7. Three caveats

Three structural problems can be noted in the interpretation of the statistics.

7.1 Innovations

The statistics are a representation of contemporary Arabic. Ideally one would want to have similar statistics from varieties of Arabic in 1800, 1500, 1200, and so on. Obviously, lack of sources for spoken Arabic precludes such detailed accounting. In general it may be assumed that Arabic dialects are becoming more dissimilar with time and that the present statistics include many innovations. An obvious case in point are variables 3 and 4. These describe the reflex of *h and *ʔ, which in WSA is mostly h and ? (see Ap 9.1.3, 4). That these are innovations in WSA is shown by the contrast in Bagirmi Arabic of fa ’har “month” vs. le ḥem “meat”. In these forms *h and *h have fallen together in WSA h. However, the indication of a former pharyngeal h is left behind in the different vowel quality. Low a remained low after *h, but was raised to e after *h. Ancestral Bagirmi Arabic must have had *h at some point.

The statistics, however, by their very logic, will in this respect show WSA to contrast with most other varieties of contemporary Arabic, which of course have retained *h. This point underscores the necessity of using the statistics as a tool of historical interpretation only in the context of a plausible attempt at reconstruction. If, as suggested, WSA and Uzbekistan Arabic have a special historical affinity, the data cited here would allow one to remove one statistic separating the two contemporary reflexes, namely presence (Uzbekistan) vs. absence (WSA) of pharyngeals. Moreover, looked at in these terms the statistics help indicate not only what innovations have occurred in a given contemporary dialect area, but also the degree to which innovations have spread. The WSA area has a mean of (1.86) for this variable, which indicates that it is close to complete in the region (2 = uniform h/?).

7.2 Minority forms

A temptation should be avoided of valorizing the statistics to the detriment of other, relevant material and variant forms. The statistics will always be incomplete. I illustrate this in two separate places, so I will be brief here. In 4.3 it was seen that the form of the active participle + suffix pronoun provides an interesting historical link between Uzbekistan Arabic and one variety of WSA. This has not been included as a variable, however, since in many dialect descriptions no information is given on the form. It would be a major mistake to ignore it, however.

In 8.5 below are mentioned forms which are attested sporadically in my WSA data, always in texts, but which, so far as my experience in explicit elicitation of data goes, does not occur regularly. Unfortunately, even the best-researched dialect areas have gaps. Precisely in regard to the question of the spread of a given form, and its status as to retention or innovation, such minority forms may have an important interpretive role to play. The sporadic presence of a final imala form in verb, ma fi' “he went”, suggests a link to Tripolitanian m fi', the Daaxila oasis in Egypt (B & W: 281) and to various Gulf dialects (Prochazka 1981: 37).

7.3 What features are criterial?

A final caveat is not to lose sight of the fact that the statistics as I have developed them here are non-weighted. Each feature is of equal status as far as numerical calculations go.
It is above all this aspect of statistics which perhaps will make philologically-orientated
Arabicists wary. One could add a feature, for instance, for the realization of *s. This is
one of the most stable of all phonemes in Arabic (discounting assimilation for voicing,
emphasis and other regular allophonic effects). Were it added, it would have the same
numerical weight as the realization of the 1SG perfect verb suffix as –tu or –t, a variable
of greater significance than is *s? Why, however, is the 1SG suffix more significant?
Dialectologists would simply intuitively regard it as such. Even statistically one could
work out a justification for according it greater significance.

Each value of a variable could be given a variation index, for instance, ranging between 0
and 100. The index is formed by dividing the number of cases by the number of different
realizations of a variable and multiplying by the percent of each variety in the sample. So
far as the present data goes, were *s entered as a variable it would have had a value of
100, indicating that *s is realized as /s/ in all varieties. The 1SG perfect suffix, on the
other hand, has a value of –tu = 33%, -t = 67% (9.2.40). For purposes of dialect
comparison it is the more heterogeneous variables which are the more interesting. In
carrying out comparisons one could manipulate the variation index in different ways, e.g.
by excluding variables whose variability is above 90% (i.e. not very variable).

8. Problems in coding

In this section I would like to discuss a number of problems which, in a sense, have no
solution, but which, at the same time, in all probability do not materially affect the main
results of the present analysis. I add the proviso “in all probability” because, since they
are problems, their significance cannot be accurately gauged until they have been solved,
or at least, worked on to a greater extent than what I have done. Furthermore, there are
problems whose solution is fairly obvious, but which, for lack of more detailed data, are
impossible to implement at the present time.

8.1 Coding values

The statistical summary relies on a numerical coding practice whose basis is, to a broad
degree, arbitrary. It is hoped that this very arbitrariness, however, guards against skewing
the results in one direction or another.

Statistics distinguishes between ordinal, interval and nominal coding scales. Interval are
those which measure the extent to which a case has a certain fixed property; ordinal
define the relative order a set of cases have of a certain property. Nominal are those in
which cases are defined to have a certain property in an arbitrary, though explicit way.
The present statistics, like much found in language phenomena, are of a nominal order:
they measure no fixed property, and the assignment of a number to one exponent or
another is arbitrary. I have, however, followed certain guidelines in assigning values to
the different cases. “1” is a form close to or identical with Standard Arabic. Thus,
between the two 2FSG object pronoun forms –ki and –ik, I have coded –ki with “1”, -ik
with “2”. I should emphasize that in according Standard Arabic the value “1” I am not
making a statement of historical origin. This is simply an arbitrary starting point, which
provides an orientation for any dialect. Note furthermore that I use Standard Arabic and
not a presumed Classical Arabic as the hinge. Classical Arabic as seen above is an
interpretive problem. Thus I codify the form CaCVV, as in kabiir with “1”, as this is the
same as Standard Arabic. It is also broadly identical to Classical Arabic. However, if
taking Sibawaih as a source for the language, Classical Arabic also has the adjective form
fi fiyl e.g. si fiyd “happy” (Kitaab II: 274), so that, barring an extensive justification of
what one understands by Classical Arabic, many variables in this data base would be ambiguous in this respect. Standard Arabic as an artificial, conventionalized model (certainly broadly based on Classical Arabic) is in this context a convenient point of initial measure.

A second informal principle is that what I perceive as more widespread forms are codified with a lower number. My horizon extends here beyond the present data set. Thus, for the 2MSG object pronoun –ak is tagged as “1”, -ik as “2”. Of course in this case the Standard Arabic form –ka is of no help.

In a few cases I have coded variables to allow for split categories, which are coded as a fraction of a number (see below). In CC-C sequence in nouns there are three outcomes in the current data, as illustrated in (1). In (1a and b) an epenthetic vowel is inserted between either the first and second or second and third consonants.

(1a) CaC-C, akl-na → akl-ña “our food”
(1b) CCa-C, akl-ña
(1c) CC-C aklña

Some dialects (e.g. JB) have only (1a), some (WSA in most cases) have only (1b), and one (Uzbekistan) appears to have (1c). In addition some dialects have (1a) in some contexts, (1c) in another. Mardin for instance has (1a) when C₂ is a sonorant, as in the example, otherwise (1c), as in kalb-ki “your.F dog”. I code these “both/and” dialects with a half point. In order for this to work, however, (1a) and (1c) need to be numerically contiguous, so in this instance, (1a) has the value “1”, (1c) the value “2” and (1b) the value “3”. In this reckoning Mardin gets a score of “1.5” (see Ap 9.1.23).

I would note that as more dialects are added to the data bank, and “both/and” dialects fill in gaps, this system would probably break down (e.g. if there emerged a dialect with both (1a) and (1b), the value between would be “2”, which is already occupied by (1c)). For present purposes, however, no contradictions arise.

Beyond these three points I use my own intuitions to one extent or another. While different codings would result in different individual values, in most cases they would not alter my final analysis. For instance, take the variable t, emphatic t. “1” is the Standard pronunciation t (note arguably different from Sibawaih’s description of t as majhuura), “2” the variant q(t)(implosive emphatic), found in various WSA dialects, while “3” is t, a non-emphatic reflex found again in the WSA area, and in Uzbekistan. Why I have coded q as closer to Standard Arabic than t is: (1) q is still emphatic and (2) it maintains a systematic contrast to t in its dialects, whereas the de-emphasized form results in a phoneme merger. Still, it is relevant to see what would happen if the coding for “2” and “3” here were reversed. The difference is only relevant for the status of Uzbekistan Arabic. All of the Mesopotamian dialects in the sample have t, and most of the WSA have q, though some have t or t. The following Table compares the results of two coding values.

<table>
<thead>
<tr>
<th></th>
<th>t = “1”</th>
<th>q = “2”</th>
<th>t = “3”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesopotamia</td>
<td>9 x 1 = 9/9 = 1</td>
<td>Mesopotamia 9 x 1 = 9/9 = 1</td>
<td></td>
</tr>
<tr>
<td>WSA</td>
<td>8 x 2 + 1 x 3 = 19/9 = 2.1</td>
<td>WSA 1 x 2 + 8 x 3 = 25/9 = 2.7</td>
<td></td>
</tr>
<tr>
<td>Uzbekistan</td>
<td>3</td>
<td>Uzbekistan = 2</td>
<td></td>
</tr>
</tbody>
</table>
The end effect is the same, when Uzbekistan is compared to Mesopotamia and the WSA area. In both cases it is closer to WSA. The effect is greater for the first manner of coding than for the second, though other than to assign Uzbekistan Arabic to Mesopotamian or to WSA Arabic, I do not consider “degree of difference” as a relevant variable.

8.2 Hierarchicalized variables

Linguistic variables, particularly phonological and morphological ones, may often be hierarchicalized. “Emphasis” is a more general category than its instantiations in individual consonants, t̪, d̪ and so on. There is an implicational relationship such that if a dialect lacks emphasis, it lacks t̪, d̪ and all other emphatic consonants. The reverse does not hold. A dialect may lack a given emphatic consonant, say emphatic d̪, but still may have other emphatic ones (e.g. s and/or d̪).

I think it preferable to oppose dialects according to their most general parameter of contrast. Abbeche Arabic, for instance, should be opposed to Nigerian Arabic (all dialects), on the basis of non-emphatic vs. emphatic. This produces one feature of contrast. Another basis of comparison would be to oppose each consonant which realizes such as opposition. On this basis, Nigerian Arabic, with phonemic d̪, d̪, s, r,  l, m would differ in six ways from Abbeche Arabic (d, t, s, r, l, m).

In practice, I have followed neither accounting practice completely. The clear parameter of difference described in the previous paragraph is muddied through the following circumstance. Consonants in the emphatic series may differ not only in terms of emphasis, but also in terms of other phonetic contrasts. To continue the example from the beginning of this section, the alveolar plosive not only contrasts in terms of emphasis, but also in terms of buccal pressure, d̪ vs. t̪. I judge this latter to be an important parameter characterizing many WSA dialects. However, once the d̪ vs. t̪ contrast is included, so to must t (non-emphatic variant), since a number of dialects have this reflex for SA t̪.

The practice which I have followed here is that a difference in emphasis alone is not included as a parameter of contrast. The difference between s (basal “onion, Nigeria” and s (basal, Abbeche) for example is not included, since, so far as the present data goes, there are no further phonetic parameters of difference implied by the s - s contrast. As soon as another difference is implied in the contrast, however, individual emphatic phonemes are included. The Abbeche Arabic reflex t for SA t is therefore included, since SA t also differs in terms of the explosive/implosive contrast described above.

Note that the same principle applies to morphological differences. In many dialects the 2 and 3PL. object pronoun forms have –n rather than –m, -kun, -hum “your.PL, their.PL (e.g many qalantu dialects), vs. –kum, -hum. To the extent that –hum always implies –kun, or the other way around, there would be no grounds for recognizing two separate variables here. Indeed, there are no dialects in this sample with –m in one of the forms and –n in the other. However, there are dialects with –ku in the 2PL, so a three-set –kum, -kun, -ku series has to be coded. This is, in the present analysis, different from the two-way –hum, - hun contrast, so the 2/3PL. forms are entered as separate variables.

It would, perhaps, be possible to organize the analysis in a more subtle way, based on the following observation. A 2PL form –kun always implies –hum, whereas either -kum or ku always imply –hum (there are no –ku/hun dialects). In this respect my classification errs
on the side of caution. While I know of no dialects with –ku/hun, it cannot in principle be ruled out that somewhere such a contrast isn’t found. Indeed, as more and more detail about Arabic dialects become known, experience has shown that what can occur, will occur, somewhere.

Of course, other considerations are implied here, which I do not address explicitly. For instance, I would tend to recognize distinctions more in the realm of inflectional morphology than in phonology. To return to the s/s difference noted above, I think it unlikely that a dialect should turn up where only this distinction occurs, but not in other emphatic sounds.

There are, however, issues within this broad thematic area which I leave unresolved. One, for instance, pertains to the class of second and third person plural forms. Some dialects have morphologically distinct masculine and feminine forms, while others do not. Among those that do, the exponent of the FPL may be variable. In Bagirmi Arabic, for instance, the 3FPL object suffix is –han, whereas in Kirenawa, as in most areas of Chad and Nigeria, it is –hin. I have not incorporated this vocalic difference into the comparison, since the variation cannot be applied in a broad way to the Mesopotamian area: few Mesopotamian dialects have a morphological FPL (only those in the extreme south). This case is analogous to the emphatic/non-emphatic contrast discussed above. In this instance, however, I recognize only the parameter, morphological FPL, yes or no. This means that any differences within the exponent of FPL forms are not considered. In this one instance the effect is to slightly increase the homogeneity of the WSA area vs. the Mesopotamian area. However, one would have to compare variation within a category in area A with a non-existent category in area B.\(^1\) The issue is noted as a problem.

8.3 One rule or two, one case or two

A similar problem relates to the status of certain phenomena as linked or independent, in linguistic terms, one rule or two. A case in point is the status of a short vowel in pre-stress position in (reconstructed) open syllables. In many dialects one needs but a single rule to account for all occurrences, or more accurately non-occurrences: short vowels do not occur in the context *CV.CV. The quality of the vowels is irrelevant:

(2) JB, Hiit etc.: kbiiy/kbiir, kleeb < *kabiir, kilaab “big, dogs”

In fewer dialects the vowel may be maintained in both contexts.

(3) Kirenawa: kabiir, kilaab

So far as these two sets of data go, and indeed the rules generalize to cover a variety such as JB in its entirety, a single rule accounts for the occurrence or non-occurrence of all short vowels in open, pre-stress syllables. On the basis of these cases the present statistics need only a single variable with two values: short vowel maintained or short vowel deleted (or not allowed).

Further data, however, indicates that the two cases do not always go hand in hand. In particular, there are a good number of dialects allowing a short low vowel, or reflex thereof, in an open pre-stressed syllable, but not a short high vowel.

\(^1\) Lest it be thought that I am minimizing the variation within WSA by ignoring this feature, a comparable category would be the suffixal copular pronouns (Jastrow 1978: 131 ff.), which are unique to various qatlu dialects (this could be captured by a “yes/no” feature), and which within this group are internally variable. I have not included this variable.
(4) MB: *cibiir, cleeb* “dogs”  
Khaweetna: *cabir, cleeb*

In MB the short low vowel is raised to *i* in an open syllable (a very widespread phenomenon), but not deleted, while in Khaweetna it is maintained as *a*. Short high vowels, however, are categorically deleted in pre-stress position in both.

On the basis of this further data, two variables need to be recognized for vowels in pre-stressed open syllables, one for short high, and one for short low vowels. Moreover, even this differentiation is not adequate. In the Anatolian qultu dialects a short low vowel will often be maintained, either as *a* or as *ɔ* in nominal forms, but in derived verb forms the *a* is lost. In Siirt (Jastrow 1978: 62) one has, for instance, *kaleem-i* “my word”, with *a* maintained in the pre-stress open syllable but in the same dialect *tqaatal* “quarrel” < *tqaatal* with the vowel of the derivational prefix *ta-* lost. Variable 17 is therefore phrased to cover only a short *a* in the pattern *CaCiiC*. Separate variables apparently are needed to cover the reflexes of short *a* in different morphological forms.

In general in cases of doubt I over-differentiate or over-specify variables. There is, for instance, an important variable pertaining to medial **CCVhC-C'V** sequences. This occurs, inter alia, in the basic imperfect verb when a vowel suffix is added. Prototypical variants are *yiktub-u* (no vowel rearrangement) vs. *yikitbu*.² Potentially comparable nominal forms are found with the feminine singulative suffix –*it* (zib’dit-u vs. zibitt-u ‘his butter’), but the interpretation of the variants here is complicated by two factors. On the one hand in Mesopotamian Arabic in particular the maintenance or deletion of the short high vowel interacts with stress (which may protect the short high vowel). On the other hand, the situation in other dialects (Nigerian Arabic in general) is influenced by the nature of the consonants involved, which makes a simple rule formulation difficult. At this stage in a broad investigation, I limit the alternation to verbs only, where the situation is fairly clear.³

8.4 Variation: morphophonological and sociolinguistic

There are basically two types of variation in language, linguistically conditioned and sociolinguistically explicable. Both types are present in the current data. The linguistic variation is, traditionally, the better known and theoretically better accounted for. In basic linguistic parlance, it is known as allo-variation (allomorph/allophone). Examples from the present data include the following:

The preformative vowel of verb: Abbeche = /a/ before -CC , /i/ or /u/ before CV, *yu-mrug* “he leaves”, *yu-murgu* “they leave”; Shukriyya = /a/ before verb stems with /a/ as stem vowel, /i/ before verbs with high stem vowel (*yafrab* “he drinks” vs. *yuktub* “he writes”).

---

² This form is usually accounted for in two steps: the stem vowel is deleted in an open syllable, *yiktitb-u* → *yiktb-u* and the medial CCC sequence is broken up by the insertion of an epenthetic vowel between the first two consonants → *yikitbu*.

³ In any instances of doubt, I present data as morphophonological rather than phonological. Jastrow (1978: 88-92) for instance reports that in nouns **CC#** sequences will frequently be broken up by an epenthetic vowel in Anatolian qultu dialects, e.g. Siirt *mahxl* “cradle”, but in verbs the –*t* of 2MSG is not accompanied by epenthesis, *nɔmt* “you slept”. In respect of epenthesis in these cases a distinction has to be drawn, inter alia, between nouns and verbs.
The 2FSG object suffix. In a number of dialects, -ki after a long vowel, -ik otherwise (similarly 3MSG objects suffix, either –u/-hu or –u/-nu in same contexts). A slightly different variation concerns the 3MSG object suffix. Generally it has a basic allomorphic variation between –V (usually either –u or –a) and –CV (either –nu or –hu), the latter after a long vowel (as in the immediately preceding –ki/ik). What is distinctive here are the allomorphic variants –hu/nu. Within the dialects with the –u suffix I distinguish further as to whether they take –nu or -hu after VV-. These instances differ from the preceding in that there are no dialects where only –nu/hu occur. I have categorized these variants as having mid values. For instance, -ki has a value of “1”, -ik a value of “2”. If a dialect has a conditioned –ki/ik variation it has the score of “1.5”.

A more difficult problem to handle is the issue of sociolinguistic variation, in part because it relatively rarely is systematically investigated. Dialects, however, increasingly are coming into contact, often under the influence of regional koines, and undergoing change as a result of this contact. Pure dialects in the sense that only a minimal amount of non-linguistically conditioned variation occurs are becoming rarer. Such variation is often remarked upon obliquely but unsystematically in standard dialect descriptions. Talay (1999: 30) for instance, notes that Khawetna (qoldt) has many occurrences of /g/ in loanwords from other Arabic varieties. While he suggests that /g/ enters the dialect via loanwords, rather than via substitution native dialect words, he has both baq “steal” (native /q/) and bawwaag “thief” (with “Bedouin” /g/), indicating that /g/ is insinuating itself into “native” vocabulary as well. Remarkign on the same alternation, Khan (1997: 56) notes that in his recordings, the gilt /g/ is largely limited to a story about a bedouin. The alternation g ~ q is thus conditioned in complex ways: lexical conditioning plays an important role and situational and discourse factors are relevant. Ideally one would want a dialectology which systematically takes into account quantitative variation. In Owens (1998: chapter 5) I use dialect maps, for instance, which divide binary variables into percentiles, so that, for instance, the sampling points for rural Nigerian Arabic for the preformative vowel /a ~ i/ (see 9.2.51) are given as categorically a, categorically i or as falling somewhere in between. The contrasts, moreover, are based on statistically-distinctive differences. Lacking such quantitative data, however, one is left to make judgement calls about how to represent the linguistic variation. Based on the discussion in Khan and Talay, for instance, I consider both dialects to have both /q/ and /g/, marked with a score of “1.5”. On the other hand for a parallel variable which is spread from the

---

4 Other variables here include CC# nouns (viewed in synchronic, syllable-structure terms). In Kirenawa, for instance, a consonant cluster is allowed where C1 is sonorant (kalb “dog”), otherwise CVC#, lham “meat”. In CB only CaC is allowed, kalib, whereas in JB CC# is categorically allowed, lahm. A second is the stem vowel of CCVC-V imperfect verbs. In Kirenawa the V is elided where C2 is sonorant; otherwise it is kept, bifarb-o “they drink” (via elission + epenthesis), vs. buktubu “they write” (no elision). In contrast to Kirenawa, many dialects either categorically do, or do not elide.

5 In Arabic sociolinguistics lexical conditioning has been relatively well studied for the influence of Standard Arabic on spoken Arabic (e.g Abdul Jawad 1981, Holes 1987); it is equally relevant to the study of dialect contact, however.
dominant gilit dialect, it appears that whereas Khaweetna has both /k/ and /c/ (< *k), Hiit has only /k/.  

Ideally features should be coded with differentiated sociolinguistic and lexically marked features.

8.5. How many variants should be recognized?

I think there really is no obvious place to stop a survey of this kind. The temptation is to continue adding as much detail as possible. A broader survey will certainly account for a greater degree of variation than that represented here. Two aspects of this problem can be briefly cited. First, greater attention could be given to morphophonological and allophonic variation. For instance, the verbal 3FSG suffix is given as –at for WSA, yet there is a widespread alternative with weak middle verbs before V-initial suffixes, namely –t, jaab-at “she brought” vs. jaab-t-a “she brought it.M”. Such alternation specific to weak middle verbs is fairly rare among Arabic dialects, hence its broad distribution in the WSA is significant and worthy of classificatory note. At the same time, it is morphophonologically restricted, and for this reason I do not include it in this survey. A detailed look at the grammars I have used as the basis of this survey will reveal a number of comparable forms which I have not included. Such “second order” variables, as they can be termed, will await a more detailed treatment.

A second type of feature worthy of note here is that which in a given dialect is rare, yet potentially of significance in understanding the spread of certain forms on a pan-Arabic basis. To mention two examples here, very occasionally there are found in Nigerian Arabic weak-final 3MSG forms of the type ma\’se “he went”. I have never elicited them, though they do occasionally crop up in texts from the Bagirmi dialect area. They are also found in Tripolitanian Arabic, Daaxila and some Gulf dialects. (see 7.2 above). A similar case is the imperfect of stem II verbs, where occasionally ablaut stems like bi-diffin “cover with earth” are found. Such ablaut forms are the normal ones in Uzbekistan Arabic and are found in Upper Egypt as well (Behnstedt and Woidich: 226). These might be interpreted as remnant forms brought originally from Upper Egypt: these forms spread from Upper Egypt and became established in Tripoli, but not, ultimately, in the WSA area.

My point in citing this second class of occurrences which are not treated in these statistics is to remind one that it may be the unusual, rare form, but one which appears to link up with similar or identical forms located far afield that help to determine the historical status of a feature. Ablaut imperfect verb stems are rare, historically contingent

---

6 Though again another variable is at play here. Khaweetna is a dialect still in situ, constantly exposed to influences from surrounding Iraqi dialects, whereas Hiit was recorded among Jewish Arab emigrants residing in Beersheva, Israel.

7 In fact, the values of the 3FSG perfect suffix before a vocalic suffix is a small chapter in itself. In Eastern Libyan Arabic and Alawitan Hatay and Cilicia, for instance, it lengthens, ELA iktilb-aat-a (or iktilb-iet-ih) “she wrote it”, Hatay qatl-iiit-u “she hit him”, in southern Mesopotamia the –t doubles, kitb-att-a “she wrote it” (urban), in Cairene it attracts irregular stress, dara b-it-ak “she hit you”, etc. Ignoring purely phonologically-determined variants, Nigerian Arabic has three segmental values for it: -t after hollow verbs before a V-initial suffix, as in jaaf-t-ak “she saw you”, -it after weak final high-V verbs before a V-initial suffix, lig-it-a “she got it.M”, otherwise –at, lig-at-hum “she got them.M”.

8 Though cf. the verbal noun of second stem verbs, ti-diffin “covering with earth” which is common in the WSA area.
developments. Rather than interpret their presence in three widely separated areas as due to independent innovation, it is more likely they go back to common point of origin.

8.6 Structural, not lexical features in the index

The data used here is largely structural (phonological and morphological structure), not lexical. Thus when it is noted that Khaweetna (Talay 1999: 111) has CǝCy stems (🐈m-yit “she got blind”), the point of identity with, say Abbeche Arabic is in the existence of a common morphological structure. The fact that the lexical distribution of these patterns is not identical (Khaweetna laq-it vs. Abbeche ligiy-at “she got”) must be explained and described elsewhere.

Similarly, in regards to CVCC# nouns, both WSA and various qultu dialects have both CVCC# and CVCǝC# nouns, the presence or absence of the epenthetic vowel determined in each case by the value of the consonants. In my coding CVCC# has the value "1", CVCǝC# the value "2" and dialects which have both classes of nouns get the middle "1.5". Note, however, that the precise rules that sanction the CVCǝC# forms are different in the two areas. In the WSA area, one has CVCǝC, unless C₂ is a sonorant, in which case the CVCC# form occurs (darb "road" vs. digin "beard"). In the Daragözü qultu dialect (Jastrow 1973: 78-9, also in Mardin, Sasse 1971: 78) CVCǝC# is required if C₃ is a sonorant hǝjǝr "stone", vs. dǝhǝf "young donkey". Again, more detail is needed to distinguish the two, though the current classification does not contradict basic facts of the two cases.

8.7 Other points

Opposing forms within a variable are minimally contrastive. This is perhaps a type of hierarchical relation (8.2), though phonology and morphology are not logically hierarchizable relative to each other. The minimal point of contrast in the 2FSG object suffix is whether the vowel occurs before or after the consonant (-ık vs. -ki). In these terms no contrast is made between MB –ıc and CB –ık (post-C form), since the c ~ k contrast is a phonological one accounted for elsewhere (see 9.1.2). Should there be a dialect which has –ıc 2FSG object, but no other forms with c (I am not aware of such), the 2FSG object suffix in this case would be given a separate value, as the c in this case would not be accountable for under other rules.

The variants of variables are opposed in their surface realizations. It might be argued that Kirenawa buktubu (with maintenance of stem u) and Shukriyya bukutbu (with deletion and epenthesis, as in n. 2) have the same underlying form and therefore are not to be differentiated. However, at some point a differentiation can be made, either at an inspection of surface forms (as done here) or in terms of application of rules (Shukriyya applied rules which Kirenawa does not).

Unless otherwise specified, the suffixal elements in the data do not themselves have another suffix. The addition of further suffixes often induces allomorphy of different types, which is not accounted for here.

In most cases in the following summary /i/ represents a high vowel, either /i/ or /u/. In works on qultu dialects this vowel is usually represented as /o/. As is well known, the quality of the vowel will be affected by factors of emphasis, vowel harmony, gutturality

---

9 It appears that Khaweetna has largely regularized the low vs. high stems according to the transitive/intranstive parameter.
and lexical specification. Each of these elements is an independent variable in and of itself, none of which are included in the current data base.

Abbreviations

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>JB</td>
<td>Jewish Baghdad Arabic</td>
</tr>
<tr>
<td>CB</td>
<td>Christian Baghdad Arabic</td>
</tr>
<tr>
<td>MB</td>
<td>Muslim Baghdad</td>
</tr>
<tr>
<td>WSA</td>
<td>western Sudanic Arabic</td>
</tr>
</tbody>
</table>

Dialects used in survey (number represents the code)

Western Sudanic Arabic
1. Kirenawa
2. Mada
3. Aajiri
4. Wulaad Eeli
5. Amm Attiman
6. Umm Hajar
7. Abbeche
9. Atia I
9 Atia II

Mesopotamia (dialect sub-group as in Table 7 given in brackets)
10. Christian Baghdad (Baghdad qultu)
11. Jewish Baghdad (Baghdad qultu)
12. Muslim Baghdad (gilit)
13. Mardin (Anatolia)
14. Daragözü (Anatolia)
15. Siirt (Anatolia)
16. Khaweeeta (non-Baghdad qultu)
19. Hiit (non-Baghdad qultu)
19. southern Mesopotamia (gilit)
30 Uzbekistan
31 Shukriyya