A Historical Reconstruction of Some Pronominal Suffixes in Modern Dialectal Arabic

Phillip W. Stokes

Department of Modern Foreign Languages and Literatures, The University of Tennessee,
Knoxville, TN 37996, USA; pstokes2@utk.edu

Abstract: The morphology of the pronominal suffixes in dialectal Arabic are of particular interest for scholars of the history of Arabic for two main reasons. First, multiple dialects attest suffixes that, from a comparative perspective, apparently retain final short vowels. The second and more complicated issue concerns the vowels which precede the suffixes in the dialects, which are thought to either have been case inflecting or epenthetic. In this paper, I take up Jean Cantineau's "embarrassing question" of how to account for the development of the vowels of the pronominal suffixes. Based on data from dialectal *tanwīn* in modern dialects, and attestations from pre-modern texts as well, I will argue that the pre-suffix vowels did originate in case inflecting vowels, but that no historical model heretofore proposed can satisfactorily account for how the various dialectal forms might have arisen. I identify two major historical developments and propose models for each. First, I suggest that dialects in which the pre-suffixal vowels harmonized with the suffix vowels developed via a process of harmonization across morpheme boundaries before the loss of final short vowels. For dialects in which one vowel is generalized, I argue that a post-stress neutralization took place, which led to a single vowel both before suffixes and *tanwīn* as well. Finally, I rely on evidence from the behavior of the suffixes to argue that the final vowel of the 3fs suffix was originally long, but that those of the 3ms, 2ms, and 2fs were most likely short.

Keywords: Arabic dialects; historical dialectology; historical linguistics

1. Introduction

The morphology of the pronominal suffixes in dialectal Arabic are of particular interest for scholars of the history of Arabic for two main reasons. First, multiple dialects attest suffixes that, from a comparative perspective, apparently retain final short vowels: dialectal *daras* "he studied", < **darasa* but *ʾabū-ki* < **ʾabVV-ki*, "your (fsg) father." This is despite the fact that most assume a complete loss of final short vowels in the pre-modern ancestors of all Arabic varieties. The second, and more complicated issue concerns the vowels which precede the suffixes in the dialects. In Classical Arabic (henceforth ClAr), this position was occupied by a short case vowel: *bāb+u/i/a +ka* "your door (nom/gen/acc)." In the modern dialects, which lack morphosyntactic case marking, these vowels vary: *bāb-Vk* "idem." Several attempts have been made to account for the development of these vowels from a functional case system ([Diem 1991](#); [Jastrow 1991](#); [Blau 2006](#)), though none has yet proved entirely satisfactory. The thorniest issue from a historical perspective is how to account for the functional shift of the pre-suffixal vowels, from morpho-syntactic inflection to their current forms, while also assuming the loss of final short vowels, which should have eliminated several of the vowels of the pronominal suffixes. Alternatively, several scholars have recently argued that the ancestors of the dialects came from varieties of Arabic that lacked case completely ([Retsö 1994](#); [Owens 2006](#), [2018](#)). In this alternative scenario, the pre-suffixal vowels were epenthetic vowels inserted to resolve increasingly un-tolerated consonant clusters.

In this paper I argue in favor of the case vowel origin and propose reconstructions that address both issues: final vowel length and the nature and development of the pre-

suffixal vowel. Following a discussion of the previous proposals and their deficiencies, I discuss both of these questions, with special focus on the problems posed by the third and second person singular forms. I will argue that data from remnants of case vowels before *tanwīn* in the dialects (dialectal *tanwīn*), as well as historical data in which case was retained longer before suffixes than elsewhere, can help understand the development of these vowels. The implications of these observations for the question of the historical length of the vowels of the suffixes will be discussed at length in Section 3. The ultimate goal of these reconstructions is to gain deeper historical insight into the most common surface forms attested across the spectrum of Arabic dialects, as well connect them together in the broader picture of Arabic linguistic history.

2. Pre-Suffix Vowel: Case Vowel or Epenthesis?

The pronominal suffix paradigms for the modern Arabic dialects do not inflect for case but nevertheless attest vowels before the pronominal suffixes. The following data (Table 1) serve as examples of the regular realizations of these suffixes, regardless of syntax (Fischer and Jastrow 1980, p. 81; Jordanian from Al-Wer (2007, p. 510)):

Table 1. Levantine and N. African Singular Pronominal Suffixes.

	Jordanian	**Cairo**	**Damascus**	**Tunis**
3ms	*-o, -V:*	*-uh, -h*	*-o, -(h)*	*-u, -h*
3fs	*-ha, -a*	*-ha*	*-a, -ha*	*-ha*
2ms	*-ak, -k*	*-ak, -k*	*-ak, -k*	*-ik, -k*
2fs	*-ik, -ki*	*-ik, -ki*	*-ik, -ki*	*-ik, -k*

The origin and nature of the vowels preceding the pronominal suffixes in the modern dialects has become rather controversial. They have traditionally been interpreted as frozen reflexes of case vowels,[1] which eventually harmonized with the short vowels of the suffixes, e.g., nominative *u* was frozen based on harmony with 3ms *hu, etc. (Birkeland 1952, pp. 12, 19; Cantineau 1936/1937, p. 180, vol. 2; Fischer and Jastrow 1980, p. 42; Diem 1991). To account for the vowel harmony, which would ostensibly be possible only after the breakdown of case, and thus presumably the loss of final short vowels, Blau (2006, pp. 87–88) appeals to the intentionality of speakers, who decided to avoid a homophony between masculine and feminine forms, which after the breakdown of case and loss of final short vowels, should have both become **bāb-k* "your (msg and fsg) door." Instead, they inserted an *a* vowel before the suffix for the masculine and an *i* to indicate the feminine. Blau cites a phenomenon in ClAr called *naql*, "transfer", wherein speakers would insert a vowel homophonous with the suffixed case vowel on noun of the pattern of CVCC: *al-bakru* (nom sing) > *al-bakuru*, etc. (ibid., p. 88). However, Blau does not address the biggest issue with his proposal. Given that he believes that a *naql* would have occurred only once a final cluster had been created, it is unclear how speakers would have retrieved the etymological *a* vowel to mark masculine or the *i* vowel for the feminine if they had been lost word-finally.

Diem (1991) offers a different explanation. He recognizes that if loss of case was caused, or at least accompanied by loss of final short vowels, then the vowels of at least some of the pronominal suffixes, like 3ms *-hu* and 2ms *-ka*, should have also been lost. If such was the case, however, then of course the final vowel, with which the case vowel would harmonize, would have been eliminated. Diem, still believing the initial vowels to be remnants of case vowels, argues that the breakdown of the case system was caused by syntactic factors, especially the redundant nature of the Semitic case system, and not the result of a phonetic loss of short final vowels. In Diem's scenario, case breakdown resulted initially in a state where the short case vowels were in free variation (ibid., pp. 301–3). Eventually, before pronominal suffixes they were harmonized with the suffix vowel and frozen; in word-final position (including, for Diem, before *tanwīn*), one case was levelled.[2]

This is the state of affairs in what Diem calls "nomadic dialects" (i.e., Bedouin dialects) (ibid., pp. 303–4). Most non-nomadic dialects eventually witnessed the total loss of these remnants due to their lack of any meaningful syntactic function.

This is an attractive explanation[3] because it solves the main objection to the traditional assumptions (see above regarding Blau's 2006 proposal), namely that case was lost due to loss of final short vowels. If case marking was reanalyzed as marked solely by word order etc., then these variants might have eventually been in free variation.[4] It should be noted here that Diem recognizes the loss of final short vowels at some point, but argues that the breakdown of case was not related to this loss (or, probably, multiple losses), and presumably preceded it.

Alternatively, Owens (2006, chp. 8) has argued that these initial vowels originated as epenthetic vowels, completely unrelated to case vowels, the quality of which varied depending on the following consonant/vowel: 3ms *hu > u-hu (addition of epenthetic vowel) > uh (loss of final short vowel).[5] Owens's arguments against reconstructing case in the ancestor of all Arabic varieties is unconvincing. There is thus no reason a priori to suggest these vowels cannot be remnants of case vowels. Still, the complete loss of case vowels could have led to their loss before pronominal suffixes as well, presumably in this case, per Blau (2006), via the levelling of pausal forms to non-pausal position, which in turn could have led speakers to insert epenthetic vowels to resolve these newly-created consonant clusters. In other words, since Ø-marked forms were levelled to other contexts where they would be expected to be preserved if the loss of case were the result of a regular phonetic change (*v > Ø/C_#), such as the construct, then it is also possible that these Ø-marked forms were levelled to the position before pronominal suffixes.

Deciding between these options is difficult due to the uneven knowledge of the contemporary Arabic dialects, and the virtual absence of historical data for the ancestors to the modern dialects. Nevertheless, a careful examination of the implications of these proposals yields meaningful insights. If we imagine that the vowels in question originated as case vowels, then we must account for speakers presumably sacrificing the morphosyntactic role of case vowels before pronominal suffixes in favor of a non-phonemic epenthetic function. Scholars have thus mostly assumed that this happened following the breakdown of case (cf. Diem 1991). Most likely, then, the loss of case marking on the un-suffixed nouns motivated the levelling of one form, as in the external masculine plural oblique form *īn*. The presence of very similar suffix forms in Aramaic, and probably for Hebrew as well, provide attractive parallels (Table 2; Aramaic and Hebrew data taken from Hasselbach 2014, pp. 204–5):

Table 2. Reconstructed Aramaic (Syriac) and Hebrew Pronominal Suffixes.

	Aramaic (Syriac)	**Hebrew**
3ms	malk-eh < *malki-hi (Gen)	malkō < *malku-hu(:) (Acc)
2fs	malk-ek̲(y) < *malki-kī [1] (Gen)	malkēk̲ < *malki-ki (Gen)
2ms	malk-ak̲ < *malka-ka (Acc)	(non-Tiberian) malkak̲ < *malka-ka (Acc) [2]

[1] As indicated by the mater lectionis <Y> of the Aramaic form, the *i* vowel here was apparently long. Benjamin Suchard has suggested to me (p.c.) that this long ī was the result of *ki > kī based on contamination with the imperfect ending. [2] The Tiberian forms are complex. In non-pausal position, Tiberian *malkə-kā*, but pausally, *malkēkā*. The latter form suggests < *malki-kah (Suchard, p.c.), or possibly analogy with the feminine singular form (Suchard 2020, pp. 194–96).

There are several reasons to favor positing etymological case vowels as the origin of the pre-suffixal vowels instead of originally vowelless consonant clusters that were only later broken up via epenthesis. First and foremost is a methodological argument. We can confidently reconstruct case for both Proto-Semitic and Proto-Arabic based on comparative and internal data. A case system in which singular (and broken plural) nominal forms are inflected for three cases (nominative, genitive, and accusative), while duals and sound masculine plurals are inflected for two (nominative and oblique) is attested in Akkadian, Ugaritic, and Canaanite, as well as ClAr (Huehnergard 2019, pp. 60–61). Classical Ethiopic

(Gǝʕǝz) also attests a remnant of the same system, with an accusative/non-accusative system in which accusative is marked with short-*a*, and non-accusative is unmarked, likely as the result of the merger and loss of final *u and *i (>ǝ > ø/C_#) (Butts 2019, pp. 128–29).

Within Arabic, it is not only ClAr that attests this nominal case marking; it is attested, to varying degrees, in pre-Islamic Arabic corpora in the Safaitic (Ancient North Arabian), Nabataean, and Greek scripts as well (Al-Jallad 2018). As argued in Van Putten and Stokes (2018), the Quranic consonantal text (*rasm*) attests a partial case system, distinct from that attested in ClAr. Finally, from a methodological perspective, we should greatly prefer an explanation for the lack of a feature that is most parsimonious and consistent with standard historical linguistic methodology. In attempting to answer these questions we are faced with the scenario either that the Semitic languages in general, and Arabic specifically, should be sub-grouped historically based on the presence or absence of a single feature, or that nominal case marking, which is known to have been lost in a number of languages, was retained completely in some, only partially in others, and lost completely in the rest. The latter is more parsimonious than the former.

Further, Owens' argument that these vowels originated as epenthetic vowels is not as strong as his claims make it seem. Owens himself reconstructs these pronominal suffixes with final short vowels, which were only lost after vowel harmony with the newly inserted epenthetic vowels (Owens 2006, p. 248 ff.). He further argues that when a pronominal suffix is added to a word like *qalb* "heart", the combination would result in a chain of three consonants, CCC, which is often (but not always; see Watson (2007, pp. 340–41)) disallowed in Arabic dialects (ibid., pp. 107–11). This resulted in the insertion of an epenthetic vowel between the stem and the suffix: *qalb* + *hā* > *qalbhā* > *qalb-a-hā* "her heart." The same process was responsible for creating 2ms *ak*: *qalb* + *ka* > *qalbka* > *qalb-a-ka* "your (ms) heart", and 2fs *qalb* + *ki* > *qalbki* > *qalb-i-ki*, and so on. After this stage of epenthetic vowel insertion, final short vowels of these suffixes were lost: *qalb-a-ka* > *qalb-ak*.

While on the face of it this explanation might seem to account for the dialect data without resorting to the notion of case vowels, it is mitigated by some further considerations. First, for Owens's explanation of these forms to obtain, we would need to assume that all dialects resolved these CCC clusters in the same way, namely by inserting the epenthetic vowel after the second of the three consonants, i.e., CCCv(:) > CCvCv(:). In fact, however, as Owens implicitly acknowledges (ibid., pp. 108–9), many dialects resolve these clusters by inserting the epenthetic vowel after the initial consonant, so CCCv(:) > CvCCv(:) (on which, see Watson (2007, pp. 340–48)). In such cases, we would expect two different allomorphs of the suffixes to exist, distributed according to the phonotactic rules of the dialect, but this is not the case. That is, if CVCC nouns were the origin of the epenthetic vowels, as Owens maintains, then we would expect a distribution which matches the two patterns of epenthesis insertion: *qalb-a-k* for CCC > CvCC, and **qalabk for CCC > CCvC. Instead, we only find forms that would need to have originated in CCvC-inserting dialects.

Another question that Owens does not meaningfully answer is why epenthesis would have been necessary in most nominal patterns. Owens's examples, cited above, are typically CVCC (including CVW/YC), i.e., *qalb-ak* "you (msg) heart" and *bēt-ha* "her house", etc. Elsewhere epenthesis is not typical. For example, most dialects do not resolve CCV clusters: Jordanian (a CVCC dialect) *katabti* "you (fsg) wrote", and not **katabiti "idem" (Al-Wer 2007); Cairene (a CCVC dialect) *ġasalti* "you (fsg) washed", and not **ġasaliti (Woidich 2006, p. 330).[6] As Owens admits, the pronominal suffixes, with the exception of 3fs (*hā) and 1cp (*nā), are to be reconstructed with final short vowels (reviewed above) (Owens 2006, pp. 239–58). Based on the dialectal data on which Owens relies, epenthesis at this initial stage, where the final short vowels of the suffixes were still present, would only be required for nouns of the pattern CVCC. It seems rather implausible that allomorphs necessary only for nouns of one (admittedly frequent) noun pattern were leveled everywhere. However, it is not impossible. If there are any dialects where Owens' reconstruction should be expected to accurately match the distribution, it is thus CCVC-patterning dialects, like Cairene (Watson 2007, p. 341).

The decisive blow to Owens's reconstruction, however, comes upon further examination of the Cairene distribution. The Egyptian realization of **qalb-a-hā* "her heart" is *ʾalb-a-ha*, **qalb-a-ka* "your (msg) hearth" is *ʾalb-ak*, and **qalb-i-ki* "your (fsg.) heart" is *ʾalb-ik*, each as Owens's reconstruction would predict. However, if the forms with epenthesis were generalized, as could be suggested by the 2ms and 2fs forms, we would expect 3fsg -aha to occur when suffixed to any noun, regardless of morphological pattern. Instead, however, we find that on nouns ending in a single consonant, the form is -ha, i.e., *baʾarit-ha* "her cow", and not ***baʾaritaha*. The same noun, however, when suffixed with a 2ms or 2fs suffix, would have the supposedly epenthetic allomorph, i.e., *baʾaritak* (msg.)/*baʾaratik* (fsg.). In other words, in a dialect that fits Owens' pattern for the creation of epenthetic allomorphs, we would have to suppose that the epenthetic allomorphs of the 2ms and 2fs suffixes were generalized everywhere, but the 3fs suffix retained its original distribution (i.e., occurring only after CC clusters). Thus, Owens's theory does not successfully explain the particulars of the distribution of any dialect type.

The question remains how to model this development from etymological case vowel to the various surface forms. As noted above, both Blau and Diem believe, as I do, that nouns originally ended in a short vowel that marked nominal case on singular, feminine plural, and most broken plural patterns. The question of their development is tied up with that of the simplification and eventual loss of Proto-Arabic morphosyntactic case inflection. For Blau, the breakdown and loss of case is related to the loss of final short vowels, and the subsequent analogical extension of these caseless forms to non-final position. For Diem, on the other hand, case breakdown was only partially, if at all, related to short final vowel loss; rather, for him it was primarily due to the low functional yield of case in Semitic and Arabic.

While both Blau and Diem are undoubtedly correct that both analogy and sound change played roles, neither provided explanations which, in the end, were able to provide satisfactory reconstructions of the process. Blau is ultimately unable to compellingly answer how short final case vowels were lost, thus leading to case loss, without the simultaneous loss of short pronominal suffix vowels. Diem, on the other hand, does not manage to explain what led to the free variation of case vowels, and ultimately the loss of case. As Blau argued successfully elsewhere (Blau 1972), the low functional yield of case in Semitic is not, in and of itself, an explanation for its loss.

Another problem that plagues both accounts detailed above is that neither fully integrates the only other feature which provides direct evidence of case vowels that remain in the dialects, namely dialectal *tanwīn*. Both Blau (2002, p. 44–46; 2006) and Diem (1991, 303 ff.) assume that the attested surface forms of dialectal *tanwīn* reflect unaltered the case vowel which was ultimately generalized. That is, in their reconstructions, any examples of dialectal tanwīn -in are automatically assumed to reflect a frozen genitive *-in*, any dialectal -an reflects accusative *-an*, etc. As argued in Stokes (2020), however, there are reasons based on the dialectal data to doubt this straightforward identification. The clearest example is found in the southern Saudi dialects of Bal-Qarn and Banī ʿAbādil, where an interesting pausal/non-pausal distribution of dialectal *tanwīn* is attested: non-pausal *bēt-in* "a house"/pausal *bēt-u* "idem." The most likely interpretation of this is a vowel merger to a high vowel in final position, wherein before *tanwīn* the surface vowel is short front -i, whereas when word-final the vowel is back high -u (itself apparently lengthened rather than lost; see Stokes (2020, pp. 655–56) for details).

Other evidence suggests a similar merger occurred in most, if not all, ancestors of the modern dialects. For example, in both Bahraini and rural (*fellāḥī*) Iraqi dialects, there is an adverbial tanwīn -an that contrasts with the standard dialectal *tanwīn* form -in; crucially, these occur in contexts which could not have been ClAr borrowings (Stokes 2020, p. 650). Rather, the reflexes of dialectal *tanwīn* attest a stage in which phonemic distinction was lost between short case vowels before tanwīn: **u,*i,*a* > V/C_#(n) (Stokes 2020, pp. 655–60). It is likely that such a merger is responsible for the surface forms of dialectal *tanwīn* in those dialects that retain it.

Finally, the traditional accounts offer only one set of reconstructions for dialects that attest a great deal of variation. In what follows, I will bring together data from dialects across the spectrum in order to provide historical explanations for each of the major patterns attested therein. Where possible and relevant, I incorporate data from dialectal *tanwīn* to inform the reconstruction. I will attempt to show that while the patterns attested in the dialects can effectively be derived from the same Proto-Arabic case-bearing situation, they require different scenarios—each of which differs from those proposed heretofore—to explain the surface manifestations. It is to the data, then, that we now turn.

3. Historical Development of Pronominal Suffixes

Obviously, there is not sufficient space here to treat every attested form. Further, the reconstructions offered here are not final; rather, it is hoped that they will constitute just one more step in the process of understanding the history of the dialects and their interrelationships. There are undoubtedly dialect-specific factors necessary to understand the particular histories of these suffixes.

Table 3 (below) illustrates most of the major strands of morphological variation in the dialectal suffix forms and will constitute the bulk of the data upon which the subsequent discussions will focus. The following should be considered a representative sampling, and are not exhaustive (Najdi from Ingham (1982, pp. 96–98; 2008, p. 328); Ṣanʿāni from Rossi (1939, p. 4), Watson (2009, p. 110), Isaksson (1991, p. 127)); Mardin Arabic from Grigore (2007, p. 228); Levantine from Brustad and Zuniga (2019, p. 411).

Table 3. Pronominal Suffixes from Sample Arabic Dialects.

	Najdi	**D̲afār**	**Ṣanʿāni**	**Mesopotamian Qǝltu (Mardin)**	**Levantine**
3ms	*C-ih, C/V-h*	*C-eh, V-h*	(Rossi) *C-ah, V-h* (Watson) *-ih, V-h*	*-u, -hu*	*-o, -u*
3fs	*(a)ha*	*-ha*	*C-aha/V-hā*	*-a, -wa, -ya*	*-a, -ha*
3mp	*-(i)hum, -ham*	*-hum*	(Rossi) *-[o]hum* (Watson) *-uhum*	*-ǝn, -wǝn*	*-(h)on, -(h)in, -hum*
3fp	*-(i)hin, -(a)hin, -han*	*-hin*	(Rossi) *-[e]hin* (Watson) *-ahin*	*N/A*	*-hin*
2ms	*C-ik, C/V-k*	*C-uk, V-k*	*C-ak/V-k*	*-ǝk, -k*	*-ak, -ek*
2fs	*C/V-(i)ć ~ š*	*C-iš, V-š*	*C-iš/V-š*	*-ki*	*-ik, -itš, -tši*
2mp	*-(i)kum, -kam*	*-kum*	(Rossi) *-[o]kum* (Watson) *-ukum*	*-kǝn*	*-kon, -kun, ---kin, -kum, -kam, -kim*
2fp	*-(i)ćin, -ćan*	*-kin*	(Rossi) *-[e]kin* (Watson) *-akin*	*N/A*	*-kin, -tšin, -tšan*
1cs	*C-i/V-ya*	*-i*	*C-i/V-ya*	*-i, -ya*	*-i*
1cp	*-(i)na, -a(na)*	*-na*	*-[a]na*	*-na*	*-na*

3.1. Third Person Masculine and Feminine Forms

Numerous and geographically widespread modern dialects attest 3ms C-*u*/C-*o*, and -*V:*. Regarding forms with -*o(h)*, Cantineau 1939) argued that it is the result of *dār* + frozen accusative *a* + *hu*, with subsequent elision of the intervocalic laryngeal:

dār-a-hu > dāra-u > dār-o.

Jastrow (1991) assumes that *dār-u(h)/dār-o*, as well as *dār-a(h)* and *dār-ih* "idem", were the result of the "freezing" of different case vowels in different dialects. In his scenario, *dār-u(h)/dār-o* attested the retention of nominative **-u*, whereas *dāra(h)* attested a frozen accusative **-a* and *dāri(h)* a frozen genitive **-i*. However, like Blau and Diem, Jastrow fails

to provide a scenario which accounts for the development of each. We are not told why the nominative case would be frozen only when followed by the 3ms suffix.

I agree that the 3ms suffix forms -*u(h)* ultimately derives from *u-hu. As we saw above, explaining these forms involves a model that can explain the freezing of the case vowel before PN suffixes before the loss of word-final short vowels. I would suggest the following steps. From the reconstructed Proto-Arabic distribution, vowel harmony across morpheme boundaries developed, wherein V^1-$CV^2 > V^2$-CV^2.[7]

1. Proto-Arabic: **bayt-V^{case}n*; **al-bayt-V^{case}*; **bayt-V^{case}-hu*
2. Harmony Rule: **bayt-V^{case}n*; **al-bayt-V^{case}*; *bayt-u-hu/*bayt-a-ka/*bayt-i-ki

    Subsequently, final short vowels were lost.

3. Loss of final short vowels: **bayt-V^{case}n;'* *al-bayt; *bayt-uh

    And finally, the neutralization of phonemic contrast before *tanwīn* occurred:

4. Neutralization of contrast before tanwīn: *bayt-Vn*; *al-bayt*; *bayt-uh*

Steps 3 and 4 are not necessarily linear; they might have been simultaneous, or 4 might have preceded 3. Unlike Jastrow et al., however, I also suggest that 3ms -*o(h)* forms originate in *u-hu as well, and not *a-hu. The raising of *u to *o* is widespread,[8] and is likely behind the realization of the 3ms-*o(h)* forms as well.

Similar types of harmonization are attested in Arabic. For example, Sibawayh (1988, p. 173) mentions a type of assimilation that takes place in pause in which CVCC nouns insert a vowel to break up the final cluster where the second vowel harmonizes with the first:

ʾiblun "a camel", > *ʿibil* "idem",

ḥulmun "a dream", > *ḥulum* "idem."

Blau's discussion of *naql* "transfer" represents a similar process of harmony, wherein a noun in pause inserts an epenthetic vowel that is the quality of the vowel which, when in non-pausal position, would mark its syntactic case:

$$bakrun > bakur$$

$$bakrin > bakir$$

$$bakran > bakar$$

This process[9] is the same kind as is proposed here ($CV^1CV^2 > CV^2CV^2$). It is important, however, to emphasize the difference between the present proposal and Blau's and Diem's. Blau's appeal to *naql* to explain the pronominal suffix is predicated on the loss of final short vowels as the driver of the breakdown of the case system, which would eliminate the source of the transfer—the final vowels of the pronominal suffixes—he proposes. In the present proposal, a cross-morpheme harmony rule develops before the loss of final short vowels. Case on word-final nouns, as well as before *tanwīn*, likely continued for some time. A round (or several rounds) of word-final short vowel loss subsequently contributed to the loss of case, although it is likely that this process was gradual. The present proposal also differs with Diem's insofar as it proposes phonological rules to account for each stage, whereas Diem posits a period of randomness in the realization of etymological case vowels that is ad hoc.

Returning to the discussion of *dār-u(h)* and *dār-o(h)*, the -*u* and -*o* in *dār-u* and *dār-o* need not therefore be frozen nominatives as such; rather, they should be considered etymological case vowels that have harmonized with the vowel of the suffix. The point is, to my view, significant and thus bears repeating. While the argument here is that the vowels that precede the pronominal suffixes were etymologically case vowels, their development in harmonizing dialects (i.e., where the vowel matches the quality of the PN suffix) represents a stage at which the vowel harmonized regularly with the vowel quality of the following morpheme. It is therefore rather meaningless to speak of "frozen nominative" when referring to the *u*. Not only does this argument account for how the

vowels began to harmonize before the final short vowels of the suffixes were lost, but it also avoids the linguistically ad hoc free variation proposed by Diem.

In addition to forms with -uh or -o(h), a minority of modern dialects, however, attest a form -hu, which quite possibly derives from *-hū. In the Arabic of Mardin, for example, post-vocalic 3ms suffix is -hu rather than a lengthened vowel -V:, e.g., abūhu "his father" (Grigore 2007, p. 228; Talay 2011, p. 913). In nearby Sīrt, as in most Qəltu dialects, short high vowels have completely merged to /ə/ in every phonetic environment (Talay 2011, p. 913). Thus Sīrt məfləs (<*muflis) "broke", and pronominal suffixes kən/ən (<*kun/*kin and *hun/*hin) "your (cpl)/their (cpl)" (Grigore and Bițuna 2012, p. 551). We would thus expect the 3ms suffix, derived from *u-hu or *uh, to be realized -əh as well. In some SW Yemeni dialects, -hu is ubiquitous whether the noun ends in a consonant or a vowel.

Comparative Semitic data suggest reconstructing Proto-Semitic 3ms suffix with a short vowel, *-su. Additionally, both Hebrew and Aramaic forms suggest a Proto-Northwest Semitic *-hu (Suchard 2020, p. 43). The Arabic dialectal forms ending in -u(h) and -o(h), if the present reconstruction is correct, also suggest a short vowel *-hu. If the Arabic dialectal -hu forms do indeed reflect etymological *-hū, it could suggest that two by-forms existed at the Proto-Arabic node: *-hu and *-hū. There is evidence of long and short vowel variants of pronominal suffixes throughout Semitic. In ClAr, for example, pronominal suffixes following short case vowels are realized long, while those after long vowels were realized short: bi-hī "with it", but fī-hi "in it" (Fischer 2002, p. 126). The opposite of this system is attested elsewhere, in, e.g., Aramaic and Gəʿəz, where suffixes with short vowels occur after short vowels, and suffixes with long vowels typically occur after long vowels, cf. Biblical Aramaic (Ezra 5:11) ʿaḏōhī "his servants." This distribution led Cantineau (Cantineau 1936/1937, vol. 2) to propose a sort of quantitative harmony, where the length of the suffix vowel harmonizes with the vowel preceding the suffix. A further explanation, offered in Hasselbach (2004), is that the suffixes with short forms are original, with the long forms the result of contamination with the independent form. It is quite possible that various analogies and developments have led to the distribution attested in the various languages (Suchard 2020, pp. 198–214). Whatever the case may be, the presence of long and short forms of the suffixes, originally tied to preceding syllable structure, is very possibly behind the different contemporary realizations of these suffixes across the dialects. This is undoubtedly the explanation for the common distribution of the 2fs suffix, cf. Damascene C-ik, but V-ki, where the vowel is always historically long (Lentin 2006, p. 548). The u forms in Qəltu dialects such as Mardin could represent a levelling of the long form to all contexts, with a subsequent loss of the laryngeal h after consonants, thus C-u but V(:)-hu.[10]

Another possibility is that -hu forms in some dialects reflect analogical protection of 3ms suffixes, based on analogy with the independent hū-based forms. Paradigm pressure can deter an otherwise regular phonetic shift from occurring. For example, the shift in West Semitic from Proto-Semitic *s[1] > h when word-initial was blocked in some roots where other forms of the paradigm would have remained *s[1]. The verb samiʿa, "he heard", should have become everywhere hamiʿa but, likely due to the imperfect forms, which would have been word-internal and thus not shifted (yismaʿ not **yihmaʿ), the initial s[1] was retained (Al-Jallad 2015).

Finally, in some instances an etymological *-hu was possibly lengthened by analogical pressure from the long vowel in the 3fs suffix, *-hā. Elsewhere in the pronominal suffix paradigm feminine forms form the basis for analogical adjustment to the masculine forms. In some cases, for example, the originally feminine -i vowel is levelled to masculine forms as well: e.g., Ğabal Fayfāʾ (Yemen) *-hum/*-hin > -him/-hin. Elsewhere, the 3fp are extended via clipping of the feminine plural imperfective suffix *-nah: e.g., Bani Abādil (Yemen) 3mp -him but 3fp -hinna (Behnstedt 1987, p. 67). Masculine forms with a doubled nasal and final a were created in many dialects based on the feminine forms: e.g., Aṣ-Ṣalt (Jordan) -hummu/-hinne ((Palva 1994, p. 461); apud (Procházka 2014, p. 142)). It is very possible then that the long feminine vowel led to the lengthening of the masculine form in some cases as well.

Other than -*u(h)*/-*o(h)*, the 3ms suffix also frequently takes the form -*i(h)*, -*eh*, or -*ah*. In a number of Najdi dialects, the vowel preceding all suffixes, singular and plural, is the high vowel *i*, with the exception of the 3fs suffix, which is *a* (Ingham 2008, p. 328). In, e.g., the southern Hijaz and Tihama, (Procházka 1988, p. 192) reports -*ah*. Elsewhere, a fronted -*eh* is reported, for example in the dialect of Yašīʿ in Yemen (Isaksson 1991, p. 126).

Dialects that attest a 3ms suffix realized as -*a(h)*, including, e.g., some Najdi dialects (Ingham 1982, p. 98) and the southern Ḥijāz (Procházka 1988, p. 126), as well as some Sudanese dialects (Owens 1984), have traditionally been interpreted as reflecting the generalization of the accusative case followed by the loss of final short u: *a-hu > *ah* (Cantineau 1939). Owens (2006, pp. 254–55) suggested instead that there were originally two 3ms suffix variants, one with a high vowel (usually *u*, *i* in Najdi), and one with a low vowel, the quality of which was specifically conditioned by the presence of emphatic consonants. As an example, he cites Eastern Libyan Arabic, where the distribution is supposedly attested. Owens then suggests that some dialects levelled the emphatic variant *ah*, while others the non-emphatic *ih*. As mentioned, Jastrow (1991, p. 170 et passim) argues that -*ah* and -*ih* reflect a frozen accusative and genitive vowel respectively.

There are reasons to doubt each of these explanations, however. To begin with, it is not clear whether dialects with -*eh* reflect either *-*i* or *-*a*. Relatedly, it is not always clear that contemporary *a* or *i* should be connected with historical *a or *i. This is especially true of, e.g., the Najdi dialects that constitute the classic examples of dialects with 3ms suffix -*ih* and -*ah*. In most Najdi dialects, historical high vowels *i and *u merged, with phonetic context determining whether the surface form was *i* or *u* (Ingham 2008, p. 327). It is also important to consider issues of transliteration and phonetic reality when discussing these forms. For example, as illustrated in Table 3 above, Sanʿāni Arabic is sometimes transliterated with -*a*, as with Rossi, and other times with -*i*, as with Watson. It is likely that presumed etymological correspondence with historical vowels influences which is used.

I believe the most obvious explanation is that the forms -*ah*, -*ih*, and -*əh* are ultimately the result of a process wherein the distinction between *i* and *a* before *h* word-finally is neutralized. In non-rounded contexts, though transliterated as *ah*, the feminine ending is actually realized as basically a schwa; e.g., Najdi *-*ah*, realized in neutral contexts as ə, i.e., *xirzəh* "bead", (Ingham 2008, p. 327). Ingham elsewhere (Ingham 1986, p. 283) notes that 3ms suffix transliterated -*ih*, "his", is often realized similarly to the *tāʾ marbūṭa*, namely as -*əh*. Such an explanation would also make sense out of the difference variation in transliteration found in a number of dialects, such as Ṣanʿāni, where Rossi (1939, p. 4 gives -*ah*, but Watson (2009, p. 110) gives -*ih*. There is also evidence for such a situation in the Negev (Palestine) dialect of the ʿAzazmih, where the reflex of the 3ms suffix is -*ih* when suffixed to most nouns, but -*ah* when suffixed to verbs whose stem vowel is low: *rās-ih* "his head", but *ǧāb-ah* "he brought it" (Shawarba 2012, p. 193). This is paralleled in some cases by the realization of the feminine nominal marker *-*ah*, which is -*it* in construct on most nouns but -*at* on verbs whose stem vowel is low: *nāg-it* "she-camel of", but *nām-at* "she slept." The same is true for Eastern Libyan itself, where Owens notes that Mitchell's descriptions (Mitchell 1952, 1960, apud Owens 1984, p. 93) give 3ms suffix-*ih*, whereas Owens transliterates it -*ah*. Owens' examples from Eastern Libyan Arabic are probably similar to other dialects, where the exact quality depends on the phonetic environment. Other dialects have been recorded with a form transliterated as -*eh* (e.g., Yašīʿ; (Isaksson 1991, p. 126)).

If correct, the traditional historical explanation for these forms, namely that they represent frozen genitive or accusative cases, becomes much less convincing. These variants cannot be explained through the same set of developments as the harmonizing group. Note that the variation attested in the vowel before *tanwīn* (Stokes 2020, pp. 654–58) mirrors that of the 3ms suffix. I suggest that the same neutralization attested before *tanwīn* occurred before PN suffixes in some dialects as well:

bayt-V^{case}-hu*/bayt-V^{case}n* > *bayt-V-hu/*bayt-Vn

Following the loss of final short vowels, the 3ms suffix would follow this non-phonemic vowel:

bēt-ih/bēt-in.

As we will see, many dialects with 3ms *-ih* or *-ah* attest 2ms *-ik* and 2fs *-ić*. Thus, it is likely that vowels in this position merged, with a high vowel surface manifestation in every instance.[11] It is unclear what might have caused such a neutralization in this position. One possibility is that the change began in the construct, where the case vowel on the initial noun of the construct chain would be in a waṣl position, and thus potentially prone to reduction:

*baytu r-raǧuli "the man's house" > **bayt Vr-raǧuli*

Another possibility is that post-stress but non-word final vowels were reduced and merged:

*bayt-u-hu/bayt-un > **bayt-V-hu/*bayt-Vn* but *al-bayt-u* and *baytu r-raǧuli*

Subsequently, the non-phonemic vowel manifested as a high front vowel in most contexts, and final short vowels were lost. The construct perhaps continued inflecting for case, but the otherwise-ubiquitous absolute form eventually replaced it.

This can be summarized by the following steps:

1. Post-stress, non-word final neutralization of vowel contrast: **bayt-ən*; *bayt-ə-hu*; but *bayt-V^{case}*
2. Generalization of high vowel realization: **bayt-in*; **bayt-i-hu*; but **bayt-V^{case}*
3. Loss of final short vowels: *bayt-in*; *bayt-ih*; *bayt*

In a few dialects, such as Al-Mahābšeh, however, 3ms is *-eh* but 2ms is *-ak*, suggesting against a generalization to all contexts. In these cases, there might have been a generalization of *-a* before pronominal suffixes, with subsequent raising of *-ah*, as with the *tāʾ marbūṭa -ah*, to *-eh* or *-ih*.

The 3fs suffix is in most dialects realized as *-a* or *–(V)ha*. The latter suggests an originally long final vowel **-hā*, but the former is ambiguous. Ahmad Al-Jallad (p.c.) has pointed out that there is some pre-Islamic Arabic epigraphic evidence for a short **-ha* by-form in line one of the Namārah inscription, which reads *mlk ʾl-ʿrb kl-h/*malk ʾal-ʿarab koll-ah/"king of all ʿArab" (M.C.E Macdonald's translation, from Fiema et al. (2016, pp. 405–6)). For either /kull-ha/ or /kull-hā/ we would expect ***kl-hʾ* in the Nabataean orthography. Cantineau (1936/1937, pp. 78, 182–83) argued based on data from the Syrian Šammarī and ʿAnaze Bedouin dialects, in which 3fs suffix was realized *C-ah* but post-vocalic *V-ha*, that the vowel was originally anceps (that is, both long and short), and that its length matched the length of the preceding case vowel. While this fits the data on which Cantineau was focusing, this does not fit many other dialects, which show only one form. So, while I agree with Cantineau that there were possibly two by-forms, one long and one short, of the 3fs suffix, based on realizations of the suffix in other Arabic dialects (see below), as well as ClAr data, it seems unlikely that we can posit a length harmonization for the ancestors of all dialects, and almost certainly not for Proto-Arabic.[12]

If we posit short and long by-forms for Proto-Arabic, then examples of 3fs suffix realized as *-ha* seem most likely explained as retentions of **-hā*. More complicated are forms such as *-ah*. If, as is maintained here, final short **-a* vowels were eventually lost in the ancestors of the dialects, then *-ah* forms are most likely explained as reflecting the same harmony across morpheme boundaries described above: **V^{case}-ha* > *a-ha* > *-ah*, with harmonization before loss of final short *a*. Alternatively, *-ah* forms could reflect an analogical restructuring of the 3fs suffix based on the 3ms forms: *-uh/-aha* > *-uh/-ah*. In that case, even these forms could ultimately have descended from an originally long final form **-hā*. From such a situation, whether due to loss of a final short **a*, or analogical restructuring, we can explain the distribution in a minority of dialects in which 3fs is *-eh/-ih*, as in the dialect of Ǧiblah, Yemen (Table 4; data from Isaksson 1991, p. 130):

Table 4. Ğiblah Forms.

3ms	*C-uh/C-h*
3fs	*-ih*
2ms	*C-ak/V-k*
2fs	*-ik*
1cs	*-i*

The short *i* vowel of the 3fs suffix in this dialect could reflect one of two paths. One possibility is that it reflects a fronting of short a before word-final -*h* (as with *tāʾ marbūṭa* and 3ms suffixes noted above). Such a distribution is also attested in, e.g., Al-Ḫugarīyeh: 3ms C-*oh*/V-*uh* and 3fs C-*eh*/V-*ih* (Isaksson 1991, p. 132). Alternatively, the 3fs suffix in Ğiblah could be analogical with the 2fs suffix.

Forms like Levantine -*a*, which lack a final laryngeal, could possibly go back to either *-ha* or *-hā*. If they reflect an originally short vowel, then their development would mirror the -*ah* forms discussed above, with elision of the final laryngeal. More likely, in my view, is that they descend from long *-hā* and reflect the loss of the laryngeal -*h* following a vowelless consonant: *bayt-V-hā > bayt-hā > bayt-ā > bayt-a* "her house". Several pieces of evidence support such a reconstruction. First, the 3fs suffix regularly patterns in terms of pre-suffixal vowel quality and insertion with suffixes that historically were heavy, namely the plural forms, in multiple dialects. That is, these suffixes are typically preceded by a vowel when the phonotactic patterns of the particular dialect dictate the insertion of an epenthetic vowel.

In Meccan Arabic, for example, epenthetic vowels are regularly *a*, which in the paradigm above is inserted before 3fs, 1cp, and 3rd and 2nd plural forms. In the dialect of Al-Maḥall (among others), the anaptyctic vowel harmonizes with the following vowel (see Table 5). The simplest explanation of this distribution is that the vowels before these suffixes originated as anaptyctic insertions, not case vowels, and are thus distinct from the vowels preceding the 3ms, 2ms/2fs, and 1cs suffixes.

Table 5. Patterns of *-CVV and *CVC suffixes (Arabian data from Isaksson (1991); Damascene from Brustad and Zuniga (2019, p. 411)).

	Mecca	**Al-Maḥall**	**Damascene**
3fs	*-[a]ha*	*-[a]ha*	*-a*
3mp/2mp	*-[a]hum/-[a]kum*	*-[o]hum/-[o]kum*	*-on/-kon*
3fp/2fp	N/A	*-[a]han/-[o]kun*	N/A
1cp	*-[a]na*	*-[a]na*	*-na*

How can this apparent contradiction be resolved? First, the 3fs and 1cp suffixes, unlike 3ms, 2ms, and 2fs, are re-constructible with long *ā* vowels, 3fs *hā* and 1cp *nā*, which pattern with the heavy CVC syllables of the 3mp/3fp *-hum/*-hin* and 2mp/2fp *-kum/*-kin* suffixes. I argue that as case inflection broke down, the original short (formerly case-inflecting) vowel was syncopated before long (CVV) and heavy (CVC) syllables[13]:

*bayt-V-hā/hum/kum > *bayt-hā/hum/kum.*

At some point later in their development, a disallowance for CCVV and CCVC syllables arose in many dialects, resulting in insertion of an anaptyctic vowel in the same slot. Elsewhere, in, e.g., the ancestors of Levantine-like dialects, no such disallowance developed, with either the retention of the laryngeal of the 3rd person forms (as in Ḏafār; see Table 3 above) or loss (as in Damascene).

The current proposal can thus be summed up by distinguishing between suffixes that consisted historically of light syllables (i.e., CV) and those that consisted of heavy ones (either CVV or CVC). Before suffixes consisting of light syllables, the final short vowel,

historically a case vowel, in some cases (e.g., Levantine) harmonized with the vowel of the suffixes, and in other cases (e.g., Najdi) lost contrast and manifested as a generalized high vowel.[14] Before suffixes consisting of heavy syllables, there was an initial syncope of the short vowel.[15] In many dialects, a subsequent phonetic rule disallowing strings of heavy syllables resulted in the insertion of an anaptyctic vowel, while in others no such insertion occurred. In dialects such as Damascene, the laryngeal was elided, but elsewhere no elision occurred (see Table 6 for reconstructions).

Table 6. Reconstruction of Meccan and Damascene.

	Meccan-Type	**Damascene-Type**
3fs/3mp	**bayt-V-hā/*bayt-V-hum >* **bayt-hā/*bayt-hum >* *bayt-a-ha/bayt-a-hum*	**bayt-V-hā/bayt-V-hun >* **bayt-hā/*bayt-hun > bayt-a/bayt-on*
2mp	*bayt-V-kum > *bayt-kum > bayt-a-kum*	**bayt-V-kun > *bayt-kun > bēt-kon*

3.2. Second Person Masculine and Feminine Forms

The 2ms suffix is in many dialects realized *-ak* after consonants and *-k* after vowels. These are derivable from **V^{case}-ka > *a-ka* (via the vowel harmony rule mentioned above) *> ak* (after loss of final short *a*), with a post-vocalic variant *-k*. A significant minority of dialects attest another vowel before the final velar stop, usually either high front *-ik*, but in a few instances high back *-ok/-uk*. In the case of 2ms *-ik*, the feminine form is almost always an affricated *-itš* or *-ić*. I have already argued that 3ms *-ih* in many cases reflects a neutralization of vowel contrast in this position. The dialects with 3ms *-ih* in which 2ms and 2fs suffixes are *-ik* and *-ić* suggest a general loss of contrast, with a high front vowel surface manifestation.

The 2ms suffix form realized as *-ok* or *-uk* has long baffled commentators (cf. Diem 1973, p. 42; Isaksson 1991, p. 127). In Ḏafār, for example, the paradigm is as follows (Table 7; data from Isaksson 1991, p. 127):

Table 7. Ḏafār singular forms.

3ms	*C-eh/V-h*
3fs	*-ha*
2ms	*-uk*
2fs	*-iš*

How can the 2ms *-uk* be explained? One possibility is that, initially, the pre-suffix vowel was universally high: **bayt-V^{case}-PN > *bayt-V^{high}-PN*. Subsequently, harmonization of the 2fs to the similarly high suffix vowel occurred: **V^{high}-ki > *i-ki > *i-ši > -iš*. Before the 2ms, however, the back high *u* was generalized, with no subsequent harmonization. It is curious that the 3ms and 2fs suffixes, both of which were followed etymologically by high vowels, apparently triggered a front high manifestation of the pre-suffix vowel, whereas the 2ms, with a low vowel, did not. Perhaps there was a pattern that led to this:

V-$CV^{high} > V^{front}$-CV^{high},

but

V-$CV^{low} > V^{back}$-CV^{low}

The feminine by-forms *-ik/-ki* require a bit of comment. Comparative evidence strongly favors a reconstruction of final short i, which we would not expect to remain in these dialects. Diem (1991, p. 301) reconstructs 2fs **-kī*. While this is a possible explanation of forms after long vowels with *-ki*, it cannot explain the nearly ubiquitous existence of *-ik* forms. Several scholars have suggested rather that *-ki* forms are due to contamination from the nominative pronoun inti and the imperfect verbal suffix, e.g., Jordanian *tuktubi* "you

(fsg) write" (<*tiktubī*) (attested also apparently in Gǝʿǝz; (Hasselbach 2004, p. 10, n. 28; Al-Jallad 2014, p. 319)). While it is possible that 2fs patterned originally with 3ms and 3fs in having short and long by-forms, with some length assimilation to explain why *-ki* remains only after long vowels, I believe the lack of 2ms long forms argues in favor of analogy with the independent and imperfective forms.

If indeed the source of the *i* on 2fs-*ki* PN suffix forms is based on analogy with the imperfect, we must still explain its predominant distribution: suffixed to nouns that end in a long vowel but absent following nouns that end in a consonant. In, e.g., Damascene Arabic, the 2fs suffix is *-ik* after a C but *-ki* after a vowel: *kitāb-ik* "your (fs) book", but *abu-ki* "your (fs) father." In other words, if the final vowel of -ki was restored from the imperfect, why was it not generalized to all contexts? Its distribution is precisely the same distribution of the 3fs suffix, which is *-a* after a C but *-ha* after a vowel: *kitāb-a* "her book", but *abū-ha* "her father." There is, in other words, a gendered difference in the paradigm. I suggest that, in many dialects, this distinction led to the creation (or retention) of a 2fs suffix with final *-i*, which was available still on the independent and prefix verbal forms:

3ms -VV/2ms -VVK:: 3fs -*ha*: 2fs X > *ki*

The symmetry also holds for other dialects. For example, in Najdi, the 2fs suffix patterns completely with the 3ms and 2ms:

3ms C-*eh*, V-*h*/2ms C-*ik*, V-*k*/2fs C-*ić*, V-*ć*

Alternately, in a small minority of dialects, the 2fs suffix patterns completely with the 3fs, against 2ms and 2fs, e.g., the Saudi dialect of Al-Maḥābšeh (Table 8).

Table 8. The 3rd and 2nd singular pronominal suffixes.

	After Consonant	After Vowel
3ms	C-*eh*	V-*h*
2ms	C-*ak*	V-*k*
3fs	C-*ha*	V-*ha*
2fs	C-*ki*	V-*ki*

We can therefore explain the 2ms and 2fs suffix forms via developments from originally short final vowels. As we have seen with the 3ms and 3fs (and the plural forms, obliquely), the vowels preceding the suffixes were originally case inflecting short vowels, which lost phonemic value and subsequently harmonized with the retained suffix vowels.

4. Suffix Vowel Length

The length of the final vowels in Proto-Semitic is notoriously difficult to determine because the reflexes of the forms differ across the attested Semitic languages.[16] Various scholars have attempted different solutions to this puzzle. Brockelmann (1908/1913, p. 74 ff.) suggested what has been the majority opinion of the past century, namely that the vowels were anceps, meaning that they could be long or short. Brockelmann's own intuition was that they were originally long, but that some were shortened because they were unstressed (ibid.).[17] Blau (1981, p. 63 ff.) held the opposite scenario, namely that the suffix vowels were originally short, but that paradigm pressure, namely the preservation of gender distinctions on suffixes, led speakers to retain them in languages like Hebrew, where short vowels were otherwise lost.

Despite the difficulties in reconstructing their length in Proto-Semitic, I follow Hasselbach (2004); Al-Jallad (2014) in analyzing most of the forms as originally short in Proto-Arabic. The reasons, discussed above, concern the consistency with which the dialects agree with ClAr. For example, virtually all modern dialects lack the final vowel on 2ms *-ka*, which seems best interpreted as the loss of final short *a*. It is necessary, and indeed methodologically preferable, to allow for multiple rounds of final-vowel reduction and loss when modeling the development of the dialects from their ancestors. We should

not constrain ourselves to solving the loss of, say, the word-final *-a* of the 2ms suffix in the same stage as the loss of final *-a* on the 3ms prefix verb; rather, we should probably factor in multiple rounds of reduction and loss, depending on dialect. The main difference between my reconstructions and those offered by Hasselbach concern the 3fs suffix, which I reconstruct long for Proto-Arabic, based on the form in ClAr, as well as its almost total retention in the modern dialects. Exceptions to these forms are, I believe, explicable by appeal to analogies (on which, see above). However, the Syrian Bedouin data gathered and analyzed by Cantineau suggest the possibility that, in the ancestors of some dialects, and perhaps even Proto-Arabic, a short by-form for 3fs suffix, and a long by-form for the 3ms suffix, based on length assimilation and/or polarization, was created. Table 9 presents proposed reconstructions of Proto-Arabic PN suffixes.

Table 9. Reconstructed Proto-Arabic Pronominal Suffixes.

3ms	*-hu*, *-hū* (?)	3mp	*-hum(ū)*[18]
3fs	*-ha* (?), *-hā*	3fp	*-hin*
2ms	*-ka*	2mp	*-kum(ū)*
2fs	*-ki*	2fp	*-kin*
1cs	*C-ī*/*V-ya*	1cp	*-nā*

5. Conclusions

In this paper I have addressed two long-standing issues in the historical study of the modern Arabic dialects. First, I argued in favor of the traditional interpretation of the vowels before the pronominal suffixes of the 3ms, 2ms, and 2fs as originating in etymological case vowels. I provided two different models to account for the modern realizations in the two major attested paradigms. Specifically, I posited two sets of rules. To account for dialects in which pre-suffixal vowels harmonized with the vowels of the pronominal suffixes, I proposed a vowel harmony rule which operated across morpheme boundaries. In order to explain forms in which a single vowel occurs prior to each suffix, I proposed a development by which post-stress vowels were neutralized, leading to a generic vowel (usually front high) surface manifestation. The remaining variants were explained within these models. I contrasted these scenarios with those offered by Blau and Diem, detailing how the arguments here account for variables unaccounted for in these previous proposals.

The second topic examined here is that of the historical length of several pronominal suffixes; specifically, the length of the 3ms, 3fs, 2ms, and 2fs suffixes was examined. I concluded that most 3ms forms can be derived from a historical short *-hu*, although a long by-form possibly existed at the Proto-Arabic node. I further argued that the 2ms and 2fs suffixes are derivable from historically short forms, *-ka* and *-ki*, and that exceptions can be explained via analogies that are readily identifiable. Finally, I argued that the 3fs suffix was in the vast majority of cases derivable from a historically long *-hā*. Furthermore, I argued that the vowels before the 3fs suffixes patterned with those of the 3rd and 2nd plural forms over against the other singular forms, which I argued is due to the fact that the etymological case vowels were syncopated at some point before suffixes of the syllable shape CVV and CVC, with the subsequent insertion of an epenthetic vowel. These vowels behave differently from those of the singular forms because they represent a separate stage of development in the dialects.

Conflicts of Interest: The author declares no conflict of interest.

Notes

[1] Most who posit a nominal case system for Proto-Arabic reconstruct it essentially as it is found in ClAr. For our purposes, the important point is that originally a pronominal suffix would be suffixed to a noun which ended in a case vowel: N + V^{case} + CV(C). So, for example: **kalb-u-hu* "his dog (nom)"/**kalb-a-hu* "his dog (acc)"/**kalb-i-hu* "his dog (gen)."

[2] Diem (ibid.) argues that the various realizations of DT, i.e., *un*, *in*, and *an*, reflect frozen forms of the three primary Arabic cases. For an alternative proposal, see Stokes (2020).

[3] That is, other than his distinction between 'nomadic' and non-nomadic dialects, for which I can find no good evidence in the realizations of the pronominal suffixes. It seems his distinction is instead based here entirely on the DT. However, DT is also present in sedentary dialects.

[4] Blau's main objection to Diem's analysis, with which he interacts directly, is that redundancy is very common in linguistic systems and cannot therefore be considered a very strong factor on its own (Blau 2006, p. 86). Rather, Blau posits that pausal forms, where *tanwīn* and case vowels were regularly lost, began to intrude into non-pausal positions as well, i.e., he argues for an analogical extension of pausal Ø-marked forms into non-pausal position. Additionally, however, Blau seems to also hold that some sort of final vowel elision took place, perhaps as a result of this pausal intrusion, but it is not clear to me how he envisions the relationships between these two processes.

[5] It is puzzling that Owens recognizes implicitly in his reconstructions the loss of final short vowels, since this is the typical explanation for the loss of case vowels, which he finds unconvincing (ibid., pp. 2–8 *et passim*). Since he allows for the loss of these final vowels, which is implicit in his suggested proto-forms and necessary to account for the data, it is not clear why that loss could not account for the absence of case in the dialects. Thus, for Owens, final vowel loss seems a very random, haphazard process.

[6] This point also argues against some kind of metathesis, i.e., *qalb-hu > qalb-uh*. First, we would expect that such a metathesis would be motivated by the same impetus as epenthesis, namely phonotactic considerations. My argument in this section is that the epenthetic argument does not account for the distribution of these vowels in any dialect. A further argument against the metathesis argument is the fact that the metathesis would have to have occurred only with pronominal suffixes. That is, if metathesis occurred in the same phonetic environment (i.e., -CCV), we would expect the same to affect suffix conjugation suffixes, such as *darasta* "you (msg) wrote" and *darastu* "I wrote." That these forms do not have dialectal forms ***darasat* or ***darasut*, but rather just *darasit* (< *darast* <* *darata/darastu*) "I/you wrote," suggests that this did not occur there.

[7] This development would have presumably affected nouns in construct when the following noun lacked the definite article: **ibnu malikin* "a king's son," > *ibna malikin* "idem." While the topic is slightly outside the scope of this paper, it seems likely that changes in both the definite and indefinite construct syntagms played a role in the reduction of case inflection.

[8] It is already attested in the pre-Islamic period in, e.g., the corpus of Graeco-Arbica, transliterations of Arabic into the Greek script from the city of Petra in modern-day Jordan (Al-Jallad 2017, pp. 145–46). The same is true in many modern dialects. For Levantine varieties, see Brustad and Zuniga (2019, pp. 407–8).

[9] In his chapter on pausal forms, called *bāb al-waqf* (Sibawayh 1988, p. 173 ff.), Sibawayh discusses these forms and the phenomenon to which Blau refers, but without using the term *naql*. Sibawayh typically reserves the term *naql* for the transfer of the vowel on a *hamza* to the preceding consonant with the subsequent loss of the *hama*: *al-ʾḥmar* "the red one," *>laḥmar* "idem." I thank an anonymous reviewer for pointing this difference out to me.

[10] Perhaps the ClAr reverse length polarization in the 3ms suffixes, wherein *V-hū* but *VV-hu*, would represent an innovation prodded by the existence of these by-forms.

[11] Alternatively, we might posit a generalization of a single case, perhaps nominative, before pronominal suffixes. After the merger of **u* and **i*, the high vowel would be generalized.

[12] In varieties described by the grammarians, which subsequently became standardized for ClAr pronunciation, a reverse length polarity developed for the 3ms suffix (*V-hū*, *VV-hu*), whereas 3fs was ubiquitously long. The dialects cannot, of course, have descended from ClAr since ClAr represents an amalgam of varieties and features. However, the reality of the reverse polarity elsewhere suggests that by-forms existed (or, to use Cantineau's term, should be reconstructed as anceps), but were distributed differently across dialects. In some, a length distinction became allomorphic; elsewhere, one form was generalized to all contexts.

[13] This phenomenon is attested and remarked upon by the Quranic reading commentator Al-Farrāʾ in his *Luġāt al-Qurʾān* (Al-Farrāʾ 2014, p. 30). He notes specifically that the case vowels *-u* (nom) and *-i* (gen) are syncopated whenever pronominal suffix with a heavy syllable follows: *bayt-kum* (nom/gen) but *bayta-kum* (Acc). Interestingly, traces of this same phenomenon can be seen in the Quranic reading traditions of Abū ʿAmr and Ibn Kaṯīr. For example, in Q2:54 the reciters of Abū ʿAmr read *ʾilā bāriʾikum* "to your (pl) creator" as *ʾilā bāriʾ-kum*, without the genitive *i* due to the following heavy syllable. I thank an anonymous reviewer for bringing these examples to my attention.

[14]　　Again, it is unclear whether in these dialects either the nominative or genitive case was levelled to all contexts before pronominal suffixes, or rather there was a general loss of phonemicity as case inflection broke down, with the resulting high vowel a result of the general preference in these dialects for open, unstressed syllables to contain high vowels.

[15]　　I thank an anonymous reviewer for pointing out that this is similar to a change from Pre-Classical to Classical Latin, in which a light syllable is deleted between two heavy syllable: e.g., *lāridī* > *lardī* "bacon". Armin Mester (1994) calls this process "prosodic trapping," in which the light syllable is not metrically footed.

[16]　　For example, the 2ms suffix reflects *-ka* in Arabic and Gəʿəz, but is present in some dialects of Hebrew, in which short final *a* should have been lost. Further, 2fs reflects *-ki* in Hebrew and Arabic, but in Gəʿəz, where *u* and *i* merged to ə, the attested form is -ki < *kī.

[17]　　This position was also suggested to me by J. Huehnergard (p.c.), who notes the cross-linguistic tendency for long vowels to be realized short when unstressed (as in, e.g., Latin).

[18]　　I have not encountered solid evidence for the retention of the so-called long pronominal forms among the modern dialects. Sibawayh discusses the plural forms in (*bābu mātuksaru fī-hi l-hāʾu llatī hiya ʿalāmatu l-ʾiḍmār*) (Sibawayh 1988, p. 195). For an in-depth discussion of these long forms, see Van Putten and Sidky (Forthcoming).

References

Al-Farrāʾ, ʾAbū ʿAlī. 2014. *Kitāb Fīh Luġāt Al-Qurʾān*. Edited by Ǧābir b. ʿAbd Allāh al-Sarīʿ. Unpublished.

Al-Jallad, Ahmad. 2014. Final short vowels in Geʿez, Hebrew ʾatta, and the Anceps Paradox. *JSS* 59: 315–27.

Al-Jallad, Ahmad. 2015. Yusapʿil or Yuhapʿil, that is the question—Two solutions to sound change s¹ > h in West Semitic. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 165: 27–39.

Al-Jallad, Ahmad. 2017. Graeco-Arabica I: The Southern Levant. In *Arabic in Context: Celebrating 400 Years of Arabic at Leiden University*. Edited by Ahmad Al-Jallad. Leiden: Brill, pp. 99–186.

Al-Jallad, Ahmad. 2018. The earliest stages of Arabic and its linguistic classification. In *The Routledge Handbook of Arabic Linguistics*. Edited by Elabbas Benmamoun and Reem Bassiouney. New York: Routledge, pp. 315–31.

Al-Wer, Enam. 2007. Jordanian Arabic (Amman). In *Encyclopedia of Arabic Language and Linguistics, Vol. II*. Edited by Kees Versteegh. Leiden: Brill, pp. 505–17.

Behnstedt, Peter. 1987. *Die Dialekte der Gegend von Ṣaʿdah (Nord-Jemen)*. Wiesbaden: O. Harrassowitz.

Birkeland, Harris. 1952. *Growth and Structure of the Egyptian Arabic Dialect*. Oslo: I kommisjon Hos Jacob Dybwad.

Blau, Joshua. 1972. On the problem of the synthetic character of Classical Arabic as against Judaeo-Arabic (Middle Arabic). *Jewish Quarterly Review* 63: 29–38. [CrossRef]

Blau, Joshua. 1981. *The Emergence and Linguistic Background of Judaeo-Arabic: A Study of the Origins of Middle Arabic*, 2nd ed. Jerusalem: Ben-Zvi Institute.

Blau, Joshua. 2002. *A Handbook of Early Middle Arabic*. Jerusalem: Hebrew University.

Blau, Joshua. 2006. Some reflections on the disappearance of cases in Arabic. In *Loquentes Linguis: Linguistic and Oriental Studies in Honour of Fabrizio A. Pennacchietti*. Edited by Pier Gorgio Borbone, Alessandro Mengozzi and Mauro Tosco. Wiesbaden: Verlag Harrassowitz, pp. 79–90.

Brockelmann, Carl. 1908/1913. *Grundriss der Vergleichenden Grammatik der Semitischen Sprachen*. Reprint 1961. Hildesheim: Georg Olms Verlagsbuchhandlung, vol. II.

Brustad, Kristen, and Emilie Zuniga. 2019. Levantine Arabic. In *The Semitic Languages*, 2nd ed. Edited by John Huehnergard and Na'ama Pat-El. New York: Routledge, pp. 403–22.

Butts, Aaron Michael. 2019. Gəʿəz (Classical Ethiopic). In *The Semitic Languages*, 2nd ed. Edited by John Huehnergard and Na'ama Pat-El. New York: Routledge, pp. 117–44.

Cantineau, Jean. 1936/1937. Études sur quelques parlers de nomades arabes d'Orient. In *Annales de l'Institut d'Études Orientales*. Paris: Librairie Larose.

Cantineau, Jean. 1939. Le pronom suffixe de 3ᵉ personne singulier masculin en arabe Classique et dans les parlers arabes modernes. *BSL XL* 1: 89–97.

Diem, Werner. 1973. *Skizzen Jemenitischer Dialekte*. Beiruter Texte und Studien, 13. Beirut and Baden-Baden: Nomos Verlagsgesellschaft.

Diem, Werner. 1991. Vom Alterabischen zum Neuarabischen Ein neuer Ansatz. In *Semitic Studies: In Honor of Wolf Leslau on the Occasion of His Eighty-Fifth Birthday, 14 November 1991*. Edited by Wolf Leslau, Alan Kaye and Thomas Leiper Kane. Wiesbaden: Harrassowitz, vol. 1, pp. 297–308.

Fiema, Zbigniew T., Ahmad Al-Jallad, Michael Christopher Macdonald, and Laïla Nehmé. 2016. Provincia Arabia: Nabataea, the Emergence of Arabic as a Written Language, and Graeco-Arabica. In *Arabs and Empires before Islam*. Edited by Greg Fisher. Oxford: OUP, pp. 373–433.

Fischer, Wolfdietrich. 2002. *A Grammar of Classical Arabic*, 3rd rev. ed. Translated from the German by Jonathan Rodgers. New Haven: Yale University Press.

Fischer, Wolfdietrich, and Otto Jastrow, eds. 1980. *Handbuch der Arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Grigore, George. 2007. *L'arabe parlé à Mardin: Monographie d'un Parler arabe Periphérique*. Piscataway: Gorgias Press.

Grigore, George, and Gabriel Biṭuna. 2012. Common Features of North Mesopotamian Arabic Dialects Spoken in Turkey (Sirnak, Mardin, Siirt). In *Bilim Düşünce ve Sanatta Cizre: (Uluslararası Bilim Düşünce ve Sanatta Cizre Sempozyumu Bildirileri)*. Edited by Mehmet Nesim Doru. Mardin: Mardin Artuklu Üniversitesi Yayınları, pp. 545–54.

Hasselbach, Rebecca. 2004. Final Vowels of Pronominal Suffixes and Independent Personal Pronouns in Semitic. *JSS* 49: 1–20. [CrossRef]

Hasselbach, Rebecca. 2014. Agreement and the Development of Gender in Semitic (Part 1+2). *ZDMG* 164: 33–64.

Huehnergard, John. 2019. Proto-Semitic. In *The Semitic Languages*, 2nd ed. Edited by John Huehnergard. New York: Routledge, pp. 49–79.

Ingham, Bruce. 1982. Notes on the Dialect of the Ḍafīr of North-Eastern Arabia. *Bulletin of the School of Oriental and African Studies* 45: 245–59. [CrossRef]

Ingham, Bruce. 1986. Notes on the Āl-Murra of eastern and southern Arabia. *Bulletin of the School of Oriental and African Studies* 49: 271–91. [CrossRef]

Ingham, Bruce. 2008. Najdi Arabic. In *The Encyclopedia of Arabic Language and Linguistics, Vol. III*. Edited by Kees Versteegh. Leiden: Brill, pp. 326–34.

Isaksson, Bo. 1991. The personal markers in modern Arabic dialects of the Arabian peninsula. *Orientalia Suecana* 40: 117–45.

Jastrow, Otto. 1991. 'Une question embarrassante'—Jean Cantineau über das Pronominalsuffix 3. Sg. m. in den arabischen Dialekten. In *Festgabe für Hans-Rudolf Singer: Zum 65. Geburtstag am 6. April 1990, Überreicht von Seinen Freunden und Kollegen*. Edited by Martin Forstner. Paris: Peter Lang, pp. 167–74.

Lentin, Jérôme. 2006. Damascus Arabic. In *The Encyclopedia of Arabic Langauge and Linguistics, Vol. I*. Edited by Kees Versteegh. Leiden: Brill, pp. 546–55.

Mester, R. Armin. 1994. The Quantitative Trochee in Latin. *Natural Language & Linguistic Theory* 12: 1–61.

Mitchell, Terence F. 1952. The Active Participle in an Arabic Dialect of Cyrenaica. *Bulletin of the School of Oriental and African Studies* 14: 11–33. [CrossRef]

Mitchell, Terence F. 1960. Syllabification and Prominence in Arabic. *Bulletin of the School of Oriental and African Studies* 23: 369–89.

Owens, Jonathan. 1984. *A Short Reference Grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Owens, Jonathan. 2006. *A Linguistic History of Arabic*. Oxford: OUP.

Owens, Jonathan. 2018. Where multiple pathways lead: A reply to Ahmad Al-Jallad and Marijn van Putten. *Abhandlungen für die Kunde des Morgenländes* 113: 95–162.

Palva, Heikki. 1994. Bedouin and sedentary elements in the dialect of es-Salṭ. In *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*. Edited by Dominique Caubet and Martine Vanhove. Paris: INALCO, pp. 459–69.

Procházka, Theodor. 1988. *Saudi Arabian Dialects*. London: Kegan Paul International.

Procházka, Stephen. 2014. Feminine and Masculine Plural Pronouns in Modern Arabic Dialects. In *From Tur Abdin to Hadramawt: Semitic Studies Festschrift in Honour of Bo Isaksson on the Occasion of His Retirement*. Edited by Tal Davidovich, Ablahad Lahdo and Torkel Lindquisit. Wiesbaden: O. Harrassowitz, pp. 129–48.

Retsö, Jan. 1994. ʾiʿrāb in the Forebears of Modern Arabic Dialects. In *Actes des Premières Jounrées Internationales de Dialectologie Arabe de Paris*. Edited by Dominique Caubet and Martine Vanhove. Paris: INCALCO, pp. 333–42.

Rossi, Ettore. 1939. *L'arabo Parlato a Ṣanʿāʾ*. Rome: Istituto per l'Oriente.

Shawarba, Musa. 2012. *A Grammar of Negev Arabic: Comparative Studies, Texts and Glossary in the Bedouin Dialect of the 'Azazmih Tribe*. Wiesband: Harrassowitz.

Sibawayh, ʾAbū Bišr ʿUṯmān. 1988. *Kitāb Sībawayh, Volume IV*. Edited by ʿAbd al-Salām Muḥammad Hārūn. Cair: Maktabat al-Ḫāniǧī.

Stokes, Phillip W. 2020. A Fresh Analysis of the Origin and Diachronic Development of 'Dialectal *Tanwīn*' in Arabic. *JAOS* 140: 637–64. [CrossRef]

Suchard, Benjamin. 2020. *The Development of the Biblical Hebrew Vowels*. Leiden: Brill.

Talay, Shabo. 2011. Arabic dialects of Mesopotamia. In *The Semitic Languages: An international handbook*. Edited by Stefan Weninger. Berlin and Boston: De Gruyter Mouton, pp. 909–20.

Van Putten, Marijn, and Phillip W. Stokes. 2018. Case in the Qurʿānic Consonantal Text. *WZKM* 108: 143–79.

Van Putten, Marijn, and Hythem Sidky. Forthcoming. Pronominal Variation in Arabic among the Grammarians, Quranic Reading Traditions, and Manuscripts. In *Formal Models in the History of Arabic Grammatical and Linguistic Tradition*. special issue of *Language and History* 65: 1. Edited by Raoul Villano.

Watson, Janet C. E. 2007. Syllabification Patterns in Arabic Dialects: Long Segments and Mora Sharing. *Phonology* 24: 335–56. [CrossRef]

Watson, Janet C. E. 2009. Ṣanʿānī Arabic. In *The Encyclopedia of Arabic Language and Linguistics, Vol. IV*. Edited by Kees Versteegh. Leiden: Brill, pp. 106–14.

Woidich, Manfred. 2006. Cairo Arabic. In *The Encyclopedia of Arabic Language and Linguistics, Vol. I*. Edited by Kees Versteegh. Leiden: Brill, pp. 323–33.