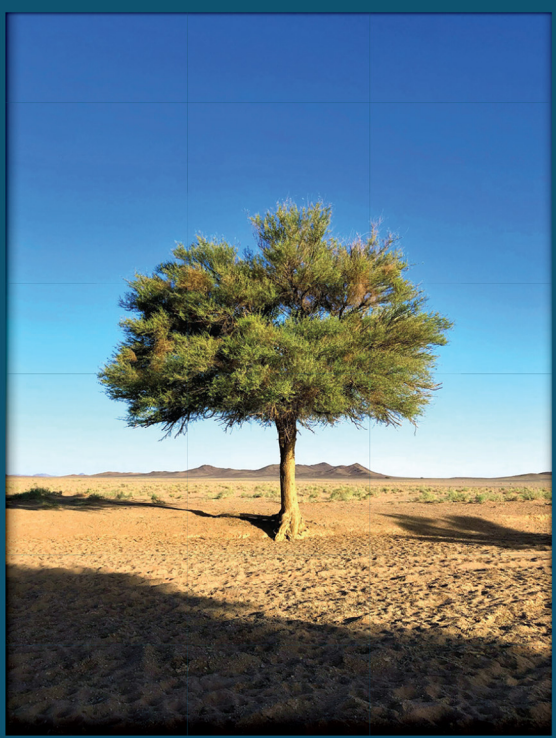
Cambridge Semitic Languages and Cultures

Diversity across the Arabian Peninsula
Language, Culture, Nature

Edited by Fabio Gasparini, Kamala Russell and Janet C. E. Watson



UNIVERSITY OF CAMBRIDGE
Faculty of Asian and Middle Eastern Studies

https://www.openbookpublishers.com

Fabio Gasparini, Kamala Russell and Janet C. E. Watson, *Diversity across the Arabian Peninsula: Language, Culture, Nature.* Cambridge, UK: Open Book Publishers, 2024, https://doi.org/10.11647/OBP.0411

Semitic Languages and Cultures 28

ISSN (print): 2632-6906

ISSN (digital): 2632-6914

ISBN Paperback: 978-1-80511-337-9

ISBN Hardback: 978-1-80511-338-6

ISBN Digital (PDF): 978-1-80511-339-3

DOI: 10.11647/OBP.0411

Cover image: Photo by Rabah Al Shammary, titled 'Wild Acacia tree, Ha'il, Arabian Peninsula', July 26, 2021; https://unsplash.com/photos/green-tree-on-brown-sand-under-blue-sky-during-daytime-e-UPgjjEwCM?utm_content=creditCopyText&utm_medium=referral&utm_source=unsplash.
Cover design: Jeevanjot Kaur Nagpal

The main fonts used in this volume are Charis SIL and Scheherazade New.

VERBAL NOUN FORMATION IN MEHRI

Anton Kungl

1.0. Introduction

Most Semitic languages show a specific category of deverbal nouns, usually referred to as 'Verbal Nouns' (henceforth VN).[1] Whereas the precise semantic and morphosyntactic properties of VNs diverge across languages, for the purpose of this study the minimal definition of VNs in Mehri will be taken from Morris (1981, 251): "[VNs being]... gerunds or nouns describing the action of the verb, they have no plural forms" (Morris 1981, 251).

[1] Ar. – Arabic (Classical Arabic/Modern Standard Arabic), EM – Eastern Mehri, IG – Idle Glottis (consonant), *JL* – *Jibbali Lexicon* (Johnstone 1981), J/S – Jibbali/Sheḥri, *ML* – *Mehri Lexicon* (Johnstone 1987), MSAL – Modern South Arabian (Languages), PS – Proto-Semitic, PMSAL – Proto Modern South Arabian, Soq. – Soqotri, VN – verbal noun, WM – Western Mehri.

The terms 'Eastern Mehri' and 'Western Mehri' are used here as shorthand for the varieties described in Johnstone (1987)/Morris (1981) and Jahn (1902) respectively. As is well known, the dialectal distinctions of Mehri do not correspond to an exact East–West divide, nor to the modern-day borders between Yemen and Oman, however, since a precise account of the different varieties of Mehri and their geographical and sociological (tribal/urban) contexts has not been provided as of yet, these—descriptively defective—terms are used in the present study.

 https://doi.org/10.11647/OBP.0411.06

The exact meaning of VNs is dependent on the semantics of the corresponding root, and not all verbal roots necessarily have corresponding VNs, at least not in accordance with the definition given above. Also, an overlap between VNs and other nominal forms, sharing the same root with a given verbal form, is to be expected, particularly in the G-stems.

In this study, a descriptive overview of VN formation in Mehri will be given, together with an investigation of certain pertinent phonological and morphological features. Previous descriptions of Mehri, with the exception of the aforementioned thesis by Morris (1981), make no or only passing mention of VNs. Looking to other Semitic languages, one would expect derived-stem VNs to display more regular patterns, while G-stem VNs would show quite divergent patterns.

As will be shown in this paper, this situation is also broadly so in Mehri, and most likely in all MSAL. However, the parameters of VN formation in derived stems, as well as some prominent VN patterns for G-stem verbs, are quite distinct from patterns in other Semitic languages. §2.0 will provide a concise discussion of different VN patterns across stems, with particular focus on various pertinent phonological and morphological features. In §3.0, a resumé of the most prominent VN patterns and their surface realisations discussed in §2.0 will be given, followed by a summary discussion of a number of questions raised in §2.0 with consideration of the broader Semitic context.

The previously published data for this study come from Johnstone's (1987) *Mehri Lexicon* and Jahn's (1902) glossary, as well as the discussion by Morris (1981). Whereas in the *Mehri*

Lexicon (henceforth *ML*), VNs appear quite underrepresented, and for most verbs no VNs are indicated,[2] in Jahn's (1902) glossary, almost every verb is listed with a corresponding VN "*Infinitiv*." However, many common VN patterns appear in non-corresponding stems, which suggests the need for caution regarding the feasibility of some of the items provided.[3] Also, as will be shown, the previously published data quite often appear to be contradictory, both within a single publication and between different publications. Thus, for the sake of expanding and correcting previous data, further data were collected by the author in a fieldwork session in Salalah in February and March of 2022. The varieties of all speakers recorded appear to be examples of Eastern Mehri (henceforth EM); unfortunately, no speakers of Western Mehri (henceforth WM) could be consulted by the author. During this session, VNs in J/S were also collected, which were previously barely attested at all.

In terms of methodology, after giving a summary overview of VN surface patterns attested for each type in §2.0, pertinent features will be investigated, on the basis of, inter alia, previously described phonological features of Mehri, such as the IG-effect (Bendjaballah and Ségéral 2014), the effects of guttural and glottalic consonants (Rubin 2018; etc.), the vowel-spread of sonorants to their left (Dufour 2017; Rubin 2018), the status of vowels (Bendjaballah and Ségéral 2017; Rubin 2018), etc.

---

[2] Yet, significantly more VNs are present in ML than in JL.

[3] Although this does not seem to be uncommon—in the data collected by the author, informants would also sometimes give non-corresponding derived-stem VNs.

2.0. Verbal Nouns in Mehri

2.1. G-Stem VNs

For G-Stem VNs, the following surface patterns are attested in ML:

Table 1: G-stem VN patterns in ML

Surface Pattern	**Attestations**
CayCəC	83
CaCC	10
CəCC	5
CəCayC	9
CēCəC	4
CōCəC	5
+ Others (lexicalised)	

As evident, by far the most productive pattern of VN formation appears to be the pattern *CayCəC*. Most other patterns seem interspersed and most likely lexicalised. In terms of the distribution between Ga- and Gb-stems, Ga-stem VNs are much more frequently attested, otherwise no significant difference can be found, with a few salient exceptions (see §2.1.2).

Table 2: Ga- and Gb-stem VN patterns in ML.

Pattern	**Ga**	**Gb**
CayCəC	*ḥəlūb* (VN *ḥayləb*)	*rīdəf* (VN *raydəf*)
CaCC	*nəkūf* (VN *nakf*)	*śīrəġ* (VN *śarġ*)
CəCC	*dəḵḵ* (VN *dəḵḵ* < *dəḵ* >)[4]	
CəCayC	*rəḥāṣ́* (VN *rəḥayṣ́*)	*wīzək* (VN *wəzayk*)
CēCəC	*ləhūh* (VN *lēhi*)	*wīṣəl* (VN *wēṣəl*)
CōCəC	*fərūd* (VN *fōrəd*)	—

[4] All attestations of the VN pattern CəCC in ML show biradicals CəCC > CəCC.

One thus observes from Table 2 that by and large no VN pattern seems to be reserved for one particular G-stem (Ga or Gb), although the pattern *CəCayC* is problematic in this regard, as will be shown below.

In the following section the patterns *CayCəC* (§2.1.1), *CəCayC* (§2.1.2), *CvCC* (§2.1.3), and *CV̄CəC* (§2.1.4) will be discussed, followed by a brief overview of some of the other (lexicalised) patterns attested (§2.1.5).

2.1.1. *CayCəC*

Of all G-stem VNs, the pattern *CayCəC* is by far the most well attested. As mentioned above, the pattern *CayCəC* is attested for both Ga- and Gb-stems. In fact, a number of examples can be found of Ga- and Gb-stem verbs derived from the same root which appear to take both *CayCəC* and *CīCəC* as their VN pattern:

(1) *CayCəC*-type VNs Ga-Gb in ML
ṯəbūr (Ga) / *ṯībər* (Gb) – *ṯaybər* (VN)
ḥəsūd (Ga) / *ḥaysəd* (Gb) – *ḥaysəd* (VN)[5]

In Johnstone's data, based on EM, this pattern generally surfaces as *CayCəC*, whereas in Jahn's data this pattern usually surfaces as *CīCəC*, with diphthongisation only appearing when the first radical is a guttural or glottalic consonant, e.g., *līṯaġ* 'to kill' versus *ġayleḵ* 'to look for, search' (Jahn 1902).

There are a few exceptions in Jahn's data, where a diphthong appears with R1 = [−Gutt., −Glott.] and, concomitantly,

[5] The nominal form *ḥaysəd* in ML is glossed as a generic noun 'envy', however, it is almost certainly a VN formation.

where no diphthong appears with R1 = [+Gutt., +Glott.]. The indicative value of these examples remains unclear. In any case, the possibility of these forms being erroneous (due to over-elicitation?) remains, particularly when viewed against the overall number of attestations of VNs in general, and *CəyCəC/CīCəC* in particular.

(2) *CəyCəC* / R1 = [−Gutt., −Glott.]: *kéynes*, *kaytb*, *śaymer*, *śayreṭ*, *kéysī*, *téylī*, *déyḥar*, *téylef*
CīCəC / R1 = [+Gutt., +Glott.]: *xīneḳ*, *ṣīber*, *ṣīdeḳ*, *ḳīṣem*

Regarding the question of the nature of the diphthong/monophthong in the first syllable, the surface presence of such a diphthong in a stressed syllable in EM would suggest an underlying glide /y/. Also, in EM distinct noun patterns *CīCəC* and *CəyCəC* are attested outside of VNs (see also Dufour 2016, 371–72). However, the question remains whether an underlying glide can be assumed for the whole of Mehri, e.g., also for the western varieties in Jahn (1902). Thus, in the following section the following factors will be considered to evaluate the underlying nature of the stressed monophthong/diphthong of the VN patterns *CəyCəC* and *CīCəC*: IG-effect between R2 and R3 (§2.1.1.1.); the presence of gutturals in R2 (§2.1.1.2.); cognates from other MSAL (§2.1.1.3.).

2.1.1.1. IG-Effect between R2 and R3

The first argument concerns the application of the IG-effect between R2 and R3.[6] If the initial syllable was entirely vocalic, one would expect shortening of the stressed vowel, since no *CV̄CəC*-type nouns with R2 + R3 = [+IG] are attested.

A significant difference between the data provided by Johnstone and Jahn lies in the application of the IG-effect in the forms in question. The forms collected in ML do not show any traces of IG, and hence they surface as *CayCəC*:

(3) *faytəḥ*, *raykət*, *bayḥəṭ*, *dayhəf*, *mayḥəś*

By contrast, amongst the forms in Jahn's glossary, almost no surface forms *CīCəC* with R2 + R3 = [+IG] appear.[7] However, amongst the surface pattern *CəCC* (<*CeCC*>/<*CiCC*>) one finds a number of forms R2 + R3 = [+IG]. Hence it would appear that *CayCəC/CīCəC* regularly displays the IG-effect between R2 and R3, and therefore, the stressed surface vowel between R1 and R2 is shortened.

(4) *ʿeks* (*ʕks* Ga), *liḥs* (*lḥs* Gb), *misḥ* (*msḥ* Ga), *nifh* (*nfh* Ga) *nefś* (*nfś* Ga), *niḥt* (*nḥt* Gb), *nikś* (*nkś* Ga), *nisf* (*nsf* Ga), *nitf* (*ntf* Ga), *nitx* (*ntx* Ga)

In the data collected by the author, surface forms *CayCəC* with R2 + R3 = [+IG] were attested, thus corresponding to the situation in ML:

---

[6] I.e., the elision of an unstressed vowel between voiceless, non-glottalic consonants (see Bendjaballah and Ségéral 2014).

[7] With the exception of *lišet*.

(5) *layḥəs – lḥays* (*lḥās*)
naykəf – nakf (*nkūf*)
bayḥəṯ – baḥṯ (*bəḥāṯ*)
āykəs – aks (*ākūs*)

Furthermore, amongst the attestations of *CəCC* only two examples of R2 + R3 = [+IG] could be found: *nəkś* (*nkūś*) and *məsḥ* (*məsḥ*). Some attestations of *CaCC* with R2 + R3 = [+IG] could be found (e.g., *nakf*, *faḥs*, *maḥś*, *natf*, *aks*, *natx*, etc.), however, in at least some of the aforementioned examples, no reason seems apparent why the vowel in the stressed syllable should be considered to be a phonetic variant of /ə/,[8] hence—at least most of—these forms cannot be explained as underlying *CayCəC*, since the shortened stressed vowel should be /ə/ in these cases and not /a/.

Therefore, it appears that in EM the IG-effect does not apply to *CayCəC*. Whereas the question remains open why apparently in some varieties of Mehri the initial diphthong is (not) shortened, the fact that it is not shortened in EM highlights the distinctness of this pattern from *CV̄CəC*-type nouns.

2.1.1.2. The Presence of Gutturals in R2

The next aspect is connected to the surface similarity between Jahn's *CīCəC* and the 3MS perfective form of Gb-stem verbs (also *CīCəC* in Mehri). If this surface similarity is to be understood as underlying identity, one would expect a similar surface outcome with R2 = [+Gutt.] roots, i.e., a lack of patterns *CayCəC/CīCəC*

[8] This would be expected when adjacent to a guttural, which is not the case in *nakf*, *natf*, *natx*, etc.

with R2 as a guttural and a number of corresponding surface forms of the type *CəCēC* or *CəCāC*. Yet, both Johnstone's and Jahn's data show a good number of VNs with R2 = [+Gutt.]:

(6) *CayCəC* / R2 = [+Gutt.] in ML
bayhər, *bayḥəṯ*, *dayhəf*, *dayxəl*, *ḏayhəb*, *mayḥəḵ*, *mayḥəś*, *rayḥəl*, *rayxəṣ*, *zayhəd*

(7) *CīCəC* / R2 = [+Gutt.] in Jahn (1902)
śayġab, *śayhar (ṭayhar)*, *déyḥar*, *ṣayhel*, *ṭeyḥan*, *zayġaf*, *bīġaś*, *līheg*, *sīḥaṭ*, *bīhel*, *dīheb (ḏīheb)*, *dīher*, *dīḥaḵ*, *līheg*, *līḥaḵ*, *nīheḵ*, *rīhez*, *rīḥeḵ*, *sīher*, *sīḥeḵ*, *tīhel*, *wīhed*, *wīhem*

Note also that among the rare VN pattern *CəCēC*, only two attestations of G-stem *CəCēC* with a guttural as R2 were found in ML (*śəḥāk* and *ṭəʾēn*), and none in Jahn (1902) (see §2.1.5). Thus, one observes that in terms of the behaviour of gutturals as R2, both Johnstone's and Jahn's data pattern in opposition to verbal forms Gb *CīCəC*, the latter presumably only showing an underlying vowel between R1 and R2.

An important note of consideration is that amongst the VNs elicited by the author, *CayCəC*-type VNs of II-Gutt. roots were almost completely lacking,[9] and usually surface as *CəCayC*. The question of the relationship between *CayCəC* and *CəCayC* will be discussed in §2.1.2.

[9] With the exception of *layḥəs* and *bayḥəṯ*, both of which were also elicited alternatively as *CəCayC* (*lḥays*, *bəḥayṯ*).

2.1.1.3. Cognates from other MSAL

Cognates from other MSAL would seem to support the interpretation of an underlying glide. Whereas intra-MSAL vowel correspondences are an intricate subject, with no one-to-one alignments, a few pertinent trends can be noted. Thus, /i/ in Mehri often corresponds to either /i/ or /e/ in J/S and Soq. Examples for the latter correspondence can be found in the 3MS perfective forms of the Gb-stem (Mehri *CīCəC* – J/S *CéCəC* – Soq. *CéCəC*) or in the G-stem passive participle (Mehri *məCCīC* – J/S *məCCéC* – Soq. *méCCeC*).

In places where an underlying /y/ is to be expected in Mehri, due to the non-conditioned surface diphthong, the reflex in J/S and (when applicable) Soq. appears to be /i/, as witnessed in the frequent MS adjective pattern *CəCayC* and its J/S reflex *CəCíC*.

Regarding potential cognates to VN *CayCəC/CīCəC* in JL, the possible cognate form *CéCəC* is attested in JL. Johnstone mentioned only one potential item of this pattern, namely the item he transcribed as *héḏər*. Amongst all J/S informants the author consulted, this item was consistently produced as *ḥíḏər*. Hence, apparently Johnstone's form is erroneous both in terms of the nature of R1 and the vowel between R1 and R2. Other examples of the same pattern show /i/ between R1 and R2 and not /e/, and there are enough VNs attested without a neighbouring nasal consonant or /r/ to ascertain that this is not due to vowel raising (see Rubin 2014, 40):

(8) VN pattern CíC(ə)C in J/S (author's data)
ḥíḏər (Ga *ḥɔḏɔ́r*), *ríkəb* (Gb *rékəb*), *ṣílb* (Ga *ṣɔlɔ́b*), *hídəm* (Ga *hɔdúm*), *ʿíks* (Ga *ʿɔkɔ́s*), *díhf* (*dɛhéf*), *fírd* (Ga *firɔ́d*), *díḥar* (*daḥár*), *nikf* (Ga *nkɔf*), *bíxəṣ* (*baxáṣ*), *bírk* (*berɔ́k*), *líṣəḳ* (*léṣḳ*), *díġəl* (*daġál*), *ġílḳ* (*ġɔlɔ́ḳ*), *ṭíʿan* (*ṭaʿán*), *ríğaʿ* (*réğaʿ*)

Note in these forms the application of the IG-effect between R2 and R3, as well as the lack of a surface vowel between R2 and R3 in forms where R2 is a sonorant (more specifically, a liquid). With regard to the former aspect, the behaviour of *CíCəC* in J/S resembles the putative behaviour of *CīCəC*-type VNs in Jahn (1902). It should be noted that this does not necessarily imply that *CíCəC*-type VNs in J/S have a synchronically underlying glide. The sound-correspondences, however, do point towards cognacy with forms which on the Mehri end would typically have an underlying glide. In the case of Soq., the cognate pattern appears to be *CíCiC* (Kogan and Bulakh 2019, 302).[10] Hence we also observe /i/ as the first vowel.

Thus, the vowel correspondence between /vy/ (ML) and /i/ (Mehri) (Jahn 1902), /i/ (J/S) and /i/ (Soq.), implies an underlying glide in Mehri, as witnessed in the aforementioned singular adjectival pattern *CəCayC*.

2.1.1.4. Summary on *CayCəC / CīCəC*

For EM, when taking into account that the first syllable shows a surface diphthong, that no IG-effect is observed between R2 and R3 (§2.1.1.1), that no simple-vowel shift is observed when R2 is a guttural consonant (§2.1.1.2. in Johnstone's and Jahn's data,

[10] Examples *ṣífin* and *ḥíṣiḳ* in Kogan, Naumkin et al. (2018).

for the author's data see §2.1.2.), and that sound-correspondences exist between Mehri and other MSAL—indicating the non-vocalic nature of the first syllable in Mehri (§2.1.1.3)—one arrives at the conclusion that the initial syllable of *CayCəC* shows an underlying glide /y/.

For the western varieties described by Jahn (1902), the situation appears somewhat more problematic. While the sound-correspondences mentioned in §2.1.1.3 apply, as well as the lack of shift with gutturals as R2 (§2.1.1.2), (almost) no forms with R2 + R3 = [+IG] could be found (§2.1.1.1). However, even with this being the case, this does not necessarily constitute a counter-argument against the notion of an underlying glide in WM, since the shortening of non-glottalised/gutturalised diphthongs in a stressed and phonologically closed syllable is attested otherwise.[11] To explain the distinction between EM *CayCəC* and WM *CīCəC* a more promising explanation could perhaps be found in different dialectal realisations of underlying /y/, rather than by assuming distinct underlying patterns.[12]

---

[11] E.g., *bayt* > *abátk* (Rubin 2018, 56). As mentioned, it remains an open question why exactly this effect does not seem to apply in the case of EM *CayCəC* / R2 + R3 = [+IG]. Any connection with a possible historical long vowel between R2 and R3? (see §3.0 and Dufour 2016, 376–78).

[12] Potentially due to a distinction in the underlying vowel of stressed underlying *vy/vw* sequences – /ə/ in WM, /a/ in EM? Therefore /əy/ > /ī/ in the western data as opposed to /ay/ > /ay/ in EM? This solution might also pattern well with non-secondary diphthongs in EM and their monophthong counterparts in WM in a more general sense.

2.1.2. *CəCayC / CəCyūC* (and *CayCəC*)

VNs of the patterns *CəCayC* and *CəCyūC* deserve separate discussion due to their peculiar distribution. Unlike other G-stem VN patterns, these two patterns appear to be restricted almost entirely to R2 = [+Gutt.] verbs, as well as to a few Gb-stem verbs R2 = [−Gutt.].[13]

(9) *CəCayC* with R2 = [+Gutt.] verbs in ML
dəhayr, *dəḥayk*, *dəḥayḵ*, *ṯəḥayl*, *śəhayd*, *śəxayl*, *ṭəʿaym*

(10) *CəCayC* with R2 = [−Gutt.] Gb verbs in ML
fəḏayr, *wəzayk*

(11) *CəCyūC* with R2 = [+Gutt.] verbs in Jahn (1902)
deheyūb, *daḥayūḵ*, *ṯaḥayūl*, *zaġayūf*, *ṭaʾayūn* (*ṭʿn*), *gāyūr* (*gʿr*), etc.

(12) *CəCyūC* with R2 = [−Gutt.] Gb verbs in Jahn (1902) *rakiūb*, *śabyūṭ*

Moreover, the VN patterns *CəCayC* and *CəCyūC* appear to be in complementary distribution (see *dəḥayḵ* – *daḥayūḵ*, *ṯəḥayl* – *ṯaḥayūl*), in ML one does not find any attestations of *CəCyūC* as a

[13] In Morris (1981, 255), the pattern *CəCayC* is described as a prominent VN pattern for *CəCēC*-type verbs, that is G-stem verbs with R2 = [+Gutt.], although there is at least one attestation of *CəCayC* in a verb R2 = [−Gutt.] in Morris's data (*fīḏər* > *fəḏayr*, Morris 1981, 254, also to be found in ML). No attestations of *CəCyūC* as a VN pattern are to be found in Morris's study. By contrast, Watson (2012, 26) notes that "...a more common verbal noun pattern for the simple verb [referring to VN *fʿūl*] is *faʕyūl*." Also, Bittner (1909, 22–23) describes *CəCyūC* as a particular VN (infinitive) pattern, mostly attested for roots II-Gutt., which he connects with **CiCāC*.

VN pattern, while in Jahn's glossary one does not find any attestations of *CəCayC* (*CəCīC*)[14] as a VN pattern.

As has been shown §2.1.1, according to the published data, the presence of a guttural as R2 cannot be taken to be a sufficient indication for the occurrence of *CəCayC/CəCyūC* alone, since *CayCəC*-type VNs with R2 = [+Gutt.] are attested. Also, as one has seen above, the patterns *CəCayC* and *CəCyūC* do not appear to be exclusively reserved for roots R2 = [+Gutt.]. Hence, if the patterns *CəCayC* and *CəCyūC* do indeed constitute distinct common VN pattern(s), it has to be assumed, that this pattern would be distinct from *CayCəC/CīCəC*.

In terms of the internal structure of roots showing the patterns *CəCayC* and *CəCyūC*, no salient features which might explain the two patterns as being conditioned variants of each other (or of *CayCəC/CīCəC*) could be found, aside from the aforementioned overrepresentation of R2 = [+Gutt.]. Thus, R1 and R3 of all forms in question are shown in the tables below.

Table 3: Positional Analysis R1 and R3 of *CəCayC/CəCyūC* / R2 = [+Gutt.] and *CayCəC/CīCəC* / R2 = [+Gutt.] in ML and Jahn (1902)

Combinations R1-R3 *1ə2ay3* (ML)		**Combinations R1-R3 *1ə2(ə)yū3* (Jahn 1902)**	
d-r	*ś-d*	*d-ḳ*	*ṯ-l (ṯ-l)*
d-k	*ś-l*	*d-b (ḏ-b)*	*z-f*
d-ḳ	*ṯ-m*	*l-m*	*ṣ́-ṯ*
ṯ-l	*w-k*	*r-l*	*b-r*
f-r		*r-b*	*ṭ-n*
		r-ṣ́	*z-ḳ*
		g-r	

[14] There is one example of a VN of the pattern *CəCīC* in Jahn (1902), *ḳəṣīd*, which is unlikely to be a cognate, since it is a probable loan from Arabic.

Combinations R1-R3 *1ay2ə3* (ML), R2 = Gutt.		**Combinations R1-R3 *1ay/ī2ə3* (Jahn 1902)**	
b-r	*m-ḵ*	*b-l*	*s-ḵ*
b-ṯ	*m-ś*	*b-ṣ́*	*s-ṯ*
d-f	*r-l*	*d-b*	*ṣ-l*
d-l	*r-ṣ*	*d-r*	*ṯ-l*
ḏ-b		*d-ḵ*	*ṯ-n*
		l-g	*w-d*
		l-ḵ	*w-m*
		n-ḵ	*z-f*
		r-ẓ	*z-r*
		r-ḵ	*ṣ́-b*
		s-r	*ṣ́-r*

Hence, one finds a diverse configuration of possible groupings of neighbouring consonants in VN patterns *CəCayC/CəCyūC* / R2 = [+Gutt.] and also an overlap with groupings of neighbouring consonants in VN patterns *CayCəC/CīCəC* / R2 = [+Gutt.], indicating that the distinction between *CəCayC* and *CəCyūC* cannot be immediately traced to factors pertaining to the neighbouring consonants (other than R2). In the absence of any other discernible factors, it would thus have to be assumed that the two patterns do indeed represent two synchronically distinct VN patterns between the different dialects of Mehri, and both similarly distinct from *CayCəC/CīCəC*. Note that both patterns are well attested otherwise: *CəCayC* is otherwise attested as the most prominent adjective pattern, while *CəCyūC* (and *CəCyōC*)[15] is attested as a plural/collective noun pattern in all dialects of Mehri.

In the data collected by the author, there is another distinct patterning of *CəCayC* and *CayCəC*. Whereas in ML, II-Gutt. roots seamlessly take *CayCəC* and *CəCayC*, almost all G-stem II-Gutt.

[15] *CəCyōC* appears to be the EM cognate of WM *CəCyūC* (in non-VN items)? Note the curious relation of EM /o/ – WM /u/.

items collected by the author took the latter pattern. The only exceptions are *layḥəs*, which was also collected as *lḥays*, and *bayḥəṯ*, which was also collected as *bəḥayṯ* and *baḥṯ*. *CəCayC* type VNs with a guttural consonant as R2 include those which are given as *CayCəC* in ML.

Table 4: *CayCəC* / R2 = [+Gutt.] in ML and their reflexes in the author's data

ML	Author's Data
bayhər	*bəhayr*
bayḥəṯ	*abḥayṯ (baḥṯ, bayḥəṯ)*
dayhəf	*dəhayf (mdəhfēt)*
dayxəl	*- (dəxōlət)*
ḏayhəb	*ḏəhayb (ḏəhīb* in Morris 1981*)*
mayḥəq	*- (maḥq)*
mayḥəś	*mḥayś (maḥś)*
rayḥəl	*rḥayl*
rayxəṣ	*- (raxṣ)*
ẓayhəd	*- (ẓahd)*

Therefore, on the basis of the author's data it almost appears as if *CayCəC* and *CəCayC* are conditioned variants of each other, based on the presence of a guttural consonant as R2. When no guttural is present, the pattern surfaces as *CayCəC*, whereas a guttural R2 seems to trigger a shift to *CəCayC*. Also, in the author's data some additional items *CəCayC* that do not show a guttural as R2 could be found:

(13) VN *CəCayC* / R2 = [−Gutt.] (author's data)
ūṣayḳ – laṣḳ (līṣəḳ)
nśayz – nśawz (nīśəz)
ṯəbayr – ṯōbər – ṯaybər (ṯībər)
awkayb – waykəb (ūkūb)
awṣayl – ūṣawl – wayṣəl (wīṣəl)

nfayś – mənfēś (nfūś)
anaydəf – ndayf – nadf (ndūf)
ūzayk (wīzək)
fəḏayr (fīḏər)[16]

Note that amongst these forms R1 is mostly either a /w/ glide or a /l/ or /n/ sonorant. Variant forms could be found for most of the items above, some of which include *CayCəC*-type VNs. Also, *CayCəC* is attested otherwise with R1 as a /w/ glide or /l/ or /n/ sonorant. Hence it might very well be that the nature of R1 is of no particular relevance for the surfacing of a given item as *CayCəC* or *CəCayC*.

(14) Other attestations of *CayCəC* / R1 = [/w/, /l/, /n/] (author's data)
layḥəs (lḥās)
laybəd – waybəd – labd (ūbūd)
nayka (nūka)
naykəf – nakf – nkfūt (nkūf)

In J/S, most examples of II-gutt. roots do not take a form resembling *CəCayC*. However, a few examples of prima-facie cognates can be found:

(15) VN *C(ə)CíC* in J/S (author's data)
ṭəḥí(h)l (*ṭaḥál*) viz. *ṭəḥayl* (*ṭəḥāl*)
ṣ(ʿ)aykḳ (*ṣaʿáḳ*) viz. *ṣ(ʿ)ayḳ – ṣ(ʿ)awḳ* (*ṣāḳ*)
śəḥík (*śaḥák*) viz. *śəḥāk – śəḥkēt* (*ləḥāk*)
fḥís – faḥs (*fḥás*) viz. *fḥays – faḥs* (*fḥās*)

[16] See Johnstone (1987); Morris (1981, 254). In the data recorded by the author, the VN to this item was given as *fəḏrēt* by all informants.

kšíš (*kéšš*) viz. *kəśś* (*kəśś*)
lfíf (*leff*) viz. *lɛff* – *məlɛff* (*ləff*)
mḥíḳ (*maḥáḳ*) viz. *maḥḳ* (*mḥāḳ*)
nfíś (*nfoś*) viz. *nfayś* – *mənfēś* (*nfūś*)
msíḥ (*mésḥ* also VN *mísḥ*) viz. *məsḥ* (*məsḥ*)

Note that among these examples, those forms that do not show a guttural as R2 either have a sonorant /n/ or /m/ as R1, or—unlike Mehri—have a biradical root. The latter feature might be a peculiarity of J/S, although further research on J/S would be needed to establish this.

Concerning *CíCəC*, the putative cognate of Mehri *CayCəC*, some attestations of *CíCəC* with R2 as a guttural could be found:

(16) VN *CíCəC* in J/S / R2 = [+Gutt.] (author's data)
dɛhéf – *díhf*
daḥár – *díḥar*
daġál – *díġəl*
ṭaʿán – *ṭíʿan*
ǧaʿár – *ǧíʿar*
ḳaḥár – *ḳíḥar*

The J/S data collected by the author hence show both *CíCəC* and *CəCíC* being attested as VN patterns for G-stem VNs of roots with a guttural as R2, hence resembling the picture from ML and Jahn (1902).

To conclude, while the data in previously published sources show conspicuous parallels between the VN pattern *CəCayC* in EM and the VN pattern *CəCyūC* in WM, there is no reason to assume that *CəCayC* and *CəCyūC* are in any ways conditioned variants of each other. This, in turn, suggests genuine dialectal

differences within Mehri in terms of the usage of VN patterns. On the basis of the author's data, the impression emerges that *CəCayC/CəCyūC* and *CayCəC* are mutually exclusive VN patterns, with *CəCayC* appearing with roots where R2 is a guttural consonant. However, in the published data, as well as in J/S, *CayCəC*-type VNs with R2 = [+Gutt.] are well attested, and also *CəCayC*-type VNs with R2 = [−Gutt.] are attested, hence underlining the distinctness of these patterns. Taking all of this together, no cogent argument emerges to assume allophony between *CayCəC* and *CəCayC/CəCyūC*, and considering the lack of said arguments, one might assume that they represent synchronically distinct VN patterns as well. Nevertheless, there remains the unsatisfying situation, where one notices significant overlap between a given (set of) noun pattern(s), namely *CəCayC/CəCyūC* and phonological (mainly II-Gutt. roots) and morphological (mainly Gb-stem) features associated to it. Perhaps an answer to this puzzle is to be found when considering this issue from a diachronic perspective, which will be left for further research.

2.1.3. *CvCC*

The patterns *CaCC* and *CəCC* are less frequently attested as VNs, however, still relatively prominent VN patterns, particularly for G-stems.

In ML, two of the *CaCC* VNs show *CayCəC* as a variant VN, *fark – fayrək* and *harḳ – hayrəḳ*. This agrees with the overall impression of the data collected by the author, where speakers would frequently oscillate between *CayCəC* and *CaCC* as G-stem VN patterns. In the data provided by Johnstone, no sonorants

were found as R3 in *CvCC*-type VNs, as expected. In the author's data, a few such examples could be found, which might very well represent over-generalisations on the part of the speaker (all items *CvCC* / R3 = [+Son.] come from the same speaker), considering the general near-absence of sonorants as R3 in *CvCC*-type nouns. Note that for all of these VNs other patterns were also given, and that two of the three attestations were given as T2-stem VNs.

(17) VNs *CaCC* / R3 = [+Son.] (author's data)
fahm – fthəmūt (*fthūm*)
fagr – fəgōrət (*fəgūr*)
wakl – tūkəlēt – ūtkəlūt – ūtkəlēt (*ūtkūl*)

In Johnstone's data three items of the pattern *CaCC* show IG consonants as R2 + R3, namely *nakf*, *ṭahs*, and *hasf*.[17] It is a priori possible that these forms represent underlying *CēCəC* or *CōCəC*. However, this cannot be ascertained, and in a more general sense, these examples would still not change the overall picture that *CvCC*-type VNs are more frequently attested than *CV̄CəC*-type VNs.

As for the pattern *CəCC*, as mentioned above (§2.1), examples in ML are drawn from biradical roots, where the VN would be identical with the 3MS perfective. This corresponds to a more general feature of *CəCC* in ML, where most attestations of this pattern across all (nominal) forms are to be drawn from biradicals or transparent Arabisms. The reasons for this odd configuration are beyond the scope of this paper. However, when

[17] The latter is listed as the VN for a D/L verb.

considering that this situation is attested for the noun pattern *CəCC* beyond VNs, it is to be assumed that the answer to this question also lies beyond the domain of VNs.

In the author's data two attestations of *CəCC* of triradical roots could be found, the aforementioned *nəkś* and *məsḥ*, with the other attestations of *CəCC* being similarly restricted to biradicals. In at least two examples of biradical roots, the corresponding VN did not appear identical to the 3MS perfective, hence *lɛff – məlɛff* (*ləff*) and *śɛdd – mśɛdd* (*śədd*).

2.1.4. *CV̄CəC*

Amongst patterns of the type *CV̄CəC*, *CēCəC*, and *CōCəC* are attested, which are also, in a more general sense, the most frequently attested noun patterns of the type *CV̄CəC*. Most of the examples given in ML show other patterns in the material collected by the author; hence, it is questionable whether the items given in ML actually represent VNs of the type *CV̄CəC* or rather simple nouns of the same root.

(18) *fēḵəś*, *wēṣəl*, *lēhi*, *āḏər*, *fōrəd*, *tōnəg*, *hōwi*, *ḥōśi*, *bōni* (H-stem), *wəṣayl*, *fayrəd*, *ḥaywi* (author's data)

A few other items of the patterns *CōCəC* and *CēCəC* were collected by the author:

(19) *lōtəġ* (ūtūġ)
ṯōbər – ṯəbayr – ṯaybər (ṯībər/ṯəbūr)
śēni – məśənay (śīni)

2.1.5. Other Patterns

The following types of patterns were only quite sparsely attested as VN patterns in ML. Hence, they likely represent examples of lexicalised VNs. The following list is not a complete set; various other yet smaller VN patterns are attested.

2.1.5.1. Prefix *m-*

A few VNs taking a prefix *m-* could be found. These forms can take the FS suffix(es) and various stressed vowel qualities and stress positions.

Table 5: G-stem VNs with a prefixed *m-* (author's data)

Pattern	**Attestations**
məCáCC	*msayr* (*səyūr*)
məCɛCC	*məlɛff* (*ləff*), *mśɛdd* (*śədd*)
məCCaC	*məśənay* (*śīni*)
máCCəC	*mádḥəḳ* (*dəḥāḳ*)
məCCəCēt	*mdəhfēt* (*dəhayf*)
məCCē	*mətwē* (*təwōh*)
məCCēC	*mənfēś* (*nfūś*)
məCCūC – məCCawC	*mərḳawd* (*rḳawd*)

2.1.5.2. Suffix *-Vn*

Whereas in J/S a suffix *-Vn* appears to be more prominently attested, in Mehri this suffix is quite sparsely attested. Nevertheless some attestations can be found, e.g., *ṭəhəḳayn* (ML), *ġəśśīn*, *ḥalmīn*, *ġafirōn*, *geḥeydōn*, etc. (Jahn 1902). Jahn's glossary shows a somewhat larger number of forms with a suffix *-Vn*, with both *-īn* and *-ōn* being attested. The examples above do not appear to show any particular phonological features which might explain

the appearance of the *-Vn* suffix. They thus should probably be considered lexicalised.

2.1.5.3. Reduplication

In ML and in the author's elicitation efforts, only a single item of (partial) reduplication could be found, namely *kəbkēb* (Ga *wkb*). In Jahn (1902) one also finds the forms *ṣaḳaṣēḳ* and *zemzēm*, both I-*w*. Also, in other MSAL more examples of partially reduplicated I-*w* VNs can be found, which might imply the presence of more of such VNs in Mehri, hence J/S *gəḥgáḥ* (JL), *zəkzák* (author's data).

2.1.5.4. *CəCv̄C*

Patterns of the type *CəCv̄C* are sparsely attested. ML shows *ṣ́əḥāk, śədɛ̄d, ṭəʾēn* and *ādōm* (the latter being the VN of a H-stem verb). In addition, the form *hīḳōy* is attested for the irregular (T-stem) verb *təḳḳ*, which should probably be understood as *CīCōC*. In Jahn (1902) more examples are to be found. However, most examples are attested for derived stems, and hence questionable in terms of their convergence with the initial definition of VNs. For G-stem verbs, the following items are attested in Jahn (1902), all but three of which show *CəCēC* (and its conditioned variant *CəCāC*, also attested for apparently non-integrated Arabisms):

(20) *ḥabēr* (*ḥbr* Gb)
mirēṣ́ (*mrṣ́* Gb)
ṭarēf (*ṭrf* Ga?)
sedēd (*sdd* Ga)
zetēt (*ṣetēt*?) (*ztt*/*ṣtt*? Ga)
śədēd (*śdd* Ga)
ftāḥ (*fthˌ* Ga)
nśāḳ (*nśḳ* Ga)
ṭayām (ṭʿm Gb)
amāl (ʿml Gb)

Only two G-stem VNs of the pattern *CəCōC* are attested, namely *ṣabōḥ* and *gizōz*, and one (presumably Arabic) pattern *CəCīC ḳaṣīd*.

2.1.5.5. Simple Base with Suffix *-ōt*

Some G-stem VNs show a stressed suffix *-ōt*, hence, ML *rābūt* (*rʿb*) or *ṯəmərēt*. These patterns are less frequently attested for G-stem VNs, and most attestations of *CəCCūt/CCəCūt* are H-stem VNs (see §2.3).

2.1.5.6. Distinct Root

In one example, the VN of a G-stem verb is formed from a distinct root, namely *šəwkūf – šənēt*. This is most likely a feature inherited from PS, as argued by Kogan and Militarev (2000, 336—noting that the verbal usage of this root in other Semitic languages might be a secondary development).

2.1.5.7. Arabisms

Furthermore, a number of items in the data collected by the author take various VN patterns which seem to be formed in analogy to their Arabic counterpart, most prominently *CəCawC* (Ar. *CuCūC*) and *CəCōCət* (Ar. *CvCāCa*):

(21) *ktūb – ktōbət* (Ar. *kitāba*)
śhēd – śhōdət (Ar. *šahāda*)
rkūb – rkawb (Ar. *rukūb*)
wīṣəl – ūṣawl (Ar. *wuṣūl*)

2.2. D/L-Stem VNs

In the D/L stem, almost all VNs appear as *təCCáyC* (ML) or as *təCCīC* in Jahn's data. However, in a few cases forms appear which take a prefix t- and usually a stressed suffix.

Table 6: Irregular D/L Stem VNs in ML

Pattern	**Attestations**
tāCəCēt	*tābəlēt* (I-ʿ)
təCCəCēt	*təwkəlēt*
təCCāt	*tərbāt* (III-ʿ)
təCōCōt	*təwōṣōt* (III-*y*)
təCCɛ̄	*tərgɛ̄* (III-ʿ) (<*təCCayC*?)

In the data collected by the author, similar forms to the ones in ML appeared.

Table 7: Irregular D/L Stem VNs in author's data/ML

Pattern	**Attestations**
tāCəCēt	*tābəlēt* (ML), *tābəlēt / tābáwlət* (author's data)
təCCəCēt	*təwkəlēt* (ML), *tūkayl / tūkəlēt* (author's data)
təCōCōt	*təwōṣōt* (ML), *tūṣōt* (author's data)

The VNs in question show a 'weak' consonant either as R1 (/ʿ/, /w/) or as R3 (/y/, /l/). Other *təCCayC*-type VNs can also be

found for roots containing 'weak' consonants as R1 (*tālaym*, *tāṯaym*, *tākayb* [all I-ʕ]), but not as R3. This might indicate that the appearance of a suffix (and the seemingly lack of a diphthong between R2 and R3) is conditioned by the nature of R3. However, more data are needed to clarify the matter.

Since a cognate pattern does not seem to be attested in Soq. as an equally productive VN pattern,[18] and the long vowel /ī/ is represented by a surface diphthong in EM, it seems that *təCCayC* might be a loan from Arabic. According to Kogan and Bulakh (2019, 302), the D/L-stem VN in Soq. is of the shape *CɛCíCo*, which implies a cognate form with a suffix *-ūt* or *-ōt* in Mehri, which is not attested for the D/L-stem. Regarding the forms with suffixes in Mehri, these do not seem to be the simple outcome of *təCCayC* and a suffix *-ʸ̄t*, since in other forms with an underlying stressed glide, stress does not usually shift when a suffix is added, i.e., the FS adjective pattern *CəCayCət* (MS *CəCayC*). Hence, these forms appear to be distinct (verbal) noun patterns. Note that, with the potential exception of animal nouns of the type *təCCēC*, no other patterns with an initial /t/ which is not part of the root are attested amongst Mehri nouns (at least in ML), which might indicate that this initial /t/ is connected in some way or another to the initial /t/ in *təCCayC* (and Arabic and presumably PS **taC-CīC*).

2.3. H-Stem VNs

VNs of H-stem verbs are attested in the following surface forms: *həCCəCʸ̄t*, *həCəCCʸ̄t*, and *C(ə)C(ə)Cʸ̄t*. The latter form is the

[18] Although nouns of the type *tvCCiC* do exist, some of which might be classified as VNs.

regular reflex of a root with R1 = [+IG]. Namely, the /hə-/ prefix does not surface, as is the case with finite verbal forms (see Rubin 2018, 131–32). As for the distribution between *həCCəCvt* and *həCəCCvt*, this is clearly explainable via the presence of a sonorant as R2 or a glottalic consonant as R1.

Table 8: Surface *həCCəCūt* versus *həCəCCūt* in ML

həCCəCūt	***həCəCCūt***
hənġəmūt	*həḳəfdūt*
hənḳəbūt	*həmərtūt*
hənsəmūt	*hənəwfūt*
hərdəfūt	*həṣərdūt*
həwgərūt	*həməwkūt* (II-*l*)
həwgəśūt	
həwrədūt	
həzbərūt	

(22) H-Stem VNs without a surface h- prefix in ML

kəbərūt, *xəṣbūt*, *xəwfūt*, *təhmīt* (!), *kəbbūt*, *hḳəṭawt*, *fəlḥawt*, *fərḳawt*, *fḳawt*

One observes in the examples from ML given above that the presence of a sonorant as R2 or, in at least one case (*həḳəfdūt*), of a guttural consonant as R1 occasions vowel insertion between R1 and R2, whereas the lack of said features yields the outcome *həCCəCūt* (e.g., *həzbərūt*). Therefore, it appears that the standard form of the base of H-stem VNs is *həCCəC_*, with the presence of a sonorant as R2 triggering vowel insertion to its left and vowel elision to its right. The data collected by the author confirm this situation.

Table 9: *həCəCCūt / həCCəCūt* (author's data)

həCCəCūt	***həCəCCūt***
həghədūt	*həḳəfdūt*
həgḥəbūt	*həmərtūt*
	həmərśawt
	həṣərdūt

The vowel quality of the suffix appears consistently as *-ūt* and its conditioned variant *-awt* in both Morris (1981) and the data collected by the author. In ML, the problem of distinguishing *-ūt* and *-ōt* applies (see §2.7), whereas in Jahn (1902), despite the larger number of forms *-ōt* attested, a number of forms showing *-ūt* (or *-awt*) could be found as well.

Table 10: *-ūt* and *-awt* H-Stem VNs in Jahn (1902)

Pattern	**Attestations**
hə1ə23ūt	*hamertūt* (*mrt* H)
hə1(ə)2awt	*hadaḥawt* (*dḥw* H)
hə1ə2ūt	*hegerūt* (*gry* H), *haġaśūt* (*ġśy* H), *haḳowūt* (*ḳwy* H), *hamelūt* (*mlʾ* H)
hə1ə3ūt	*haġaṭūt* (*ġwṭ* H)
hə12ūt	*henhūt* (*nhy* H), *h(a)uzūt* (*wzʿ* H)
hū2ūt	*hūfūt* (*wfy* H)

Almost all of these examples (with the exception of *hamertūt*) are III-*w* or III-*y*, with one item II-*w* (*haġaṭūt*), and thus might mirror the situation in EM, where nominal *-ōt* is attested with a number of *tertiae infirmae* roots (usually III-ʾ), but is quite rare otherwise.

2.4. T-Stem VNs

For T-stem VNs, previously published sources give contradictory accounts. Consider the attestations of T-stem VNs in ML and Jahn (1902):

Table 11: Surface T-stem VN patterns in ML

	T1	**T2**
CətəCCv̄t	8	3
əCtəCCv̄t	4	6
CtəCCv̄t	0	0
CətCəCv̄t	0	0

Table 12: (Selection of) Surface T-stem VN patterns in Jahn (1902)

	T1	**T2**
CətəCCv̄t	4	5
əCtəCCv̄t	0	0
CtəCCv̄t	6	9
CətCəCv̄t	1	3

One is thus confronted with at least four surface varieties of T-stem VNs: *CətəCCv̄t*, *CtəCCv̄t*, *əCtəCCv̄t*, and *CətCəCv̄t*, all of which are spread across both the T1- and T2-stems,[19] begging the following questions:

(a) Is the initial vowel in the pattern *əCtəCCv̄t* in ML underlying or epenthetic?

(b) Is the presence or lack of a vowel between R1 and the *t*-infix a result of phonological processes (i.e., IG-effect and sonorant as R1)?

(c) Are the vowels between the infix and R2 and between R2 and R3 underlying (considering the distinct distribution between finite T-stem verbal forms T1/T2)?

When reviewing the evidence in ML against the backdrop of these questions, an incoherent picture emerges.

Concerning the question of the initial vowel (a)—in Johnstone's data most surface *əCtəCCv̄t* forms show a sonorant or IG

[19] A few additional surface configurations can be found in Jahn (1902), all of which seem explainable via IG/Sonorant effects, see below.

consonant as R1 (e.g., *əḥtərfūt*, *əntəġsūt*, *əwtəlmūt*, *ərtəḳyūt*, *əḥtəfḳawt*, *əftərtōt*, *ərtəfōt*, *ərtəwōt*). However, not all do, e.g., *əġtyəṭawt* (T1), *əġtəwṭōt* (T2). In these two counterexamples, a glide and /l/ are attested as R2. However, it does not seem immediately clear why this should influence the behaviour of R1 (at least in the case of the T1 verb). If one wants to accept these items as valid, one might imagine that the initial surface vowel actually represents the initial underlying vowel of finite T2-stem perfective verbs. However, if this was the case, it would seem completely unexplainable why this vowel would surface with some T1 and T2 stem verbs and not with others, namely with those of the surface pattern *CətəCCvt* (*lətəwḳawt*, etc., see below), where one would expect to see a pattern *əCtəCCvt* regardless of the nature of the initial consonant.

Concerning the question of a vowel between R1 and the t-infix (b)—if there was an initial underlying vowel, one might aprioristically assume that no vowel would surface between R1 and the *t-* infix. However, this is not always the case. Drawing upon phonological factors, one might assume, that the presence of a vowel is conditioned by the nature of R1—if R1 is an IG consonant or a sonorant, no surface vowel is to be expected; otherwise, one might expect a surface vowel to be present. However, in ML there are examples where this does not seem to apply, for no evident reason: *śətəwḳawt* – T1, not (*ə*)*śtəwḳawt* or (*ə*)*śtūḳawt*; *mətənyūt* – T2, not (*ə*)*mtənyūt*; and *mətəwyūt* – T1, not (*ə*)*mtəwyūt* or (*ə*)*mtūyūt*.

Concerning the rest of the base (c)—one notes that in Jahn's data at least, the IG-effect and the effect of sonorants do

seem to apply, hence the *CətCəCv̄t* pattern in items such as *gitfiōt*, *wutxerōt*, *wutkelōt*, and *atwurōt*, with all but *atwurōt* showing an IG consonant as R2. The underlying status of a vowel (or lack thereof) between the infix and R2 is, of course, not prejudiced by the presence of a surface vowel.

Hence, if one takes Johnstone's sparse data seriously, no coherent picture on the formation of T-stem VNs emerges, with both *CətəCCv̄t* and *əCtəCCv̄t* being attested for both T1- and T2-stems, and with no coherent application of phonological factors to the initial sequences of the noun. Jahn's data seem to suggest—despite inconsistencies[20]—that no underlying initial vowel exists and that the surface differences of T-stem VNs might be due to well-known phonological effects, hence suggesting a default T-stem VN pattern *Cət(ə)C(ə)Cv̄t*.

To clarify the picture these and further items of T1- and T2-stems were checked by the author:

[20] Non-application of IG: *ḥaterfōt*, *ḥatemiyōt* (both T1), *ḥateḟṭōt* (T2); lack of vowel after R1 in *ṣterfōt* (T1).

Table 13: T1- and T2-stem VNs (author's data)

Pattern	**Attestations**
CətəCəCūt	*ḳətənəmūt* (T1)
CətəCCūt	*gətənbūt* (T1, T2),[21] *ġətərbūt* (T2), *ḳətūbūt* (T1, II-*l*), *ūtəlmūt* (T2)[22]
əCtCəCūt	*əntfəzūt* (T1), *əntəġsūt* (T2), (*ə*)*rtəḳəyūt* (T1), *ərtfōt* (T1)[23]
CətCəCūt	*ūtxərūt/wətxərūt* (T2), *mətwəyūt* (T1), *gətfəyūt* (T1), *ūtkəlūt* (T2)[24]
CtəCCūt	*ftərtūt* (T1), *ftərkūt* (T1), *ftərḳawt* (T1), *ḥtərfūt* (T1), *ktūṯūt* (T1, II-*l*)
CtCCūt	*ftḥsūt* (T1), *ḥtfḳawt* (T1, T2), *ftsḥawt* (T1)

Concerning the appearance of an initial vowel (a)—the presence or lack of an initial vowel seems to be simply the result of the presence or lack of an IG consonant or sonorant as R1. Hence, any initial vowel should probably be considered epenthetic (or a definite article), rather than a cognate of the initial vowel of finite T2 perfect forms. Concerning the initial sequence *Cət_* or *Ct_* (b)—this seems to be also exclusively governed by the presence of an IG consonant or sonorant as R1. Hence, in the aforementioned examples no vowel surfaces between R1 and /t/, and in the case of R1 = [+Son.], an initial (epenthetic) vowel surfaces.

Also, the IG-effect was observed between the t- infix and R2, as in Jahn (1902), e.g., *gətfəyūt*, *ūtkəlūt*, etc. Hence, it seems that the aforementioned rules do in fact apply, and that the small number of contradictory examples given in ML contain a number of errors.

---

[21] It is also attested with an initial vowel, as in *əgtənbūt*, which was described as the definite article by informants.

[22] The form *ūtəlmēt* is also given for this verb. This an *ūtkəlēt* are the only example of T-stem VNs with a suffix *-ēt*, perhaps implying that these are errors.

[23] R3 = /ʿ/, hence *-ōt* (to be observed in other nominal patterns).

[24] See above, fn. 23.

Concerning the rest of internal structure of the base (c)—it should be noted that the assimilation of R2 to the *t*-infix of T1 stems, known from finite T1 perfect-forms, is not observed in T1 (and in T2) VNs. One hence finds the pattern *CətəCCv̄t* in these cases, e.g., *ḳətəllūt* (*ḳáttəl*), *nṭəbbūt* (*náṭṭəb*). This indicates the presence of an underlying vowel between the infix and R2.

Furthermore, if one accepts the notion that no underlying initial vowel is present, it would seem reasonable to assume that an underlying vowel must be present between R1 and the *t*- infix. Hence, one might assume that the initial sequence of T-stem VNs is underlyingly *Cət_*, although this is not empirically proven. As for the question of an underlying vowel between R2 and R3, this question is hard to prove on the basis of the Mehri data. J/S VNs would, however, suggest the presence of an underlying vowel, since /b/ and /m/ are dissolved as R3:

(23) *əġtɔ̄rɔ́t* (T2 *aġtéréb*), *əbtulūt* (T2 *əbtélím*)

Hence, the underlying pattern of the base of T-stem VNs can be assumed to be either *CətəCCv̄t* or—in the light of J/S, perhaps more likely—*CətəCəCv̄t*. Note the presence of an initial vowel in the aforementioned two J/S T2-stem VNs, the non-epenthetic status of which is underlined by the shift /w/ > /b/ in *əbtulūt*.

Concerning the vowel quality of the stressed suffix, no significant differences between T1- and T2-stems could be observed, with both surfacing as *-ūt* in Morris (1981) and in the data collected by the author. In more western varieties of Mehri, again, *-ūt* is less prominent than *-ōt*, although two attestations of a suffix *-ūt* could be found in Jahn (1902), e.g., *fteḳedūt* and *stiḥawt* (both T2).

2.5. Š-Stem VNs

In the case of the Š-stems, both *šəCCəCvt* and *šəCəCCvt* are attested. Similar to the H-stem, the distribution of *šəCCəCvt* and *šəCəCCvt* seems to be conditioned by the presence of a sonorant as R2 or a glottalic consonant as R1, and not by the stem of the corresponding finite verbal form (Š1/Š2). The majority of attestations from ML show *šəCCəCvt*, and the two examples of *šəCəCCvt* (*šəhəwbūt* and *šənəwṣawt*) show /w/ as R2. However, the examples of *šəCəCCvt* given by Jahn show a sonorant as R2.

Table 14: *šəCCəCvt* and *šəCəCCvt* in ML and Jahn (1902)

Stem	Pattern	Attestations
Š1	*šəCCəCūt*	*šəwgəśūt, šəwkəfūt, šəxbərūt, šəwzūt, šəśyəḳawt, šəġbərōt, šəġfərōt, šənḥərōt, šāsərōt*
Š2	*šəCCəCvt*	*šəwhəḳāt, šədxəlēt, šəbśərēt, šərgəlēt, šəxṭərūt*
Š1	*šəCəCCūt*	*šəhəwbūt* (+*šxargōt, šaḳarśōt* [Jahn 1902])
Š2	*šəCəCCvt*	*šənəwṣawt* (+*šḳarbōt* [Jahn 1902])

In ML one item is listed with a long vowel between R1 and R2, namely *šəxārəgūt* for the corresponding Š2 verb *šxērəg*. While this form might very well be an error, since it stands alone in the previously published sources, and neither this nor other Š2-stem items were produced with a long vowel in this position by the author's informants, it should be noted that 'secondary' stress would fall on exactly this position in the corresponding J/S forms (i.e., Š2 *šədɔ̄xɔlɔ́t*), showing (at least) phonetic lengthening of the vowel between R1 and R2.

In terms of the vocalism of the suffix, Jahn (1902) and Morris (1981) only record *-ōt/-ūt* and *-ūt*, respectively. In ML, however, one finds *-ūt* for Š1 and *-ēt* for Š2, with a few exceptions of *-ūt* for Š1. This image is mirrored in the data collected by the

author, where some informants would use *-ūt* and others *-ēt* for Š2-stem VNs.

2.6. Quadriliteral, Quinqueliteral, Reduplicated VNs

Morphologically more complex forms, such as quadriliterals, 'quinqueliterals', reduplicated forms, etc., broadly pattern in a similar fashion to the derived-stem VNs, insofar as they also generally take a stressed suffix. Attestations from both previous sources and from the author's data are not particularly numerous and examples are given in Table 15 (overleaf).

For simple quadriliterals of the type *aCáCCəC*, one finds the suffix *-ēt* and (with the potential exception of Johnstone's *əmərḥəbēt*?) no trace of an initial vowel, though Jahn (1902) attests three exceptions ending in *-ōt*: *karbelōt*, *maseblōt*, and *xarbəśōt*.

Table 15: Quadriliteral, 'quinqueliteral', and reduplicated VNs in previous sources and author's data

	Johnstone (1987)	**Jahn (1902)**	**Morris (1981)**	**Author's data**
4rad	*bərḳāt* (IV-ʾ) *əmərḥəbēt*	*karbelōt markaḥāt* *marḥabēt maseblōt* *xarbeśōt*	*dərbəśēt* *kərbəlēt* *ḳərbəṭāt, tərḏəmēt* *xərbəśēt*	*kərbəlēt*
n-	*ənḥəṭməlūt* *ənḳəlādūt*	*naḥayrrōt*	*ḳayta* *ḳərbəṭāt* *tərtərēt* *xərbəśēt* *xəśxəśāt* *ənḥətməlūt* *ənḥəṭbəlūt*	*ḳəfərrāt* *ngərdəśēt* *nšərxfēt* *əḳəfərrawt* *(ənḳəfərrawt?)* *ngərdəśūt*
š-	—	—	—	*šədərbəšēt*
redup.	*dəgdəgēt dəmdəmēt*	*kalkalōt* *ḳalḳalōt* *ḥuwaḥáwt*	*dəgdəgēt* *ḏəbḏəbēt* *ḳəśḳəśāt*	*dəmdəmēt*
5rad/R	*āfērráwt śxəwəllūt*	*kiriddōt* *metxoulīl*	*zəḥəwllēt* *śəġayrrāt*	*āfərrāt* *ḥəwərrāt/ḥawr* *mśxūlīl*

For patterns with a prefixed n-, the suffix appears as *-ūt/-ōt* when the prefix is preserved. There are, however, attestations where the prefix does not appear, and in these cases the suffix surfaces as *-ēt*, as to be seen in Morris (1981) and in the author's data.

In the case of a quadriliteral root prefixed by a *š-* prefix, the single example collected by the author shows the suffix *-ēt* (*šədərbəšēt*).

Reduplicated forms without an additional morpheme show the suffix *-ōt* (*-ūt*) in Jahn (1902) and *-ēt* in all other sources.

'Quinqueliterals' take either the suffix *-ēt* or *-ūt* (*-ōt* in Jahn 1902). The distinction in the vowel quality of the suffix does not seem to exactly match Rubin's distinction of Qw- and Qy-stem verbs (Rubin 2018, 160–61), insofar as one finds Qw-type VNs with *-ūt* (*śxəwəllūt* [Johnstone 1987]) and *-ēt* (*zəḥəwllēt* [Morris 1981]). More data from more speakers would be needed to investigate these forms in further detail.

2.7. Vowel Quality of the Suffix(es) of Derived-Stem VNs

As one has seen, various vowel qualities are attested for the suffixes of those derived-stem VNs with stressed suffixes. In Jahn's data the most frequently attested suffix is *-ōt*, although some attestations of *-ūt* can be found as well, mainly with III-*infirmae* roots. In Morris's data, all derived-stem suffixes show the suffix *-ūt*. In ML, the transcription of the suffix varies, and is given as *-ōt* before ⟨ġ⟩ and *-ūt* after ⟨ġ⟩, with some exceptions. In the case of ML, a significant divergence is found in the Š2-stem suffixes, which show *-ēt* and *-ūt*. In the author's data, the derived-stem

suffixes also surface as *-ūt*, with the exception of the Š2-stem, where some speakers would use *-ēt*, and others would use *-ūt*. Hence, most derived-stem VNs appear on the surface identical to their corresponding 3FS perfective forms, as shown in the table below.

Table 16: Vocalism of derived-stem VN suffixes and the 3FS perfective

	3FS perfective	**ML**	**Jahn (1902)**	**Morris (1981)**	**Author's data**
D/L	*-ēt*	*(-ēt*, 1 *-ōt)*	*(-ōt)*	*(-ēt)*	*(-ēt*, 1 *-ōt)*
H	*-ūt*	*-ūt/-ōt*	*-ōt (-ūt)*	*-ūt*	*-ūt*
T1	*-ōt*	*-ūt/-ōt*	*-ōt*	*-ūt*	*-ūt*
T2	*-ūt*	*-ūt/-ōt*	*-ōt (-ūt)*	*-ūt*	*-ūt*
Š1	*-ūt*	*-ūt/-ōt*	*-ōt (-ūt)*	*-ūt*	*-ūt*
Š2	*-ēt*	*-ēt; -ūt*	*-ōt (-ūt)*	*-ūt*	*-ēt; -ūt*

As can be seen, the only derived-stem VN where the vocalism of the stressed suffix is not identical with that of the corresponding 3FS perfective form is the case of the T1-stem, and the (speaker-based?) variation in the Š2-stem, whereas in all other derived stems (disregarding the obvious example of the D/L-stem), the vocalism of the VN and the corresponding 3FS perfective form seems to be identical. When considering exclusively the factor of vowel quality, one might infer that the distinction of the T1-stem might be due to the collapse of T1- and T2-stem VNs, with the common T-stem VN being based on the T2-stem VN, hence explaining the non-correspondence of T1-stem VNs and their 3FS perfective forms in terms of suffix vocalism. However, as has been shown, this assumption fails to account for the apparent lack of the underlying initial vowel of T-stem VNs. Whereas more data might be needed to explore this question in greater detail, on the basis of the data gathered by the author, it seems that at

least the T-stem VNs are formally distinct from their corresponding 3FS perfective forms, which might imply formal distinctness of the other derived-stem VNs as well.

3.0. Summary and Outlook

A summary of the most prominent G-stem VNs and derived stem VNs is given in the table below.

Table 17: Overview of the most prominent Mehri VN patterns

	Stem	Notes
CayCaC	*Ga* + *Gb*	Surface *CīCaC* in Jahn (1902)
CəCayC	*Gb, II-Gutt.*	Only EM in Johnstone (1987), Morris (1981), and author's data
CəCyūC	*Gb, II-Gutt.*	Only WM in Jahn (1902) and Sima (2009)
CaCC	*Ga* + *Gb*	Mainly triradical
CəCC		Mainly biradical
CV̄CaC	*Ga* + *Gb*	*CōCaC* and *CēCaC*
Others	*Ga* + *Gb*	Many, lexicalised
təCCayC	*D/L*	Surface *təCCīC* in Jahn (1902); some *təCCəCēt/ təCCōt* attested for D/L
həCCəCūt	*H*	*-ūt* in EM, *-ōt* in WM; *həCəCCūt* / R2 = [+Son.] or R1 = [+Glott.]
šəCCəCūt	*Š1, Š2*	*-ūt* in EM, *-ōt* in WM; *šəCəCCūt* / R2 = [+Son.] (+Glide?); *-ūt/-ōt* attested for Š2 in WM and at least amongst some speakers of EM
šəCCəCēt	*Š2*	Amongst some speakers of EM; *šəCəCCēt* / R2 = [+Son.] (+Glide?).
CətəC(ə)Cūt	*T1* + *T2*	*-ūt* in EM, *-ōt* in WM; other surface forms *əCtəCCūt*, *CtəCCūt*, *CətCəCūt*, *CtCCūt*, etc.
CəCCəCēt	*IVrad. redupl.*	*-ōt* also attested in Jahn (1902)
ənCəCCəCūt	*NQ*	Forms are also attested without initial *n-*, in these cases suffix *-ēt*
šəCəCCəCēt	*ŠQ*	*šədərbəšēt*

One has thus seen that amongst G-stem VNs the most prominent and apparently productive VN pattern is *CayCəC* (EM). This pattern is used for both Ga- and Gb-stem verbs and appears to be

distinct from nominal *CīCəC* due to the presence of an underlying glide. In other dialects of Mehri, this pattern appears to surface as *CīCəC*, and at least partly shows behaviour of a form with an underlying glide, with further research into the dialectology of Mehri being needed to ascertain the distinct behaviour of glides in Mehri. Cognates from other MSAL are also known for this pattern, which seem to strengthen the notion of an underlying glide, as well as the presence of this type of VN formation—with Dufour (2016), possibly historical **CaCīC* as a VN pattern in PMSAL. The VN patterns *CəCayC* and *CəCyūC* mostly show a guttural consonant as R2, and almost appear to be in complementary distribution with *CayCəC*, at least amongst the speakers consulted by the author. However, no conditioning factors between *CayCəC* and *CəCayC*/*CəCyūC* could be found for the corpus as a whole. Hence it seems most reasonable to consider these patterns as distinct for the moment.

Other G-stem VN patterns seem largely unpredictable and hence lexicalised, with a slight tendency towards *CvCC*-type patterns. In terms of the distribution between Ga- and Gb-stems, the only VN patterns which show distinct distribution (II-Gutt. and Gb) are *CəCayC* and *CəCyūC*.

For derived stems and quadri/quinqueliteral roots (with the exception of the D/L stem), the VN patterns in Mehri show similarities with the corresponding 3FS perfective forms. While a solution for the determination of the distribution of vocalism akin to the one proposed by Bendjaballah and Rubin (2020) might seem appealing, particularly when assuming the merger of T1- and T2-stem VNs into a VN pattern based upon the T2-stem

(see §2.4), the T-stem VNs also show the clearest formal distinctness from their corresponding 3FS perfective forms, implying underlying distinctness of at least the T-stem VN from its corresponding (T1 and T2) 3FS perfective forms, in spite of surface similarities. In broad terms, other MSAL show similar ways of derived-stem VN formation, insofar as the derived-stem affix surfaces and a stressed suffix *-ý(t)* is added.[25] Note, in particular, that the suffix vocalism of derived stem J/S cognates corresponds to the situation in Mehri insofar as that the stressed suffix vowel is identical in quality (/ɔ/ in J/S) to that of the 3FS perfective .

As has been mentioned, Dufour (2016, 376–78) has suggested historical **CaCīC* as the origin of the G-stem VN pattern *CayCəC*. Another suggestion comes from Bittner (1909), who suggested original **CiCC*. However, Bittner's proposal seems outdated, based on the different framework used. A VN pattern *CaCīC* (and patterns derived from it) is also found in other Semitic languages of the Arabian peninsula (at least in Classical Arabic), and more productively in Ethiosemitic (e.g., Gəʿəz *CaCiC* < **CaCīC*), with Mehri (and MSAL) *CayCəC* falling apparently into a broader southern Semitic (in an areal sense) trend of **CaCīC* as a G-stem VN pattern.

In the case of the derived-stem VNs, a salient point to mention for a comparative perspective lies in the fact that these forms show both the derived-stem affix and a stressed suffix *-ýt*. The underlying structure remains unclear for most derived-stem VNs.

---

[25] E.g., J/S *tfkír* (D/L), *ğitenfɔ́t/ktɔlṯɔ́t* (T1), *ftɔkɔrɔ́t* (T2), *šɔġfɔ̄rɔ́t* (Š1), *šəlōḥɔḳɔ́t* (Š2), *nğɔrdɔśɔ́t/ḳafarrɔ́t* (NQ), etc. (own data). For Soq. cognates see Kogan and Bulakh (2019).

However, the case of the T-stem VN urges caution against the assumption that the surface similarities of derived stem VNs and their 3FS perfective counterparts are to be understood as formal identity. In any case, regular and fully productive derived-stem VN formation by simple suffixation of a stressed suffix *-vt* onto what appears to be the base of the derived stem—although quite possible, with an internal configuration of underlying vowels distinct from the corresponding 3FS perfective—appears to be a particular MSAL feature.

Hence, VN formation in Mehri adheres to the common Semitic notion of a plethora of VN patterns for the G-stem(s) with more regular patterns for the derived stems. Some of the patterns attested show parallels with other relatively productive VN patterns in most Semitic languages, whether through borrowing (D/L *təCCayC*?) or retention (*CvCC*?), and the most productive G-stem VN pattern *CayCəC* appears to show parallels to forms particular to Arabic and Ethiosemitic (as VNs). Finally, as alluded to previously, putative cognates for the Mehri VN patterns discussed here are to be found in other MSAL as well, thus alluding to a basically similar repertoire of VN formation in PMSAL (and hence potentially providing an interesting set of PMSAL innovations), although more research is needed, particularly on the other MSAL, to further advance this question.

References

Bendjaballah, Sabrina, and Philippe Ségéral. 2014. ‘The Phonology of “Idle Glottis” Consonants in the Mehri of Oman

(Modern South Arabian)'. *Journal of Semitic Studies* 49/1: 161–204.

———. 2017. 'The Vocalic System of the Mehri of Oman: Stress, Length and Syllabic Structure'. *Brill's Journal of Afroasiatic Languages and Linguistics* 9: 160–90.

Bendjaballah, Sabrina, and Aaron Rubin. 2020. 'The 3fs Perfect in Omani Mehri'. *Journal of Semitic Studies* 65/2: 511–29.

Bittner, Maximilian. 1909. *Studien zur Laut- und Formenlehre der Mehri-Sprache in Südarabien. I. Zum Nomen im engeren Sinne.* Vienna: Hölder.

Dufour, Julien. 2016. 'Recherches sur le verbe sudarabique moderne'. Habitation, École Pratique des Hautes Études.

———. 2017. 'Nouns and Adjectives of the Shape $C_1\acute{V}C_2(ə)C_3(-)$ in Jibbali (Śḥri) and Mehri'. *Brill's Journal of Afroasiatic Languages and Linguistics* 9: 191–217.

Jahn, Alfred. 1902. *Die Mehri-Sprache in Südarabien* III. Vienna: Hölder.

Johnstone, Thomas Muir. 1981. *Jibbali Lexicon.* Oxford: Oxford University Press.

———. 1987. *Mehri Lexicon and English-Mehri Word-List.* London: Routledge.

Kogan, Leonid. 2015. *Genealogical Classification of Semitic: The Lexical Isoglosses*. Berlin: De Gruyter.

Kogan, Leonid, and Alexander Militarev. 2000. *Semitic Etymological Dictionary. Volume I: Anatomy of Man and Animals*. Münster: Ugarit-Verlag.

Kogan, Leonid, and Maria Bulakh. 2019. 'Soqotri'. In *The Semitic Languages. Second Edition*, edited by John Huehnergard and Naʿama Pat-El, 280–320. London: Routledge.

Kogan, Leonid, and Vitaly Naumkin (eds). 2018. *Corpus of Soqotri Oral Literature* II. Leiden: Brill.

Morris, Miranda. 1981. 'The Phonology and Morphology of the Mahri Noun'. PhD dissertation, SOAS.

Rubin, Aaron. 2014. *The Jibbali (Shaḥri) Language of Oman: Grammar and Texts*. Leiden: Brill.

———. 2018. *Omani Mehri: A New Grammar with Texts*. Leiden: Brill.

Sima, Alexander. 2009. *Mehri-Texte aus der Jemenitischen Šarqiyah*. Wiesbaden: Harrassowitz.

Watson, Janet. 2012. *The Structure of Mehri*. Wiesbaden: Harrassowitz.