b, §3.5.1.1.

of "culture and language" for their district to join the province of Mazandaran (Fars New Agency 2003).

It would also be interesting, if data were available, to compare Šahmirzādi with the Mazandarani dialects immediately to its north, to find out whether it is a true insular Caspian outlier or if there is some areal continuity. Further northeast in Parvar, as far as the poems of Gudarzi (2009) reveal, the dialect is not Šahmirzādi but close to the Mazandarani varieties spoken across the Alborz ridge in Pāji and Sangedeh,[29] the southernmost villages of the Dodānga, which constitute the uplands of Sāri (Fig. 1). The data are scanty, but already show certain shared features, such as *xošten* (§§3.3.2, 5.5), which is found also in Firuzkuhi. My informants told me that Šahmirzād used to have mutual trading with communities immediately to the north via mountainous trails as well as with Firuzkuh along the caravan road via Aftar and the Bašm pass.

Another venue to consider in a typological study is along the south Alborz belt that stretches from Semnān westward. Take for instance the Šahmirzādi word *bušu* 'go!' which is shared by the varieties spoken in the southern foothills of Alborz,[30] from as far east as Ṭāleqān down to Šemirān, and further east to Damāvand and the Semnān area, while Mazandarani proper, i.e., the varieties spoken on the east of Čālus river and north of the Anti-Alborz, use the irregular form *bur* exclusively. This wave-like distribution poses yet another challenge in the provenance and linguistic position of Šahmirzādi.

29. On which I have limited field notes.
30. Borjian 2013b.

ABBREVIATIONS

Small capital letters signify the commonest pronunciation within and across languages.
Colon (:) stands between the present and past stems of verbs.

Aft.	Aftari
Chris.	Christensen
Dabir.	Dabir-Moqaddam
Ir.	Iranian
Kalb.	Kalbāsi
Lasg.	Lāsgerdi
Maz.	Mazandarani
Morgen.	Morgenstierne
Pers.	Persian
Pur.	Purkarim
pl.	plural
Sang.	Sangesari
Semn.	Semnāni
sg.	singular
Sorx.	Sorxaʾi
Sotu.	Sotuda
Šahm.	Šahmirzādi
Žuk.	Žukovskij

REFERENCES

Anonby, E. J. 2003. Update on Luri: How Many Languages? *Journal of the Royal Asiatic Society* 13.2: 171–97.

Borjian, Habib. 2005. Šenāsa-ye feʿl dar māzandarāni-e šarqi. *Guyeš-šenâsi* 1.3: 13–19.

_____. 2008. The Komisenian Dialect of Aftar. *Archiv Orientální* 76.3: 379–416.

_____. 2010. Kalārestāq. In *Encyclopaedia Iranica* XV/4. Pp. 367–73. New York: Columbia Univ.

_____. 2013a. The Tabaroid Dialects of South-Central Alborz. *Acta Orientalia* 66.4: 427–41.

_____. 2013b. Perso-Tabaric Dialects in the Language Transition Zone Bordering Mazandaran. *Studia Iranica* 42.2: 195–225.

Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb*. Leiden: Brill.

Christensen, Arthur. 1935. *Contribution à la dialectologie Iranienne*, vol. 2: *Dialects de la région Sèmnān: sourkhéī*, *lāsguerdī*, *sängesärī*, *et chämerzådī*. Copenhagen: Andr. Høst & Søn.

Dabir-Moqaddam, Moḥammad. 2014. *Radašenâsi-e zabānhā-ye irāni*. 2 vols. Tehran: Samt.

Dorn, B. A. 1860. *Ketāb-e Kanz al-asrār-e Māzandarāni: Beiträge zur Kenntnis der Iranischen sprachen*, vol. 1: *Masandaranische Sprache.* St. Petersburg: Kaiserlichen Akademie der Wissenschaften.

Ethnologue: Languages of the World, 20th ed. 2017. Dallas: SIL International. Online at http://www.ethnologue.com/language/srz.

Fars News Agency (Xabargozāri-e Fārs). 2003. Taqāżā-ye šahmirzādihā barā́ye elḥāq-e be Māzandarān. 19 Šahrivar 1392 Š./9 September 2003; at farsnews.com/newstext.php?nn=13920618000832.

Glottolog 3.1. 2017. Jena: Max Planck Institute for the Science of Human History. Online at: http://glottolog.org/resource/languoid/id/shah1253.

Gudarzi, M.-R. 2009. *Sut o nâješ: Sorudahā o tarānahā-ye buni-e Parvar.* Semnān: Hablerud.

Ḥasandust, M. 2010. *Farhang-e taṭbiqi-mowżuʿi-e zabānhā o guyešhā-ye irāni-e now*. 2 vols. Tehran: Farhangestān.

Kalbāsi, Irān. 2009. *Farhang-e towṣifi-e gunahā-ye zabāni-e Irān*. Tehran: Pažuhešgāh-e ʿOlum-e Ensāni.

Kolbitsch, A. 2008. Unpublished audiotape eliciting a 227-item basic vocabulary from two Šahmirzādi speakers, Šahmirzād.

Lazard, Gilbert. 1963. *La langue des plus anciens monuments de la prose persane***.** Paris: Klincksieck.

_____. 1970. Etude quantitative de l'évolution d'un morphème: La postposition râ en persan. In *Mélanges Marcel Cohen*, ed. D. Cohen. Pp. 381–88. Berlin: De Gruyter.

Lecoq, P. 1989. Les dialectes caspiens et les dialectes du nord-ouest de l'Iran. In *Compendium linguarum iranicarum*, ed. Rüdiger Schmitt. Pp. 296–312. Wiesbaden: Reichert.

Morgenstierne, Georg. 1960. Stray Notes on Persian Dialects. *Norsk Tidsskrift for Sprogvidenskap* 19: 72–140.

Perry, John R. 1996. Persian in the Safavid Period: Sketch for an état de langue. In *Safavid Persia: The History and Politics of an Islamic Society*, ed. C. Melville. Pp. 269–83. London: I. B. Tauris & Company.

Rastorgueva, V. S., and D. I. Èdel'man. 1982. Prikaspijskie jazyki: giljanskij, mazanderanskij (s dialektami Šemerzadi i Velatru). In *Osnovy iranskovo Yazykoznanija*, vol. 3: *Novoirankie Yazyki: Zapadnaja gruppa*, *prikaspijskie jazyki*, ed. V. S. Rastorgueva. Pp. 447–54. Moscow: Soviet Academy of Science.

Skjærvø, Prods Oktor. 2009. Middle West Iranian. In *The Iranian Languages*, ed. Gernot Windfuhr. Pp. 196–277. London: Routledge.

Sotuda, Manučehr. 1963. *Farhang-e semnāni*, *sorxa'i*, *lāsgerdi*, *sangesari*, *šahmirzādi*. Tehran: Tehran Univ. Press. (Rpt. 1977.)

Stilo, Donald. 1981. The Tati Language Group in the Sociolinguistic Context of Northwestern Iran and Transcaucasia. *Iranian Studies* 14: 137–87.

_____. 2001. Gilān x. Languages. In *Encyclopaedia Iranica*, vol. 10, fasc. 6. Pp. 660–68. New York: Columbia Univ.

Windfuhr, Gernot, and John R. Perry. Persian and Tajik. 2009. In *The Iranian Languages*, ed. Gernot Windfuhr. Pp. 416–44. London: Routledge.

Yoshie, Satoko. 1996. *The Sāri Dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Žukovskij, Valentin A. 1922. *Materialy dlija izučenija persidskix narečij*, vol. 2. Petrograd: Russian Academy of Science.