Stem

Two broad perspectives underlie the research concerned with the question of what should be taken to be the set of basic and descriptively adequate notions needed for Arabic morphology. In a time-honored view, called here the root-based approach, words are formed from (consonantal) → roots, sequences of consonants identifying a common invariant among various related word forms. In the other, less widely explored view, the stem-based approach, words are formed from stems, i.e. forms that may consist of vowels as well as prosodic features such as vocalic or consonantal length. In this view, consonantal roots are considered to be a by-product or an emergent property of the organizational principles in the linguistic grammar.

This entry presents the stem-based view. Since linguistic morphology is primarily concerned with systems of relations between words, it should first be noted that, as in other languages with rich inflectional morphology, Arabic organizes words in paradigms. These can be described as sets of words built from combinations of stems with inflectional markers, the latter designating various morphosyntactic categories. As an example, consider a fragment of the Arabic verbal paradigm in Table 1. The Arabic verb is described as having two sets of forms or 'Tense/Aspect' categories, known as the imperfect and the perfect. The words in Table 1 illustrate the indicative mood of the imperfect aspect of the lexeme 'to write'. Fully inflected words are formed by placing /ktub/ in the context of the appropriate prefix—suffix pair. These pairs consist of {ya—u, ta—u, ta—u, ta—īna, ʾa—u, ya—ūna, ya—na, ta—ūna, ta—na, na—u, ya—āni, ta—āni, ta—āni, ta—āni}, the exponents of the morphosyntactic categories of Mood (Indicative), Person (1st, 2nd, 3rd), Number (singular, plural, dual), and Gender (masculine, feminine). In this article, the set of phonological forms created from the exponents of the morphosyntactic categories of a particular paradigm is referred to as the inflectional context of that paradigm, and the form /ktub/ is referred to as the verbal stem. To avoid ambiguity, the term 'stem' refers to

that phonological form of a lexeme to which an affix is attached. This sense of 'stem' is essentially the same as that assumed in modern lexeme-based theories of morphology such as those of Matthews (1972), Aronoff (1992), and Anderson (1992), and consequently it is not specific to Arabic or Arabic-like morphologies.

Table 1. Imperfect, Indicative of *kataba* 'to write'

		singular	plural	dual
3	masc.	*ya-ktub-u*	*ya-ktub-ūna*	*ya-ktub-āni*
	fem.	*ta-ktub-u*	*ya-ktub-na*	*ta-ktub-āni*
2	masc.	*ta-ktub-u*	*ta-ktub-ūna*	*ta-ktub-āni*
	fem.	*ta-ktub-īna*	*ta-ktub-na*	*ta-ktub-āni*
1		*ʾa-ktub-u*	*na-ktub-u*	

The vowel of the verbal stem must be lexically specified: [ya-ktub-u] 'he writes', [ya-lbas-u] 'he dresses', [ya-ḍrib-u] 'he hits'. There are also a few minimal pairs of stems distinguished solely on the basis of this vowel: [ya-smar-u] 'he is brown' vs. [ya-smur-u] 'he spends the night conversing', and [ya-ḥzun-u] 'he saddens' vs. [ya-ḥzan-u] 'he is sad'. Given these facts, some researchers have proposed that the verbal morphology of triliterals builds on the stem /ktub/ (Schramm 1962, 1991:1403; Kuryłowicz 1972:34, 43; McOmber 1995:179; Ratcliffe 1998:33; Benmamoun 1999:176, among others). For example, Kuryłowicz (1972:43) writes, "The fundamental form of the Sem. conjugation, the so-called 'imperfect(ive)' yaqtul(u), shows a characteristic vowel after (R2) which is unpredictable, i.e. independent of any grammatical rule, hence basic. Therefore the verbal root is not a consonantal skeleton (q-t-l), but contains an essential vocalic component (u of qtul)". Put in present terms, Kuryłowicz's view consists of the claim that the verbal morphology is stem based. This may be a plausible hypothesis, but it is not the generally accepted view in Arabic linguistics. For verbs, traditionally, morphology is assumed to operate on the consonantal root (see Cantineau 1950; Fleisch 1956; McCarthy 1979; Yip 1988; Hoberman

1988; Goldenberg 1994, among others; and Goldenberg 1994 and Hoberman 1995 for two recent reviews). As Schramm (1991:1402) writes, "The conventional statement of Semitic morphological typology for the last thousand years or so has always reflected the view that all verbs and most nouns are to be derived by a process of interdigitating discontinuous consonantal root morphemes, expressing lexical content, and vocalic pattern morphemes which express grammatical content". For nouns, in contrast, the stem-based hypothesis has made significant contributions to the understanding of the lawful relationships between noun forms. Some important studies on Arabic singular/plural morphology (→ number), in particular, have established that surface properties of the noun stem such as vocalic and consonantal length condition in crucial ways the form of the corresponding plural form (see Hammond 1988; McCarthy and Prince 1990, and references therein).

In recent work, however, the stem-based view for Arabic verbal morphology has witnessed more systematic development, for instance by McCarthy (1993), McOmber (1995), Ratcliffe (1998, Chap. 2), Benmamoun (1999), and Gafos (2003) for Classical Arabic. For notable examples of the stem-based view for modern Arabic dialects, see Cowell's (1962) grammar of Syrian Arabic and Heath's (1987) monograph on Moroccan Arabic. In what follows, some of the virtues of the stem-based approach to Classical Arabic verbal morphology are sketched, starting with the set of facts related to doubled verbs, also known as biconsonantal or geminated verbs, whose explanation has consistently relied on the root-based view. In the perfect, geminated verbs show two allomorphs, [madd] and [madad], as shown in Table 2 for the lexeme 'to stretch'. Henceforth, [madd] will be called the GEMINATE allomorph and [madad] the STRONG allomorph – strong due to its resembling the nonalternating, so-called strong verbs like [katab] 'to write'.

The distribution of the allomorphs is also shown on the left in Table 2. The geminate allomorph occurs before vowels, and the strong allomorph before consonants. Which one of these two allomorphs underlies the alternation? One answer to this question is suggested by the following observation. The shape of [madad]

Table 2. Perfect of *madda* 'to stretch'

		singular	plural	dual	allomorph distribution
3	masc.	madd-a	madd-ū	madd-ā	
	fem.	madd-at	madad-na	madd-atā	Geminate [madd] /__V
2	masc.	madad-ta	madad-tum	madad-tumā	Strong [madad] /__C
	fem.	madad-ti	madad-tunna	madad-tumā	
1		madad-tu	madad-nā		

is the shape of the nonalternating triliteral verbs like [katab-a] 'he wrote', [katab-tu] 'I wrote'. Conventionally, triliterals are assumed to be the 'canonical' verbs in Arabic, and by extrapolation rather than logical necessity, their shape is assumed to be the canonical shape for verbs. This assumption has been inherited in all discussions of doubled verbs known (Wright 1896:68–71; Cantineau 1946:133; Brame 1970:119; McCarthy 1979:265–267). Specifically, this assumption implies an analysis that consists in the following steps. The root /md/ first assumes the shape of a CvCvC sequence. Because the root consists of only two consonants, its final consonant /d/ extends to occupy two positions, hence /madad/. This intermediate form is then converted to [madd] before a vowel-initial suffix, via a process of syncope, as in /madad+V/ [maddV], and in some cases via a process of metathesis as in (the imperfect) /ya+mdud+V/ [yamuddV]. However, as McCarthy (1986:247–248) observes, this analysis treats the alternation as 'morpholexical' in character. There does not seem to be any reason why /madad/ should change to [madd] or why /ya-mdud-/ should change to [ya-mudd-]. If this alternation were phonological, it would falsely predict that /katab-/ should change to [katb-] and /ya-ktub/ should change to [ya-kutb-] before a vowel. It can thus be seen that the alternation, as formulated in the syncope/metathesis rule, is arbitrary in the sense that there is no phonological motivation for the particular form that this alternation takes.

However, there is an alternative (Gafos 2003): the underlying stem is /madd/, and [madad] is

a surface variant of /madd/. Surprisingly, this alternative has not been pursued. It is standard methodology in generative grammar that, given an alternation like [madd] ~ [madad], we consider at least the two hypotheses outlined above, and contemplate their consequences for the rest of the grammar. If /madd/ is the basic verbal stem, then suffixation with a vowel-initial suffix gives [madd-a], an attested form. Suffixation with a consonant-initial suffix, however, results in an illicit triconsonantal sequence, */madd-tu/. As in many other languages with geminates, Arabic bans geminates from syllable codas (for syllabification in Arabic, see Angoujard 1988; Broselow 1992; Itô 1986; Farwaneh 1995). The illicit consonant sequence is therefore split to satisfy syllabification, [madad-tu]. There is no need for morphological stipulation or intermediate, unmotivated steps. The alternation is driven by pure phonotactic canons or constraints that govern the admissible sequencing of phonemes in the language.

The geminated verb allomorphy is also found in the imperfect. As Table 3 shows, the conditioning of the two allomorphs is the same as that in the perfect. The geminate allomorph occurs before vowels, the strong elsewhere: [ya-mudd-u], [ya-mdud-na]. The forms in Table 3 illustrate the indicative mood of the imperfect aspect of the lexeme 'to stretch'. The other verbal moods built on the imperfect stem (subjunctive, jussive, imperative, and the rare → energicus) are in all relevant respects similar to the indicative. That is, prefixes are vowel-final and suffixes are vowel-initial, consonant-initial, or null. Moreover, the geminated verb alternation in these moods is identical to that found in the indicative (e.g. jussive 3rd pers. masc. sg. [ya-mdud], 3rd pers. fem. sg. [ta-mudd-a]).

If we assume that /mudd/ is the stem, the alternation follows the same pattern as in the perfect: in combination with a V-initial suffix,

Table 3. Imperfect, Indicative of *madda* 'to stretch'

		singular	plural	dual
3	masc.	*ya-mudd-u*	*ya-mudd-ūna*	*ya-mudd-āni*
	fem.	*ta-mudd-u*	*ya-mdud-na*	*ta-mudd-āni*
2	masc.	*ta-mudd-u*	*ta-mudd-ūna*	*ta-mudd-āni*
	fem.	*ta-mudd-īna*	*ta-mdud-na*	*ta-mudd-āni*
1		*ʾa-mudd-u*	*na-mudd-u*	

the stem surfaces as in [ya-mudd-u], but with a C-initial or null suffix, a geminate coda would result, *[ya-mudd(-na)]. The ban on coda geminates enforces alternation to [ya-mdud(-na)].

Up to now, discussion of verbal allomorphy has been confined to Form I of the Arabic verb. If, as argued, the allomorphy is due to phonological principles rather than morphologically conditioned idiosyncrasies of certain forms, then it is predicted that the alternation will be found whenever its phonological conditions are met. This prediction is confirmed. The alternation is also met in verbs of Form IX, XI, and QIV under conditions identical to Form I. Traditionally, Form IX is identified with the pattern *ktabab* (Wright 1896:43). As far as known, all subsequent work in the generative tradition has assumed that *ktabab* is the canonical Form IX of verbs. However, stems in Form IX surface as *ktabab* only before consonant-initial suffixes, for reasons familiar by now. Representative examples of the alternation are given under the verbal part of Table 4.

Form IX verbs like [i-ħmarr-a] 'he blushed' are related to adjectives of color and bodily defects, here [ʾaħmar-u] 'red' and its corresponding nominal form [ħumr-un] 'red.plural' (for an illustration of exactly this morphology

Table 4. Alternation in Form IX (Perfect)

Adjectives		Verbal alternation in Form IX (perfect)		
ʾa-ḫmar-u	'red'	*i-ḫmarr-a*	*i-ḫmarar-tu*	'he/I blushed'
ʾa-ṣfar-u	'yellow'	*i-ṣfarr-a*	*i-ṣfarar-tu*	'he/I became yellow'
ʾa-qbal-u	'cross-eyed'	*i-qball-a*	*i-qbalal-tu*	'he/I became cross-eyed'
cf. Verbs		Form I (perfect)		
/radd/	'to return'	*radd-a*	*radad-tu*	'he/I returned'

in a modern dialect, consider Form IX verbs from Syrian, e.g. [ʔaħmar] 'red' ~ [ħmarr] 'to blush', [ʔaṣfar] 'yellow' ~ [ṣfarr] 'to turn pale', and so on; Cowell 1962:101, 250). In the verbal form [i-ħmarr-a], putting aside the transparently epenthetic [i-] and the suffix [-a], the final consonant is the long version of its corresponding segment in the noun or the adjective. There are a few different ways to state the morphological link between the verb and its derivationally related forms. One such way is to derive the verbal stem by adding a suffixal mora (μ) to the simpler stem /ħmVr/, underlying the noun or the adjective: /ħmVr-/$^{\text{Stem}}$ + /ħmarr-/$^{\text{Verb-stem}}$. What is important for current purposes is that once the verbal stem is placed in its paradigm, it is clear that what is involved in the allomorphy [i-ħmarr-a] ~ [i-ħmarar-tu] is the by now familiar phonologically determined alternation.

The same alternation applies to quadriliteral verbs in Form QIV, [i-ʃmaʕall-a] 'he hastened', [i-ʃmaʃlal-tu] 'I hastened' and [i-ṭmaʔann-a] 'he was tranquil', [i-ṭmaʔnan-tu] 'I was tranquil'. These examples are all perfect forms. In the imperfect, the same alternation is found, e.g. [ya-ṭmaʔinn-u] and [ya-ṭmaʔnin-na] 'he/they [fem.] are tranquil' (Schramm 1962:362). The conditions for the alternation and the form that this alternation takes are identical throughout.

Another systematic property of the verbal system is that all verbal moods are based on the imperfect form CCvC, Kuryłowicz's 'fundamental form'. The indicative is shown in Table 1. The subjunctive differs from the indicative in superficial ways that do not affect the ensuing discussion. The jussive and the imperative are presented in Table 5 and Table 6, respectively. In the rare variant of the jussive, the → energicus (Schramm 1962:364), the affixes are in all relevant respects similar to the other moods, that is, all prefixes are vowel-initial and suffixes are vowel- or consonant-initial.

Table 5. Jussive of *kataba* 'to write'

		singular	plural	dual
3	masc.	*ya-ktub*	*ya-ktub-ū*	*ya-ktub-ā*
	fem.	*ta-ktub*	*ya-ktub-na*	*ta-ktub-ā*
2	masc.	*ta-ktub*	*ta-ktub-ū*	*ta-ktub-ā*
	fem.	*ta-ktub-ii*	*ta-ktub-na*	*ta-ktub-ā*
1		*ʾa-ktub*	*na-ktub*	

Table 6. Imperative of *kataba* 'to write'

		singular	plural	dual
2	masc.	*u-ktub*	*u-ktub-ū*	*u-ktub-ā*
2	fem.	*u-ktub-ī*	*u-ktub-na*	*u-ktub-ā*

In the remaining discussion, an important question is addressed that has not been addressed so far by the proponents of stem-based morphology. Observe that there is no contrast between [ya-CCvC-u] and [ya-CvCC-u]. Both [ya-CCvC-u] and [ya-CvCC-u] are phonologically well formed, but only the former is attested. Whence the [CCvC] invariance of the fundamental form?

This question turns out to have a simple answer when we take into account some independent properties of the language. First, Arabic does not allow complex syllable onsets or codas. Second, whereas all prefixes in Tables 1, 5, and 6 end in vowels, some suffixes begin with a consonant or are null. Thus, a /CvCC/ stem would raise a phonotactic problem before a consonant-initial or null suffix, since *[Cv-CvCC-Cv] is banned. A /CCvC/ stem presents no phonotactic problem because prefixes end in vowels. A [Cv-CCvC-Cv] is permissible because the first stem consonant can be parsed as a coda. Hence, we can begin to see how the inflectional context coupled with phonotactics requires that the CC cluster be at the left edge of the stem.

It is instructive to contrast this approach to a well-known alternative. To account for the lack of *[ya-CvCC-u] or, equivalently, the lack of contrast between [ya-CCvC-u] and *[ya-CvCC-u], the lexicon is restricted to include only /CCvC/. The lexicon is thus preconfigured so that [ya-CvCC-u] surface forms cannot arise, and this is done by imposing a restriction on the set of admissible grammar inputs or a 'morpheme structure constraint' (Chomsky and Halle 1968), as is commonly known. Conceptually, this approach is quite different from the one argued for presently, which seeks to derive the observed pattern as the lawful consequence of systemic factors, here, the inflectional context and phonotactics. Saying that the pattern is derived means that there is no unique locus in the grammar or the lexicon where the ban against [CvCC] forms or /CvCC/ stems is stated. Rather, it is the interaction of a few independ-

ent factors that effectively bans these forms (see Kisseberth 1970 on phonotactic 'conspiracies' and apparent constraints on inputs).

A more important reason in support of the proposed model derives from its predictive power. Observe that morpheme structure constraints do not make any predictions beyond their highly specific assertions, e.g., there is no /CvCC/ verbal stem in the Arabic lexicon. The model promoted here instead employs general principles in a theory of grammar, and consequently makes predictions beyond specific data. The stem-in-paradigm approach predicts that in a different paradigm with vowel-initial suffixes /CvCC/, stems would be possible. The example needed to test this prediction is provided by the morphology of the noun. As shown in Table 7, the inflectional context for nouns consists of vowel-initial suffixes ([*stem*-un] in the indefinite, [ʾal-*stem*-u] in the definite). It is thus expected that the /CvCC/ stem banned in the verb should now be possible in the noun. This is indeed the case as shown by a few representative forms from the well-populated class of triliteral nouns, [nafs-un] 'soul', [baḥr-un] 'sea', [qufl-un] 'lock', [burd-un] 'robe', and so on.

Table 7. Noun endings

	masc. sg.	masc. pl.	fem. sg.	fem. pl.
nominative	- *un*	- *ūna*	- *atun*	- *ātun*
genitive	- *in*	- *īna*	- *atin*	- *ātin*
accusative	- *an*	- *īna*	- *atan*	- *ātin*

To review, two related specific ideas are promoted here. The first is that we can make sense of the alternation between /madd/ and /madad/ if we assume that the verbal stem is /madd/. This is the basic form of the stem on which inflectional affixes are attached. The other surface form of a doubled verb, /madad/, results by splitting the geminate of /madd/ when that form is combined with a consonant-initial or null suffix. This happens because Arabic does not permit geminate codas as in *[madd-tu] or [madd]. The geminated verb alternation is not arbitrary, in the sense of being morphologically conditioned, and there is no need for rules that sometimes result in metathesis and sometimes deletion of vowels. The second related idea is that the inflectional structure of the paradigm

coupled with phonotactics provides a powerful source of constraints on the theoretically possible diversity of stem forms within that paradigm. The paradigm molds stems to fit the inflectional context of their realizations. This allows us to explain why certain stem shapes are found while others are not attested.

The purpose of this entry is to present the hypothesis that Arabic morphology, the system of lawful relations between words, must have access to more richly specified underlying representations than is allowed by consonantal roots. Specifically, core areas of the verbal morphology require reference to stems specified for properties such as vocalism and consonantal length, e.g. /ṣubb/ 'to pour [liquid]', /ḥabb/ 'to love passionately', /ṣal/ 'to arrive', which are not admissible as part of consonantal roots. It is at the stem level where generalizations about the morphology and phonology of the Arabic verb can be observed and stated in the form of a testable theory.

BIBLIOGRAPHICAL REFERENCES
Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
Angoujard, Jean-Pierre. 1988. "Gémination et redoublement". *Langues Orientales Anciennes Philologie et Linguistique* 1.1–15.
Aronoff, Mark. 1992. "Stems in Latin verbal morphology". *Morphology now*, ed. Mark Aronoff, 5–32. Albany: State University of New York Press.
Benmamoun, Elabbas. 1999. "Arabic morphology: The central role of the imperfective". *Lingua* 108.175–201.
Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.
Broselow, Ellen. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic Linguistics* IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 7–45. Amsterdam and Philadelphia: J. Benjamins.
Cantineau, Jean. 1946. "Esquisse d'une phonologie de l'arabe classique". *Bulletin de la Société Linguistique de Paris* 43.93–140.
——. 1950. "Racines et schèmes". *Mélanges William Marçais*, 119–124. Paris: Maisonneuve.
Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
Cowell, W. Mark. 1962. *A reference grammar of Syrian Arabic based on the dialect of Damascus*. Washington, D.C.: Georgetown University Press.
Eid, Mushira (ed.) 1995. *Perspectives on Arabic Linguistics*, VII. Amsterdam and Philadelphia: J. Benjamins.
Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure*. Ph.D. diss., University of Utah.
Fleisch, Henri. 1956. *L'arabe classique: Esquisse*

d'une structure linguistique. Beirut: Imprimerie Catholique.

Gafos, Adamantios. 2003. "Greenberg's Asymmetry in Arabic: A consequence of stems in paradigms". *Language* 79.317–355.

Goldenberg, Gideon. 1994. "Principles of Semitic word-structure". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 29–64. Wiesbaden: O. Harrassowitz.

Goldsmith, John. 1995. *Handbook of phonological theory*. Oxford: Blackwell.

Hammond, Michael. 1988. "Templatic transfer in Arabic broken plurals". *Natural Language and Linguistic Theory* 6.247–270.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. New York: State University of New York Press.

Hoberman, Robert. 1988. "Local and long-distance spreading in Semitic morphology". *Natural Language and Linguistic Theory* 6.541–549.

——. 1995. "Current issues in Semitic morphology". Goldsmith (1995:839–847).

Itô, Junko. 1986. *Syllable theory in prosodic phonology*. Ph.D. diss., University of Massachusetts, Amherst. (Publ., New York: Garland Publishers, 1989.)

Kaye, Alan S. 1991. *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday*. Wiesbaden: O. Harrassowitz.

Kisseberth, Charles. 1970. "On the functional unity of phonological rules". *Linguistic Inquiry* 1.291–306.

Kuryłowicz, Jerzy. 1972. *Studies in Semitic grammar and metrics*. London: Curzon Press.

Matthews, P.H. 1972. *Inflectional morphology*. Cambridge: Cambridge University Press.

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss, Massachusetts Institute of Technology. (Publ., New York: Garland Publishers, 1985.)

——. 1986. "OCP effects: Gemination and antigemination". *Linguistic Inquiry* 17.207–263.

——. 1993. "Template form in prosodic morphology". *Papers from the Third Annual Formal Linguistics Society of Midamerica Conference*, ed. L. Smith Stvan, 187–218. Bloomington, Ind.: IULC Publications.

—— and Alan Prince. 1990. "Foot and word in prosodic morphology: The Arabic broken plurals". *Natural Language and Linguistic Theory* 8.209–282.

McOmber, Michael L. 1995. "Morpheme edges and Arabic infixation". Eid (1995:173–189).

Ratcliffe, Robert R. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in Arabic and comparative Semitic*. Amsterdam and Philadelphia: J. Benjamins.

Schramm, Gene M. 1962. "An outline of classical Arabic verb structure". *Language* 38.360–375.

——. 1991. "Semitic morpheme structure typology". Kaye (1991:1402–8).

Wright, William. 1896. *A grammar of the Arabic language*. 3rd ed. Cambridge: Cambridge University Press.

Yip, Moira. 1988. "The Obligatory Contour Principle and phonological rules: A loss of identity". *Linguistic Inquiry* 19.65–100.

——. 1997. "Lexicon optimization in a language without alternations". *Current trends in phonology: Models and methods*, ed. Jacques Durand and Bernard Laks, 757–788. Salford, Manchester: European Studies Research Institute.

ADAMANTIOS I. GAFOS
(Department of Linguistics, New York University)