Consonant Spreading in Arabic Stems

Kenneth R. BEESLEY
Xerox Research Centre Europe
Grenoble Laboratory
6, chemin de Maupertuis
38240 MEYLAN
France
Ken.Beesley@xrce.xerox.com

Abstract

This paper examines the phenomenon of consonant spreading in Arabic stems. Each spreading involves a local surface copying of an underlying consonant, and, in certain phonological contexts, spreading alternates productively with consonant lengthening (or gemination). The morphophonemic triggers of spreading lie in the patterns or even in the roots themselves, and the combination of a spreading root and a spreading pattern causes a consonant to be copied multiple times. The interdigitation of Arabic stems and the realization of consonant spreading are formalized using finite-state morphotactics and variation rules, and this approach has been successfully implemented in a large-scale Arabic morphological analyzer which is available for testing on the Internet.

1 Introduction

Most formal analyses of Semitic languages, including Arabic, defend the reality of abstract, unpronounceable morphemes called ROOTS, consisting usually of three, but sometimes two or four, consonants called RADICALS. The classic examples include **ktb** (ك ت ب)[1], appearing in a number of words having to do with writing, books, schools, etc.; and **drs** (د ر س), appearing in words having to do with studying, learning, teaching, etc. Roots combine non-concatenatively with PATTERNS to form STEMS, a process known informally as INTERDIGITATION or INTERCALATION. We shall look first at Arabic stems in general before examining GEMINATION and SPREADING, related phenomena wherein a single underlying radical is realized multiple times in a surface string. Semitic morphology, including stem interdigitation and spreading, is adequately and elegantly formalizable using finite-state rules and operations.

da**r**a**s**	'study'	verb
du**r**i**s**	'be studied'	verb
da**rr**a**s**	'teach'	verb
du**r**uu**s**	'lessons'	noun
di**r**aa**s**a(t)	'study'	noun
da**rr**aa**s**	'eager student'	noun
ma**d**ra**s**a(t)	'school'	noun
ma**d**aa**r**i**s**	'schools'	noun
ma**drs**iyy	'scholastic'	adj-like
ta**dr**ii**s**	'instruction'	noun

Figure 1: Some stems built on root **drs**

1.1 Arabic Stems

The stems in Figure 1[2] share the **drs** root morpheme, and indeed they are traditionally organized under a **drs** heading in printed lexicons like the authoritative *Dictionary of Modern Written Arabic* of Hans Wehr (1979).

A root morpheme like **drs** interdigitates with a pattern morpheme, or, in some analyses, with a pattern and a separate vocalization morpheme, to form abstract stems. Because interdigitation involves pattern elements being inserted between the radicals of the root morpheme, Semitic stem formation is a classic example of non-concatenative morphotactics. Separating and identifying the component morphemes of words is of course the core task of morphological analysis for any language, and analyzing Semitic stems is a classic challenge

[1] The Arabic-script examples in this paper were produced using the ArabTeX package for TeX and LaTeX by Prof. Dr. Klaus Lagally of the University of Stuttgart.

[2] The taaʔ marbuuṭa, notated here as (t), is the feminine ending pronounced only in certain environments. Long consonants and long vowels are indicated here with gemination.

for any morphological analyzer.

1.2 Interdigitation as Intersection

Finite-state morphology is based on the claim that both morphotactics and phonological/orthographical variation rules, i.e. the relation of underlying forms to surface forms, can be formalized using finite-state automata (Kaplan and Kay, 1981; Karttunen, 1991; Kaplan and Kay, 1994). Although the most accessible computer implementations (Koskenniemi, 1983; Antworth, 1990; Karttunen, 1993) of finite-state morphotactics have been limited to building words via the concatenation of morphemes, the theory itself does not have this limitation. In Semitic morphotactics, root and pattern morphemes (and, according to one's theory, perhaps separate vocalization morphemes) are naturally formalized as regular languages, and stems are formed by the intersection, rather than the concatenation, of these regular languages. Such analyses have been laid out elsewhere (Kataja and Koskenniemi, 1988; Beesley, 1998a; Beesley, 1998b) and cannot be repeated here. For present purposes, it will suffice to view morphophonemic (underlying) stems as being formed from the intersection of a root and a pattern, where patterns contain vowels and **C** slots into which root radicals are, intuitively speaking, "plugged", as in the following Form I perfect active and passive verb examples.

```
Root:     d r s        k t b        q t l
Pattern:  CaCaC        CaCaC        CaCaC
          -----        -----        -----
Stem:     daras        katab        qatal


Root:     d r s        k t b        q t l
Pattern:  CuCiC        CuCiC        CuCiC
          -----        -----        -----
Stem:     duris        kutib        qutil
```

Prefixes and suffixes concatenate onto the stems in the usual way to form complete, but still morphophonemic, words; and finite-state variation rules are then applied to map the morphophonemic strings into strings of surface phonemes or orthographical characters. For an overview of this approach, see Karttunen, Kaplan and Zaenen (1992).

Following Harris (1941) and Hudson (1986), and unlike McCarthy (1981), we also allow the patterns to contain non-radical consonants as in the following perfect active Form VII, Form VIII and Form X examples.

```
              Form VII Form VIII Form X

Root:          k t b   k  t b      kt b
Pattern:     nCaCaC    CtaCaC    staCCaC
             ------    ------    -------
Stem:        nkatab    ktatab    staktab
```

In this formalization, noun patterns work exactly like verb patterns, as in the following examples:

```
Root:          k t  b     k t b      kt  b
Pattern:       CiCaaC     CuCuC     maCCuuC
               ------     -----     -------
Stem:          kitaab     kutub     maktuub

Gloss:         "book"     "books"   "letter"
```

Where such straightforward intersection of roots and patterns into stems would appear to break down is in cases of gemination and spreading, where a single root radical appears multiple times in a surface stem.

2 Arabic Consonant Gemination and Spreading

2.1 Gemination in Forms II and V

Some verb and noun stems exhibit a double realization (a copying) of an underlying radical, resulting in gemination[3] or spreading at the surface level. Looking at gemination first, it is best known from verb stems known in the European tradition as Forms II and V, where the middle radical is doubled. Kay's (1987) pattern notation uses a **G** symbol before the **C** slot that needs to be doubled.[4]

[3] Gemination in Arabic words can alternatively be analyzed as consonant lengthening, as in Harris (1941) and as implied by Holes (1995). This solution is very attractive if the goal is to generate fully-voweled orthographical surface strings of Arabic, but for the phonological examples in this paper we adopt the gemination representation as used by phonologists like McCarthy (1981).

[4] Kay's stem-building mechanism, using a multi-tape transducer implemented in Prolog, sees **G** on the pattern tape and writes a copy of the middle radical on the stem tape without consuming it. Then the following **C** does the same but consumes the radical symbol in the usual way. Kay's analysis in fact abstracts out the vocaliza-

```
Root:       k t b       d r s
Pattern:    CaGCaC      CaGCaC
            ------      ------
Stem:       kattab      darras
```

In the same spirit, but with a different mechanism, our Form II and Form V patterns contain an **X** symbol that appears after the consonant slot to be copied.

```
Root:       k t b       d r s
Pattern:    CaCXaC      CaCXaC
            ------      ------
Stem:       katXab      darXas
```

As in all cases, the stem is formed by straightforward intersection, resulting in abstract stems like **darXas**. The **X** symbol is subsequently realized via finite-state variation rules as a copy of the preceding consonant in a phonological grammar (**/darras/**) or, in an orthographical system such as ours, as an optionally written shadda diacritic (دَرَّس). Finite-state rules to effect such limited local copying are trivially written.[5]

2.2 Gemination/Spreading in Form IX

Spreading, which appears to involve consonant copying over intervening phonemes, is not so different from gemination; and indeed it is common in "spreading" verb stems for the spreading to alternate productively with gemination. The best known example of Arabic consonant spreading is the verbal stem known as Form IX (the same behavior is also seen in Form XI, Form XIV, Form QIV and in several noun forms). A typical example is the root **dhm** (د ه م), which in Form IX has the meaning "become black".

Spreading is not terribly common in Modern Standard Arabic, but it occurs in enough verb and noun forms to deserve, in our opinion, full treatment. In our lexicon of about 4930 roots, 20 have Form IX possibilities (see Figure 2). Most of them (but not all) share the general meaning of being or becoming a certain color.

byḍ	ب ي ض	'become white'
ħmr	ح م ر	'turn red' 'blush'
ħwl	ح و ل	'be cross-eyed' 'squint'
xḍr	خ ض ر	'become green'
xḍl	خ ض ل	'be moist'
dhm	د ه م	'become black'
rbd	ر ب د	'become ashen' 'glower'
rfḍ	ر ف ض	'drip' 'scatter' 'break up'
zrq	ز ر ق	'be blue in color'
zwr	ز و ر	'alienate'
smr	س م ر	'become brown'
swd	س و د	'become black'
ʃqr	ش ق ر	'be of fair complexion'
ʃmṭ	ش م ط	'turn gray'
ṣfr	ص ف ر	'turn yellow/pale'
ṣhb	ص ه ب	'become reddish'
ʕwj	ع و ج	'be crooked' 'be bent'
ɣbr	غ ب ر	'be dust-colored'
qtm	ق ت م	'be dark-colored'
kmd	ك م د	'become smutty/dark'

Figure 2: Roots that combine with Form IX patterns

McCarthy (1981) and others (Kay, 1987; Kiraz, 1994; Bird and Blackburn, 1991) postulate an underlying Form IX stem for **dhm** that looks like **dhamam**, with a spreading of the final **m** radical; other writers like Beeston (1968) list the stem as **dhamm**, with a geminated or lengthened final radical. In fact, both forms do occur in full surface words as shown in Figure 3, and the difference is productively and straightforwardly phonological. For perfect endings like **+a** ('he') and **+at** ('she'), the final consonant is geminated (or "lengthened", depending on your formal point of view). If, however, the suffix begins with a consonant, as in **+tu** ('I') or **+ta** ('you, masc. sg.'), then the separated or true spreading occurs.

From a phonological view, and reflecting the

tion, placing it on a separate transducer tape, but this difference is not important here. For extensions of this multi-tape approach see Kiraz (1994; 1996). The current approach differs from the multi-tape approaches in formalizing roots, patterns and vocalizations as regular languages and by computing ("linearizing") the stems at compile time via intersection of these regular languages (Beesley, 1998a; Beesley, 1998b).

[5] See, for example, the rules of Antworth (1990) for handling the limited reduplication seen in Tagalog.

dhamm+a	اِدْهَمَّ	'he turned black'
dhamam+tu	اِدْهَمَمْتُ	'I turned black'

Figure 3: Form IX Gemination vs. Spreading

notation of Beeston, it is tempting to formalize the underlying Form IX perfect active pattern as **CCaCX** so that it intersects with root **dhm** to form **dhamX**. When followed by a suffix beginning with a vowel such as **+a** or **+at**, phonologically oriented variation rules would realize the **X** as a copy of the preceding consonant (**/dhamm/**). Arabic abhors consonant clusters, and it resorts to various "cluster busting" techniques to eliminate them. The final phonological realization would include an epenthetical **/ʔi/** on the front, to break up the **dh** cluster, and would treat the copied **m** as the onset of a syllable that includes the suffix: **/ʔid-ham-ma/**, or, orthographically, اِدْهَمَّ. When followed by a suffix beginning with a consonant, as in **dhamX+tu**, the three-consonant cluster would need to be broken up by another epenthetic vowel as in **/ʔid-ha-mam-tu/**, or, orthographically, اِدْهَمَمْتُ. However, for reasons to become clearer below when we look at biliteral roots, we defined an underlying Form IX perfect active pattern **CCaCaX** leading to abstract stems like **dhamaX**.

2.3 Other Cases of Final Radical Gemination/Spreading

Other verb forms where the final radical is copied include the rare Forms XI and XIV. Root **lhj** (ل ه ج) intersects with the Form XI perfect active pattern **CCaaCaX** to form the abstract stem **lhaajaX** ("curdle"/"coagulate"), leading to surface forms like **/ʔil-haaj-ja/** (اِلْهَاجَّ) and **/ʔil-haa-jaj-tu/** (اِلْهَاجَجْتُ) that vary exactly as in Form IX. The same holds for root **ṣhb** (ص ه ب), which takes both Form IX (**ṣhabaX**) and Form XI (**ṣhaabaX**), both meaning "become reddish". In our lexicon, one root **qʕs** (ق ع س) takes form XIV, with patterns like the perfect active **CCanCaX** and imperfect active **CCanCiX** ("be pigeon-breasted"). Other similar Form XIV examples probably exist but are not reflected in the current dictionary.

Aside from the verbal nouns and participles of Forms IX, XI and XIV, other noun-like patterns also involve the spreading of the final radical. These include **CiCCiiX** and **CaCaaCiiX**, taken by roots **nħr** (ن ح ر), meaning "skilled/experienced", and **rʕd** (ر ع د) meaning "coward/cowardly". The **CaCaaCiiX** pattern also serves as the broken (i.e. irregular) plural for **CuCCuuX** stems for the roots **zʕr** (ز ع ر) meaning "ill-tempered", **ṣħr** (ص ح ر) meaning "thrush/blackbird", **lɣd** (ل غ د) meaning "chin", and **ṭħr** (ط ح ر) and **ṭxr** (ط خ ر), both meaning "cloud". When an **X** appears after a long vowel as in **ṭuxruuX**, it is always realized as a full copy of the previous consonant as in **/ṭuxruur/** (طُخْرُور), no matter what follows.

2.4 Middle Radical Gemination/Spreading

Just as Forms II and V involve gemination of the middle radical, other forms including Form XII involve the separated spreading of the middle radical. A preceding diphthong, like a preceding long vowel, causes **X** to be realized as a full copy of the preceding consonant, as shown in the following examples.

```
Root:     ħd    b
Pattern:  CCawXaC
Stem:     ħdawXab
Surface:  ħdawdab
Form:     Form XII perfect active
Gloss:    "be vaulted" "be embossed"

Root:     xʃ    n
Pattern:  CCawXiC
Stem:     xʃawXin
Surface:  xʃawʃin
Form:     Form XII imperfect active
Gloss:    "be rough"

Root:       xḍ    b
Pattern:   muCCawXiC
Stem:      muxḍawXib
Surface:   muxḍawḍib
Form:      Form XII active participle
Gloss:     "become green"
```

tamm+a	tamam+tu
تَمَّ	تَمَمْتُ

Figure 4: Biliteral Form I Stems

```
Root:     xḍ    r
Pattern:  CCiiXaaC
Stem:     xḍiiXaar
Surface:  xḍiiḍaar
Form:     Form XII verbal noun
Gloss:    "become green"
```

A number of nouns have broken plurals that also involve spreading of the middle radical, contrasting with gemination in the singular.

```
x f ʃ     "bat"   singular gemination
xufXaaʃ    خُفَّاش

x f ʃ     "bats"  plural spreading
xafaaXiiʃ  خَفَافِيش

d b r     "hornet"  singular gemination
dabXuur   دَبُّور

d b r     "hornets"  plural spreading
dabaaXiir  دَبَابِير
```

A few other patterns show the same behavior. While not especially common, there are more roots that take middle-radical-spreading noun patterns than take the better-known Form IX verb patterns.

3 Biliteral Roots

As pointed out in McCarthy (1981, p. 396-7), the gemination vs. spreading behavior of Form IX stems is closely paralleled by Form I stems involving traditionally analyzed "biliteral" or "geminating" roots such as **tm** (also characterized as **tmm**) and **sm** (possibly **smm**) and many others of the same ilk. As shown in Figure 4, these roots show Form I gemination with suffixes beginning with a vowel vs. full spreading when the suffix begins with a consonant. However Form IX is handled, these parallels strongly suggest that the exact same underlying forms and variations rules should also handle the form I of biliteral roots.

However, the Form I perfect active pattern, in the current notation, is simply **CaCaC** (or

Root:	k t b	k t b
Pattern:	CaCaC	CaCaC
Lexical:	katab+a	katab+tu
Surface:	katab a	katab tu
Orthography:	كَتَبَ	كَتَبْتُ

Figure 5: Ordinary Form I behavior

Root:	t m X	t m X
Pattern:	CaCaC	CaCaC
Lexical:	tamaX+a	tamaX+tu
Surface:	tamma	tamamtu
Orthography:	تَمَّ	تَمَمْتُ

Figure 6: Biliteral **tm** formalized as **tmX**

idiosyncratically for some roots, **CaCuC** or **CaCiC**). As shown in Figure 5, there is no evidence, for normal triliteral roots like **ktb**, that any kind of copying is specified by the Form I pattern itself.

Keeping **CaCaC** as the Form I perfect active pattern, the behavior of biliteral roots falls out effortlessly if they are formalized not as **sm** and **tm**, nor as **smm** and **tmm**, but as **smX** and **tmX**, with the copying-trigger **X** as the third radical of the root itself. Such roots intersect in the normal way with triliteral patterns as in Figure 6, and they are mapped to appropriate surface strings using the same rules that realize Form IX stems.

4 Rules

The TWOLC rule (Karttunen and Beesley, 1992) that maps an **X**, coming either from roots like **tmX** or from patterns like Form IX **CCaCaX**, into a copy of the previous consonant is the following, where Cons is a grammar-level variable ranging freely over consonants, LongVowel is a grammar-level variable ranging freely over long vowels and diphthongs, and C is an indexed local variable ranging over the enumerated set of consonants.

```
X:C <=>
    :C \:Cons+  _  %+: Cons   ;
    :C LongVowel _ ;
    :C X: : _ ;
    where C in ( b t θ j ħ x d ð r z
                 s ʃ ṣ ḍ ṭ ð̣ ʕ ɣ f q k
                 l m n h w y ) ;
```

The rule, which in fact compiles into 27 rules, one for each enumerated consonant, realizes underlying **X** as surface **C** if and only if one of the following cases applies:[6]

- First Context: **X** is preceded by a surface **C** and one or more non-consonants, and is followed by a suffix beginning with a consonant. This context matches lexical **dhamaX+tu**, realizing **X** as **m** (ultimately written اِدْهَمَمْتُ), but not **dhamaX+a**, which is written اِدْهَمَّ.
- Second Context: **X** is preceded by a surface **C** and a long vowel or diphthong, no matter what follows. This maps lexical **dabaaXiir** to **dabaabiir** (دَبَابِير).
- Third Context: **X** is preceded by a surface **C**, another **X** and any symbol, no matter what follows. This matches the second **X** in **samXaX+tu** and **samXaX+a** to produce **samXam+tu** and **samXam+a** respectively, with ultimate orthographical realizations such as سَمَّمْتُ and سَمَّمَ.

In the current system, where the goal is to recognize and generate orthographical words of Modern Standard Arabic, as represented in ISO8859-6, UNICODE or an equivalent encoding, the default or "elsewhere" case is for **X** to be realized optionally as a shadda diacritic.

5 Multiple Copies of Radicals

When a biliteral root like **smX** intersects with the Form II pattern **CaCXaC**, the abstract result is the stem **samXaX**. The radical **m** gets geminated (or lengthened) once and spread once to form surface phonological phonological strings like **/sammama/** and **/sammamtu/**, which become orthographical سَمَّمَ and سَمَّمْتُ respectively. And if both roots and patterns can contain **X**, then the possibility exists that a copying root could combine with a copying pattern, requiring a full double spreading of a radical in the surface string. This in fact happens in a single example (in the present lexicon) with

[6]The full rule contains several other contexts and fine distinctions that do not bear on the data presented here. For example, the **w** in the set C of consonants must be distinguished from the w-like offglide of diphthongs.

Root:	m k X
Pattern:	CaCaaXiiC
Abstract stem:	makaaXiiX
Surface:	makaakiik
Gloss:	"shuttles"

Figure 7: Double Consonant Spreading

the root **mkX**, which combines legally with the noun pattern **CaCaaXiiC** as in Figure 7. In the surface string **makaakiik** ("shuttles"), orthographically مَكَاكِيك, the middle radical k is spread twice. The variation rules handle this and the **smX** examples without difficulty.

6 System Status

The current morphological analyzer is based on dictionaries and rules licensed from an earlier project at ALPNET (Beesley, 1990), rebuilt completely using Xerox finite-state technology (Beesley, 1996; Beesley, 1998a). The current dictionaries contain 4930 roots, each one hand-coded to indicate the subset of patterns with which it legally combines (Buckwalter, 1990). Roots and patterns are intersected (Beesley, 1998b) at compile time to yield 90,000 stems. Various combinations of prefixes and suffixes, concatenated to the stems, yield over 72,000,000 abstract words. Sixty-six finite-state variation rules map these abstract strings into fully-voweled orthographical strings, and additional trivial rules are then applied to optionally delete short vowels and other diacritics, allowing the system to analyze unvoweled, partially voweled, and fully-voweled orthographical strings.

The full system, including a Java interface that displays both input and output in Arabic script, is available for testing on the Internet at `http://www.xrce.xerox.com/research/mltt/arabic/`.

References

Evan L. Antworth. 1990. *PC-KIMMO: a two-level processor for morphological analysis.* Number 16 in Occasional publications in academic computing. Summer Institute of Linguistics, Dallas.

Kenneth R. Beesley. 1990. Finite-state description of Arabic morphology. In *Proceedings of the Second Cambridge Conference on Bilingual Computing in Arabic and English*, September 5-7. No pagination.

Kenneth R. Beesley. 1996. Arabic finite-state morphological analysis and generation. In *COLING'96*, volume 1, pages 89--94, Copenhagen, August 5-9. Center for Sprogteknologi. The 16th International Conference on Computational Linguistics.

Kenneth R. Beesley. 1998a. Arabic morphological analysis on the Internet. In *ICEMCO-98*, Cambridge, April 17-18. Centre for Middle Eastern Studies. Proceedings of the 6th International Conference and Exhibition on Multilingual Computing. Paper number 3.1.1; no pagination.

Kenneth R. Beesley. 1998b. Arabic stem morphotactics via finite-state intersection. Paper presented at the 12th Symposium on Arabic Linguistics. Arabic Linguistic Society, 6-7 March, 1998, Champaign, IL.

A. F. L. Beeston. 1968. *Written Arabic: an approach to the basic structures.* Cambridge University Press, Cambridge.

Steven Bird and Patrick Blackburn. 1991. A logical approach to Arabic phonology. In *EACL-91*, pages 89–94.

Timothy A. Buckwalter. 1990. Lexicographic notation of Arabic noun pattern morphemes and their inflectional features. In *Proceedings of the Second Cambridge Conference on Bilingual Computing in Arabic and English*, September 5-7. No pagination.

Zelig Harris. 1941. Linguistic structure of Hebrew. *Journal of the American Oriental Society*, 62:143–167.

Clives Holes. 1995. *Modern Arabic: Structures, Functions and Varieties.* Longman, London.

Grover Hudson. 1986. Arabic root and pattern morphology without tiers. *Journal of Linguistics*, 22:85--122. Reply to McCarthy:1981.

Ronald M. Kaplan and Martin Kay. 1981. Phonological rules and finite-state transducers. In *Linguistic Society of America Meeting Handbook, Fifty-Sixth Annual Meeting*, New York, December 27-30. Abstract.

Ronald M. Kaplan and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics*, 20(3):331–378.

Lauri Karttunen and Kenneth R. Beesley. 1992. Two-level rule compiler. Technical Report ISTL-92-2, Xerox Palo Alto Research Center, Palo Alto, CA, October.

Lauri Karttunen, Ronald M. Kaplan, and Annie Zaenen. 1992. Two-level morphology with composition. In *COLING'92*, pages 141–148, Nantes, France, August 23-28.

Lauri Karttunen. 1991. Finite-state constraints. In *Proceedings of the International Conference on Current Issues in Computational Linguistics*, Penang, Malaysia, June 10-14. Universiti Sains Malaysia.

Lauri Karttunen. 1993. Finite-state lexicon compiler. Technical Report ISTL-NLTT-1993-04-02, Xerox Palo Alto Research Center, Palo Alto, CA, April.

Laura Kataja and Kimmo Koskenniemi. 1988. Finite-state description of Semitic morphology: A case study of Ancient Akkadian. In *COLING'88*, pages 313–315.

Martin Kay. 1987. Nonconcatenative finite-state morphology. In *Proceedings of the Third Conference of the European Chapter of the Association for Computational Linguistics*, pages 2–10.

George Kiraz. 1994. Multi-tape two-level morphology: a case study in Semitic non-linear morphology. In *COLING'94*, volume 1, pages 180–186.

George Anton Kiraz. 1996. Computing prosodic morphology. In *COLING'96*.

Kimmo Koskenniemi. 1983. Two-level morphology: A general computational model for word-form recognition and production. Publication 11, University of Helsinki, Department of General Linguistics, Helsinki.

John J. McCarthy. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, 12(3):373–418.

Hans Wehr. 1979. *A Dictionary of Modern Written Arabic.* Spoken Language Services, Inc., Ithaca, NY, 4 edition. Edited by J. Milton Cowan.