DRAFT
Overview of Palestinian Arabic Phonology

Jack Halpern <jack@cjki.org>
The CJK Dictionary Institute

Version: V0.96
Created: April 9, 2021
Updated: December 4, 2021

1. Introduction

1.1 Goals [E]

This paper describes the phonology of the urban variety of South Levantine Arabic, commonly known as Palestinian Arabic (PA), and introduces a new romanization system for PA, called **DARS.** It aims to provide an overview of PA phonology from the point of view of the language learner; that is, to help learners gain a solid grasp of the PA sound system, and to provide a pedagogically-oriented description of the PA phoneme stock
[add linguistic analysis]

The paper does not aim to describe the underlying theory or mechanisms of such phenomena as vowel harmony, epenthesis, elision, assimilation, prosody and neutralization, which are treated in other works [x, y, z]. Unlike CODA [x], a morphophonemic conventional orthography for dialectal Arabic using the Arabic script, DARS is phonemic (more precisely *phonemo-phonetic*) orthography with focus on the practical needs of the learner and on serving as a practical guide for PA speech technology.

1.2 Main Features

Though the first version of DARS is fine tuned for accurately transcribing Palestinian Arabic, it has various distinctive features that make it ideally suitable for the phonemic transcription of Arabic dialects in general. These include:

- Accurately represents all phonemes, especially short vowels
- Has two versions: one pedagogic oriented and one more academic oriented.
- Explicitly represents indicates some allophonic variation, especially back /a/
- Explicitly and precisely indicates long vowels, vowel neutralization, vowel lengthening and vowel elision
- Explicitly and precisely indicates the different kinds of epenthetic vowels (detailed rules in $4.3.2)
- Outs special emphasis on accurately representing word stress (detailed rules in $7.3)
- Consists of a friendly set of symbols
- DARS+ provides a set of substitute symbols for easy typing, along with conversion tools
- Represents explicitly realization resulting from such phenomena as velarization, retraction, elision and degemination

1

1.3 Conventions [EE]

DARS transcriptions are given in italics (DejaVu sans) (*ba),* phonemes between slashes (/ba/), and IPA transcriptions are enclosed in brackets ([ba]) (except inside tables). The IPA column is a fairly accurate phonetic representation of PA phonemes, but needs further research and verification to achieve higher precision. A superscript ▲ following an Arabic letter in the tables indicates that that letter does not, or cannot, correctly represent the PA phoneme (technically, a grapheme-phoneme mismatch).

2. Romanization vs. Arabic script

2.1 Why romanization

Romanization uses the Latin alphabet to represent a non-Roman script, such as Chinese or Arabic. This includes transcription and transliteration. *Transliteration* represents the source graphemes (not phonemes) with Latin letters, as in محمد /mHmd/ (mostly useless for learners but used in NLP). *Transcription* represents source words in the Latin script in a manner that reflects their pronunciation [xy]. This includes *phonetic* transcriptions, which transcribe actual speech sounds, as in [muħëmmëd] (IPA), or *phonemic* transcriptions, which transcribe the phonemes (not phones) of the source language, as in /muHammad/. The DARS system introduced here is basically a hybrid phonemic-phonetic system optimized for learners,

Do learners of Palestinian Arabic need a romanization system? The answer is a definitive yes. The same answer applies to other Arabic dialects, and even to MSA in the beginner stages [x]. This paper demonstrates that the Arabic script, even when fully vocalized, is unsuitable for representing the phonemes of PA, which is basically true of other Arabic dialects. This is especially true for learners striving to learn how to pronounce PA correctly.

Educators who have not studied this issue deeply unthinkingly promote the Arabic script based on misguided preconceptions, citing "tradition and culture", "beauty of the script", "etymological link to MSA", and the like, in their defense. To make up for the inadequacy of the script, they often list many complex rules, such as how to pronounce short vowels in different contexts. But no matter how detailed the rules, they are inadequate because (1) PA phonology is very different from MSA, and (2) because the script is incapable of representing the many phonemes in PA, let alone such important features as word stress and vowel neutralization.

[mention CAVE developing PA CAVE]
[add somewhere just the pleasure of reading no guessing]

2.2 Ambiguity of Arabic script

2.2.1 Highly ambiguous. A major issue is that unvocalized Arabic is highly ambiguous [xx]. In MSA, a string like كاتب can represent seven distinct wordforms, such as *kātibun, kātibin* and *kātibu*. Since PA has many more vowels than MSA, it is even more ambiguous. A truly phonemic transcription represents all source phonemes unambiguously. These issues are dealt with in detail in Halpern [2009a, 2007].

In addition, there are no accepted norms for writing specific dialects in the Arabic script, including Palestinian Arabic. PA written in the script is inconsistent, has various orthographic variants, and, above all, is incapable of representing the large number of PA-specific phonemes. For educated native speakers, it is somewhat usable (but not ideal) for two reasons: (1) because they already *know* how to pronounce their dialect correctly, and (2) they know MSA and can grasp the etymological relations between MSA and their dialect.
[see Matthews]

Here are some specific reasons why the Arabic script is inadequate, especially for learners.

1. PA has six short basic vowel phonemes, in addition to two non-phonemic epenthetic vowels, as opposed to three in MSA. So even fully vocalized PA does not have enough diacritics for representing about 9 vowels with 14 allophones (see tables 1 and 2 below).

2. Various Arabic script consonants represent multiple PA consonants while multiple script consonants represent a single phoneme (see Table 6).

3. The two epenthetic (§x.x below) vowels in PA are not -- cannot -- be represented in the script.

4. Word stress, of utmost importance to the learner, cannot be represented in the script.

5. Neutralization (§x.x) --which long vowels to shorten and which double consonants to undouble -- cannot be represented.

6. Important allophones, especially /a/ vs. /A/, cannot be represented.

The Arabic script is not capable of representing the last four features at all, and even fully vocalized Arabic cannot unambiguously represent the first two. For these reasons, a romanized transcription is of critical importance, especially in the initial stages.

2.2.2 Short vowel ambiguity. It goes without saying that PA written in unvocalized Arabic is highly ambiguous. But even PA written in vocalized Arabic, which is not common, is also ambiguous. *Kasra* can represent three vowels with five allophones, *damma* two vowels with three allophones, and *fatha* two vowels with three allophones (see Table 1). Surprisingly, even *sukuun* actually has three values: *absence of vowel,* epenthetic /ᵉ/, and epenthetic /ᵒ/. For example, ب is pronounced with no vowel in قَبْلِي *'abli,* as /bᵉ/ in قَبْل مَا *'abᵉl ma,* and as /bᵒ/ in صُبْح *ṣubᵒḥ*

2.2.3 Long vowel ambiguity. Though long vowels are marked explicitly in the script by ا and ي and و, they too are ambiguous as a result of **vowel neutralization** (see §4.5). For example, the long vowel marker ا is neutralized (shortened) in unstressed vowels so that أَنَا يَابَانِي is not pronounced as *ʾanā yābānī,* but as *ʾán̲a y̲abān̲i̲,* where the underlined vowels indicated that they have been neutralized. Thus even the vocalized أَنَا يَابَانِي, not to speak of the unvocalized أنا ياباني, cannot possibly represent the correct pronunciation of *án̲a y̲abān̲i̲.*

2.2.4 Consonant ambiguity. Consonants too are ambiguous, but to a lesser extent. ث, for example, is normally pronounced [s] or [t], but in some words as [θ], while ذ is normally pronounced [z] but in some words as [ð]. "Some words" means using the MSA pronunciation in a higher register of language.

Particularly troublesome is ق, which in urban PA is normally pronounced as the glottal stop [ʔ] but in some words as [q] like in MSA (e.g. *mūsiqa*), and in some regions as [k] or [g]. Surprisingly, the emphatic, though pronounced [sˤ] as in MSA, in such words a صغير it is pronounced as [z], that is *zghīr,* not as *ṣghīr.* Famously, قَهْوَ serves as a shibboleth for subdialect identification. It can be pronounced as *ʾahwe* (urban), *q̈ahwe* (Druze), *kahwe* (?), and *gahwe* (Bedouin)
[what regions]

2.2.5 Complex rules
It is not possible to predict these variations by rules. Even if the learner masters the complex rules and exceptions, they are simply inadequate. He/she must constantly engage in a hit-and-miss guessing game with a high chance of getting it wrong. Let us take the word …...............
[give examples of rules /e/ at end or foreign words]

[add detailed analysis of one word

2.3 In defense of romanization
To address the shortcomings of the Arabic script, there is a need for a phonemic transcription system that can fully and accurately represent the phonemic and prosodic features of PA, enabling learners to read and write dialectal Arabic correctly and effectively without rules and without guesswork.

Professional educators have long recognized that from a pedagogical point of a romanized transcription is essential. In *Eastern Arabic* [x], considered an excellent books though outdated, Rice and Sa'id argue:

> The transcription is better suited to represent the sounds and forms
> of modern spoken Arabic than is the traditional Arabic orthography, and
> is also a more efficient instrument for purposes of linguistic analysis
> and stating of grammatical facts

In the front matter of the *Olive Tree Dictionary* [x], arguably the best PA dictionary ever published, Elihay explains why he does not use the Arabic script.

> "There is another, more scientific reason for preferring transcription: Arabic script suits MSA, with its... limited range of sounds. The sounds of PA are more plentiful, and more complex: its vowels are more numerous and more variable.... When Arabs write PA... they do so in rather a hit-and-miss style... PA as written in Arabic script is ...understood only by someone who already knows the spoken language well. "English" transcription, on the other hand... allows us to give the vowels their due and to depict the sounds of spoken Arabic with great precision. All books on Arabic dialects make use of transcription."

[add Mathews]

Academic publications on the phonology of Arabic dialects naturally use romanized transcription since it is literally impossible to represent dialectal phonemes with the Arabic script. Pedagogical materials like books and websites generally also use romanization. The few that don't do their users a disservice by presenting them with ambiguous and inaccurate pronunciation.

2.4 Books and reference works
A major website for the study of Palestinian Arabic, *palestinianarabic.com* is full of useful information for learning PA. Under the section "Palestinian Arabic Coursebooks", the following three books are recommended.

1. **Speaking Arabic:** A Course in Conversational Eastern Arabic (Palestinian)
   by J. Elihay

2. **Eastern Arabic:** An Introduction to Palestinian Arabic
   by Frank A. Rice and Majed F. Sa'id
   Georgetown University Press

3. **Colloquial Palestinian Arabic:** An Introduction to the Spoken Dialect (Arabic Edition),
   by Nasser M Isleem
   Alucen Learning

The first two are for beginner to intermediate learners, and the last one for advanced students who already know MSA. Naturally, the first two use romanization. The first, consisting of four volumes, is an outstanding work -- one of the best language courses anywhere. It introduces a superb romanization system which forms the basis of the DARS system described in this paper.

In addition, the following book for advanced students is an excellent reference work for studying verbs, rather than a coursebook.

4. **The Most Used Verbs in Spoken Arabic:** Jordan & Palestine

by Fridrik E. Tiedemann Jr. (Author)

This book, as well as Colloquial Palestinian Arabic, are useful for more advanced students but unfortunately both have a major drawback: only the Arabic script is used.

5. **Olive Tree Dictionary**
A Transliterated Dictionary of Conversational Eastern Arabic (Palestinian)
by J. Elihay
Minerva Publishing House

This is no doubt the best dictionary for PA, by the same author of (1) *Speaking Arabic.* This outstanding dictionary meets very high lexicographic standards and puts great emphasis on precise pronunciation and word stress. Though the subtitle mistakenly uses the term "transliterated", this dictionary uses an excellent phonemic transcription, which served as a basis of the DARS system introduced here.

[add references]

3. DARS Transcription [E]

DARS, an acronym for **Dialectal Arabic Romanization System**, refers to a transcription system developed by the author based on the excellent romanization system invented by J. Elihay used in the *Olive Tree Dictionary* [1], and on the author's pedagogically oriented MSA romanization system called CARS, described in Halpern [2].

DARS is a pedagogically-oriented, innovative phonemic-phonetic transcription system whose goal is to aid learners accurately pronounce Arabic dialects, and enable linguists to analyze the phonological and phonetic characteristics of specific dialects. It's ultimate goal is to cover all major Arabic dialects, but this first version is limited to Palestinian Arabic, to be followed by Egyptian Arabic.

The most important characteristic of DARS is its readability: it encodes phonological, allophonic, and prosodic information in an intuitive, easy to read set of symbols that represent PA phonemes unambiguously, as well as explicitly indicates word stress, vowel neutralization, vowel length, and epenthetic vowels. The system is completely unambiguous. Each symbol represents a single phoneme. On the other hand, a small number of phonemes are unambiguously mapped to alternative symbols for pedagogic simplicity (see §x.x).

DARS functions as a precise representation of PA phonemes, including some allophones that are critical to correct pronunciation, such as the back /a/ represented by /ɑ/ (open back unrounded vowel). It is essentially a *phonemic orthography,* or, more precisely, a *phonemic-phonetic orthography,* not an etymological orthography, and certainly not a graphemic transliteration of the Arabic script. These terms are discussed in detail in Halpern [2].

The symbols used in DARS are intuitive and easy to read, but some are difficult to type and display. For this reason we devised a companion system, called DARS+, a set of ASCII

symbols that can be used as substitutes for DARS symbols in email, text files, and the like. It was carefully designed to be easy to type and to enable round-trip conversion to DARS. We have also developed a tool to perform such round-trip conversion.

In Arabic DARS is written دَرْس *dars,* which by coincidence means 'lesson'. This aptly describes the pedagogic goals of this transcription system.

4. Vowels [E]

4.1 Representing Vowels [E]
The tables below describe the phonemes of urban Palestinian Arabic, roughly covering northern and central Israel, the West Bank and western Jordan. They have been researched and compiled rather carefully, but this is still a work in progress. Reviewers are welcome to comment and point out errors, especially in the section "4.2 Short vowels." The tables introduce the symbols to represent all vowel phonemes in DARS, DARS+ and IPA.

4.2 Short vowels [E]
PA has six short vowels and various allophones (as opposed to three in MSA). For grammatical and etymological reasons, vowels, especially short vowels, are represented in various ambiguous ways in the Arabic script, or not represented at all. From the script alone, there is no way to determine the short vowels without a prior knowledge of the language. Even with diacritics, there is no way to indicate the difference between the following pairs since they use identical diacritics: /a/ and /ɑ/, /i/ and /e/ , /o/ and /u/, and certainly no way to distinguish such subtle allophonic differences as between [ʊ] and [u] and [ɪ] and [i].

Table 1. Short Vowels [E]

No	DARS	DARS+	IPA	Arabic	Comments	Example	Elihay
S1	**a**	a	**1** a~æ **2** a~ʌ	◌́ *fatha*	**1** as in Spanish *gato,* tends to [æ] in some regions, as in *kalb* **2** *as final,* tends to [ʌ] as in *but*	ʾabu	a
S2	**ɑ**	A	ɑ	◌́ *fatha*	in vicinity of emphatics, as in *father* but shorter, like *lot* in American English	ʾaḥra	ɑ
S3	**i**	i	**1** ɪ~e **2** i **3** ɵ	◌̣ *kasra*	**1** *as medial,* as in *bit,* tending to [e] **2** *as final,* strongly tends to [i], as in [bɪddi] **3** before emphatics like /ṣ/, /ḍ/, /ṭ/, pronounced further back, almost like /o/, e.g. *btíḍḥak* is similar to *btóḍḥak*	**1** ʾíḥna **2** bíddi **3** btíḍḥak	i
S4	**e**	e	e~i	◌̣	Roughly as in *led* but tends to [i], as in *lāzem.* Occurs in unstressed		e

					closed syllables and final Cicc?? (…). Occurs only as final, except in loanwords.	*wāled*	
S5	**o**	O	ʊ	◌ُ *damma*	*as medial* in unstressed closed syllables, between [o] and [u] *as* in *book,* e.g. final 3rd person possessive in *bíddo*	*bēto*	O
S6	**u**	u	**1** ʊ **2** u	◌ُ *damma*	**1** *as medial,* between [o] and [u] (…) as in *book,* **2** *as final,* as in *boot* but short. In contexts other than /o/.?	*'ujra*	u
S7	**-**	**-**	**-**	◌ْ *sukuun*	*Sukuun* (no vowel) is not shown in DARS+ , e.g. *3índi* (عِنْدِي)	*mufrad*	**-**

4.3 Vowel epenthesis

4.3.1 Extra short vowels [E] PA has an interesting feature which gives it a distinctive flavor: the insertion of a extra short vowel, technically called *epenthesis,* between consonant clusters for ease of pronunciation, or euphony. Technically referred to as *anaptyxis* [x], the process of adding epenthetic vowels (also called helping vowels) in PA occurs only in unstressed syllables. Epenthetic vowels can also be added to the beginning of a word, referred to as *prothesis.*

Omitting these vowels makes PA speech sound unnatural, but unfortunately for the learner, romanized Arabic does not normally indicate these vowels, not to speak of the Arabic script. Some systems do indicate them but don't distinguish between epenthetic vowels and normal vowels, so that they look exactly the same.

Epenthetic vowel are very short. PA has two such vowels, an extra short /e/ and extra short /o/, and one short semivowel /y/.

Table 2. Extra Short Vowels [E]

No	DARS	DARS+	IPA	Arabic	Description	Example	Elihay
E1	e	ê	ə~ĕ	-	mid-central extra short /e/	tends to [ĕ], as in *lektāb*	e
E2	o	ô	ə~ŏ	-	mid-central extra short /o/	tends to [ŏ], as in *shugol*	o
E3	**yy**	ŷ	ji	-	palatalized short /i/?	stressed possible?	y

The most common epenthetic vowel is /e/. For example, إلْكْتَاب 'the book' goes through epenthesis and is pronounced *lektāb,* not *'ilktāb,* as one might expect. The /e/ is an ambiguous

schwa sound between [i] and [e], tending to [ĕ], used only in unstressed syllables. It is somewhat similar to the the final *a* in *sofa*. Note that /ᵉ/, the most common one, may sound similar to /ᵒ/, as in *shugᵒl* 'work'. /ᵒ/ is basically an extra short [o] or [ŏ]. An example of prothesis is *ᵉktāb*, where /ᵉ/ is added to the beginning.

4.3.2 Insertion rules [Edit!] Epenthetic vowels are mostly inserted to break up consonant clusters, which includes all words ending in CC, such as *kálb → kálᵉb* and *katábt → katábᵉt*. In some cases they are optional so the same word can be pronounced with or without them (see below). Elihay [x] indicates these vowels explicitly in his dictionaries and learning materials [x]. The rules below are based on Elihay.

1. the definite article followed by two consonants becomes *lᵉ* : Usually this only if the first consonant is not a sun letter, as in *ʾilktāb → lᵉktāb,* but sometimes it is applied to sun letters, so that both *ʾizzghír* and *lᵉzghír* are possible.

2. words beginning with two consonants at sentence beginning or following a final consonant may have /ᵉ/ inserted, as in *ktāb → ᵉktāb* and *wálad kbír → wálad ᵉkbír*.

3. words ending in two consonants generally have /ᵉ/ or /ᵒ/ inserted, such as at sentence end and before a consonant followed by a vowel, as in *kīf_ishshúgᵒl, kíf kān ʾabᵉl, ʾábᵉl sána.*

4. before the following pronoun enclitics -ni, -ha, -na, -kom, -ku, -hom, -hen, as in *ʾbᵉl-kom* and shughᵒl-na.

5. epenthetic vowels are not inserted (1) before vowels, (2) before the article *il-* (in its various forms like *is-, it-*), (3) before he following the pronoun enclitics *-i, -ak, -ek, -o,* and (4) words beginning with two consonants because rule #2 applies, e.g. *ʾabl snín → ʾabl ᵉsnín*
[confirm contradiction]

4.3.3 Representing epenthetic vowels [E] The Elihay and DARS romanization systems, and rarely some websites, indicate epenthetic vowels explicitly, so that it is not necessary to memorize the rules. But on the whole epenthetic vowels are not indicated in learning materials.

In DARS, the superscript /ᵉ/ and /ᵒ/ are used to make it clear that they are short and fleeting, sometimes optional. Since superscript symbols are difficult to type, in DARS+ they are represented by "^". e.g. *l ᵉktāb* is typed as "le^ktaa/b" ("/" represents the accent mark). It is not clear how to precisely represent these vowels in IPA. Basically, they are ambiguous mid-central vowels with /ᵒ/ tending toward [o] and /ᵉ/ tending towards [e]. The IPA for /ʸ/, which is used less frequent, is probably [ʲ]. We could not locate the IPA for epenthetic vowels in any sources, so this requires further investigation.

It goes without saying that it is not possible to represent epenthetic vowels in the Arabic script. Most native speakers are probably not even aware that such vowels exist, but

instinctively know when to insert them. With practice, learners can learn how to do that too, but in the initial stages it is best to use materials that indicate these vowels explicitly, and to learn the basic insertion rules.

4.4 Long vowels [E]

Long vowels must be distinguished from short vowels. In the script they are indicated by ا, ي and و, and in DARS by macrons. PA has six long vowels: /ā/, /ā/, /ī/, /ē/, / ō/ and /ū/, only three of which can be represented by the script (/ā/, /ī/ and /ū/). It is important to note that *only stressed long vowels are pronounced long.* In unstressed positions they are (almost always) pronounced short. Thus أَنَا is pronounced *ʾánạ*, not *ʾánā*, where /ạ/ indicates the neutralized (shortened) vowel.

The Arabic script can easily mislead one to pronounce all long vowels as long. The neutralization of long vowels, which is shown explicitly by the DARS romanization, is not a self-evident phenomenon. In the table below, ب represents non-emphatic "normal" consonants, whereas ص represents the emphatic consonants ص ض ط ظ ع ح as well as ر and occasionally others such as ب and م.

Table 3. Long Vowels

No	DARS	DARS +	IPA	Arabic	Comments	Example	Elihay
L1	**ā**	aa	aː	بَا	as in *aah,* following non-emphatic consonants	*lᵉktāb*	ā
L2	**ā̱**	AA	1 ɑː 2 ɑ	صَا صَ	**1** back /a/ pronounced long as in *father* **2** back /a/ pronounced short after emphatics (→ 3.6)		ā̱
L3	**ī**	ii	1 iː 2 i	بِي	**1** *as medial,* like *ee* in *meet:* **2** *as final,* like *ee* in *meet* but short	**1** *jīb* **2** *jib-li*	ī
L4	**ū**	uu	1 uː 2 u	بُو	**1** *as medial,* as in *tool* **2** *as final,* like in *boot* but short	*shūf*	ū
L5	**ē**	ee	eː eː~iː	بِي، بَيْ	*as medial,* tends towards [iː]	*wēn*	ē
L6	**ō**	oo	oː	بَوْ، بُو	long version of /o/ (S5)	*ʾọda*	ō

NOTE: The vowels in the Arabic columns below are shown with a dummy supporting consonant, since it is difficult to represent vowels independently in the Arabic script.

4.5 Velarization and Retraction [E]

PA has two series of phonemes referred to as *emphatic* or *velarized* consonants. The *primary* series consists of ص /ṣ/, ض /ḍ/, ط /ṭ/ and ظ /ẓ/ (they have no equivalents in English). and the *secondary* series consists of ب /b/, ل /l/, م /m/ and ر /r/.

The presence of a primary series consonant triggers v*owel retraction,* an important phenomenon in both MSA and PA. This means that the front vowel /a/ is retracted and is pronounced as the back vowel /ɑ/ (as in *father,* but short) in certain environments, especially *before and after* the primary series. Thus ضَ, صَ, طَ, ظَ are pronounced as /ṣɑ/, /ḍɑ/, /ṭɑ/ and /ẓɑ/. The long /ā/ behaves in similar way but is long, as in *father.*

The secondary series can also trigger vowel retraction. /r/ does so frequently, while /b/, /m/ and /l/ much more rarely. In *maraḍ* [mɑrɑḍ]*, for example, /ɑ/* occurs both before and after the /r/. In *daftarak* [daftɑrak] retraction occurs only *before* /r/, whereas in *maryam* no retraction occurs at all.

The behavior of /b/, /m/ and /l/ is even more erratic. There is no way to predict when they become velarized. They rarely trigger vowel retraction, and when they do there is no way to predict when and in what position. Thus in *mabṣūṭa* and *madrase* /m/ is followed by an ordinary front /a/, but in *qawắm* /m/ is preceded by a retracted *ắ* . In *ᵉnshalla,* /a/ is retracted both before and after the /l/ and becomes /*ɑ*/, whereas in *hallaq* it remains a front vowel /a/ both before and after.

Sometimes non-emphatic consonants trigger retraction too. Foe example, in *bayāḍ*, not only does /b/ trigger retraction, but strangely /y/, not an emphatic, does so too. And in *kafar*, /k/ is followed by /*a*/, probably by presence of the final /r/ though they are not contiguous.

The emphatics also cause /i/ to be centralized, that is, to be pronounced further back, as in *btiḍḥak*, where the /i/ is pronounced [ɪ], not [I], as for example …...............

As can be seen, vowel retraction is a complicated phenomenon. The Arabic script does not, (is incapable of) indicating retracted vowels. The rules are complex, have many exceptions, and for the secondary series retraction is basically unpredictable. Even if explicit rules did exist, neither learners nor algorithms can be expected to master them. The only sure way is indicate all retracted vowels explicitly, as is done in the DARS and Elihay romanizations and rarely in other systems.

5. Neutralization

5.1 Neutralized Vowels [E]

Vowel neutralization, the shortening long vowels to short vowels, is of great importance in both PA and MSA. For MSA, this topic is treated fully in Halpern [3].

<div align="center">

Table 4. Neutralized Long Vowels
[E] but add examples

</div>

No	DARS	DARS+	IPA	Arabic	Commments	Example	Elihay
N1	**a̲**	aa_	a	-	neutralized /a:/	*yā̆_bā̆ni*	-
N2	**ɑ̲**	AA_	ɑ	-	neutralized back /a:/	<...>	-
N3	**i̲**	ii_	i	-	neutralized /i:/	<...>	-
N4	**o̲**	oo_	o	-	neutralized /o:/	<...>	-
N5	**u̲**	uu_	u	-	neutralized /u:/	<...>	-
N6	**e̲**	ee_	e	-	neutralized /e:/	<...>	-

5.2 Representing neutralization [E]

In MSA, the neutralization rules are somewhat complex. In PA, basically the rule is simple. All unstressed long vowels are neutralized. In the Arabic script, neutralized vowels are written as long vowels, as in أنا *ʾánā* 'I' and ياباني *yābā̆nī* 'Japanese', just like normal long vowels, but in DARS the neuralization is shown explicitly by the uderline in *ʾánā_* and *yā_bā̆ni*.

Vowel neutralization is not represented in the script, though theoretically one can write أنَّا as أنَّ (making Arabs cringe). Thus the Arabic script, or a direct transliteration of it, makes it seem as if all long vowels are pronounced long, whereas in fact they are not.

In DARS, there are two ways to represent neutralized vowels, pronounced short in both cases.

1. The underlined vowel indicate that the underlying vowel is long but has been shortened. For example, أنا ياباني is written *ʾánạ yạbā̆nị*, not *ʾánā yābā̆nī*.

2. The neutralized vowels are written as they are actually pronounced; that is, as short vowels, as in *ʾána yabā̆ni*.

The advantage of the first method is that it is etymological; that is, it makes it clear that the original vowel was long but is shortened (academically oriented). The advantage of the second method is that it is simple and phonemic; that is, it represents the vowel as it is actually pronounced, which is short (pedagogically oriented).

If the romanization system does not indicate neutralization explicitly, the learner has to know where the stress falls, since unstressed vowels are neutralized. But learners have no way of knowing that, even if they master the somewhat complex rules (which requires a knowledge of phonotactics). Some of the better learning materials, such as Elihay [x] and Yoda [y], show vowel neutralization explicitly by representing final neutralized vowels as short vowels

(Option 2). This is a pedagogically sound decision and is what we recommend when using DARS for learning and reading, but for linguistic analysis it is better to use Option 1 (underlining), which shows explicitly which vowels are neutralized.

The are some other, more subtle issues, such as pronouncing neutralized vowels as **half long,** but this is an advanced topic that will be omitted here (see Halpern [x]).

5.3 Vowel Elision [E]

Another kind of neutralization, called *vowel elision,* occurs when pronoun suffixes (enclitics) are attached to verb conjugations ending in a consonant. For example, the long vowel in the imperative *qūl* 'tell' is neutralized in *qúl-li* 'tell me', *baqūl* becomes *baqúl-lak* 'I'll tell you', and *bajīb* becomes *bajíb-lak* 'I'll bring you'. Another kind of vowel elision occurs in verbs whose third person masculine form begins with *i*, as in *inbāll* 'it got wet', as opposed to verbs beginning with *yi*, as in *yíftaḥ* 'open'. In such cases the initial *i* is elided when the previous word ends in a vowel, e.g. *ma_bállᵉsh*.

Vowel elision also occurs (EAPD #6)

5.4 Vowel Lengthening

As already pointed out, vowel neutralization occurs when pronoun suffixes are attached to a verb. For example, the long vowel in *qūl* is neutralized to *u* as in *qúl-li* 'tell me!' (see §.x.x). On the other hand, sometimes the opposite phenomenon occurs: *vowel lengthening*. This happens when enclitics such as pronouns and negation suffixes are attached to words ending in vowels.

For example, the *a* in *wa/ra* 'behind' becomes * á* in *warák* 'behind you' and *waráʰ* 'behind him', and the final *u* in *jābu* becomes *ū* in *ma jā_bū-sh* 'they didn't bring'. In such cases the stress always shifts to the lengthened vowel. It should be noted that lengthened vowels *can* be represented in the Arabic script, as in [[waraAk]] (but of course the retracted *ā* is shown as a nornal *ā*).

Another example is stress shift when enclitics are attached to nouns. For example, in *mádrase* 'school' the stress shifts from the first syllable to the last syllable in *madrásti* 'my school' because the pronounc clitic *i* is added.

6. Consonants

6.1 Representing Consonants

The consonants in PA are fairly similar to those in MSA, but there are several ambiguous consonants. That is, one MSA consonant can be realized as several distinct consonants in PA.

DARS represents consonants phonemically. The glottal stop in the Arabic script is represented by *hamza* (ء) In urban PA, both ء [ʔ] (*hamza*) and ق [q] (voiceless uvular plosive) are normally pronounced as a glottal stop. For the convenience of learners DARS represented both by /ˈ/ (except in academic contexts where /q/ represents ق but still pronounced [ʔ]. In

some regions ق is ronounced [q] (or they may be in free variation). However, some words in PA, like مُوسِيقَى *mūsi̱q̇a̱* 'music', ق is always pronounced [q] as in MSA. In DARS this is represented by /q̇/. (Note that in *mūsi̱q̇a̱* the two underlined neutralized long vowels /i̱/ and /a̱/, which must be pronounced short).

An example of an ambiguous consonant is ث, pronounced [θ] in MSA. This is normally pronounced [t] or [s] in PA and occasionally [θ] for some words derived from MSA. Similarly, ذ is pronounced [d] or [z] in PA and occasionally [ð], as in MSA. For example, مَثَلًا 'for example' is pronounced *mátalan*, *másalan* or sometimes *máthalan*, as in MSA.

Theoretically such consonants can be made unambiguous in the Arabic script by replacing ث with تس or ت for [s] and [t], but this is unnatural and probably makes most Arabs cringe. So in practice, ث and ذ are ambiguous when PA is written in the Arabic script.

[Baku-san replace ^ with superscript ▲]

Table 6. Consonants

No	DARS	DARS+	IPA	Arabic	Comments	Elihay
C1a C1b	ʾ '	' "	ʔ	ء	*Hamza,* pronounced as the glottal stop, can be written as /'/ instead of /ʾ/ for simplicity	'
C2	**b**	b	b	ب	Like *b* English in *box* but sometimes velarized (→ 3.6)	b
C3	**p**	p	p	ب	Like *p* English in *spear,* as in بْرُوبَة *próva.* Used only in some loanwords.	p
C4	**t**	t	t	ت ث^	Like *t* in English *step.* Used for both ت and ث but in some literary words ث may be pronounced /th/, as in *thaqáfe*	t
C5	**j**	j	ʒ	ج	Like *s* in English *pleasure* but like *j* in *jazz* in Northeastern Levantine and rural Palestine	j
C6	**d**	d	d	د ذ^	Like *d* in English *dart.* Used for both د and ذ but in some literary wordsذ may be pronounced /dh/, as in *dháki* (→C25)	d
C7	**r**	r	r~ɾ	ر	Voiced alveolar flap or trill similar to Spanish *r* in *pero.* A double *r* is similar *rr* in Spanish *perro.*	r
C8	**z**	z	z	ز ذ^	Like *z* in English *zone.* Used for bothز and ذ but in some literary words ذ may be pronounced /dh/, as in *dháki* (→C25)	z
C9	**s**	s	s	س ث^	Like *s* in English *sat.* Used for both [[s]] and [[v]] but in some literary wordsث may be pronounced /th/, as in *thaqáfe* (→C23)	s
C10a C10b	**3** ʿ	3	ʕ	ع	Voiced pharyngeal fricative, a guttural sound like /ḥ/ (→C32) but voiced (no English equivalent). Use /3/ instead of /ʿ/ for simplicity. /ʿ/ is more academic.	ʿ
C11	**f**	f	f	ف	Like *f* in English *fix*	f
C12	**v**	v	v	ف^	Like *v* in English *verse,* as in بْرُوبَة *próva.* Used only in some loanwords.	v
C13a C13b	ʾ **q**	' q	ʔ	ق	In urban speech as glottal stop [ʔ] but [q], [k] or [g] in some subdialects. Can replace /ʾ/ with /'/ for simplicity. /q/, also pronounced [ʔ], is more academic.	'
C14	**ӟ**	q:	q	ق	The strange looking /ӟ/, pronounced [q], distinguishes it from /q/ (→C13). This is a voiceless uvular plosive like a back [k] with tongue in contact with uvula. In urban speech usually pronounced [ʔ] but as [q] in some literary words such as *músiqa.*	ӟ

C15	**k**	k	k	ك	Like *k* English in *kangaroo*	k
C16	**l**	l	l	ل	Like *l* English in *learn* but sometimes velarized (→ 3.6)	l
C17	**m**	m	m	م	Like *m* English in *man* but sometimes velarized (→ 3.6)	m
C18	**n**	n	n	ن	Like *n* English in *nation*	n
C19	**h**	h	h	ه	Like *h* English in *he*	h
C20	□^h	h^	h	ه	Like /h/ but almost silent. After vowels at word end indicates the 3rd person singular pronoun and *lengthens* the previous vowel, as in <mark>wára</mark> → wará^h 'behind him'	h
C21	**w**	w	w	و	Like *w* in English *wall*	w
C22	**y**	y	j	ي	Like *y* English *yellow*	y
C23	**th**	th	θ	ث	/th/ in some literary words such as *thaqāfe* but usually replaced by /t/ or /s/, as in *míthl* → *mít^el* and *mathalan* → *masalan*	t̲
C24	**kh**	kh	x	خ	Voiceless velar fricative, as *ch* in German *ach* or Scottish *loch*	ḫ
C25	**dh**	dh	ð	ذ	/dh/ in some words of literary origin such as *dhíhni,* but usually replaced by /d/ or /z/, as in *dhahab* → *dahab* and *dhaki* → *zaki*	d̲
C26	**sh**	sh	ʃ	ش	like English /sh/ in *shoot*	š
C27	**gh**	gh	ɣ	غ	Voiced velar fricative, a gargling sound similar to Parisian [ʁ] (voiced uvular fricative), as in <mark>ghāli</mark> (no English equivalent)	ġ
C28	**ṣ**	S	sˤ	ص	Emphatic (velarized) consonant (→ 3.6)	ṣ
C29	**ḍ**	D	dˤ	ض	Emphatic (velarized) consonant (→3.6)	ḍ
C30	**ṭ**	T	tˤ	ط	Emphatic (velarized) consonant (→ 3.6)	ṭ
C31	**ẓ**	Z	ðˤ	ظ	Emphatic (velarized) consonant (→3.6)	ẓ
C32	**ḥ**	H	ħ	ح	Voiceless pharyngeal fricative, an emphatic /h/ produced by tightening the throat muscles (no English equivalent)	ḥ

6.2 Article assimilation

The rules for article assimilation for PA are similar to those of MSA. That is, before words beginning the sun letters *d, dh, ḍ, s, sh, ṣ, t, ṭ, z, ẓ, r* the *l* of *il*, the *l* is assimilated into the first letter of the following word. For example, *il-sínima* →*is-sínima*, *il-dárs* →*id-dárs*. In the case of moon letters, no assimilation occurs, e.g. *il-bét* remains *il-bét*. Unlike MSA, sometimes *j* is sometimes also assimilated, as in *il-jár* →*ij-jár*.

It should be noted that in the case the article prcedes a word beginning with two consonants the first of which is a sun letter, the *i* is dropped and the *l* is followed by the epenthetic vowel *e*, e.g. *il-ktāb → l ᵉktāb.*

6.3 Gemination

Both MSA and the dialects have the well known phenomenon of *geminate* or double consonants (also called long consonants), as in بدِّي *bíddi* 'I want', marked by the diacritic *shadda* over the د. They are pronounced like short consonants, but longer. Normally they are not shown in the Arabic script but are shown explicitly in DARS by doubling the cononant as in *bíddi*. For ease of reading, a hyphen is placed between diagraphs, as in fish-sh.

6.4 Degemination

Both MSA and the dialects have the well known phenomenon of *geminate* or double consonants (also called long consonants), as in بدِّي *bíddi* 'I want', marked by the diacritic *shadda* over the د. They are pronounced like short consonants, but longer. Normally they are not shown in the Arabic script but are shown explicitly in DARS by doubling the cononant as in *bíddi*. For ease of reading, a hyphen is placed between diagraphs, as in fish-sh.

It is noteworthy to note that sometimes geminated consonants are *degeminated* (undoubled) in actual pronunciation, as in *bíddi* becoming *bídi* or possibly *bídᵈi* and *ʾissalmak* becoming *ʾisalmak*. Naturally such subtelties cannot be represented in the Arabic script. A related phenomenon is when a normal non-pharyngeal consonant becomes a pharyngealized geminate, as in *rabaṭtu* realized as *rabaṭtu*. Another example is the compound verb like *3am-(b)tākol* 'you are eating', in which the b beomes weak or is omitted altogether.

Degeminated consonants are fairly common, and since they affect pronunciation is a rather clear manner it is best to show them in DARS by putting the second consonant in paretntheses (bíd(d)i) or as a superscript (*bíddi*). But pharyngealization is a hardly audible phonetic phenomenon that need not be shown since DARS is essentially a phonemic transcription.

7. Word Stress

7.1 The role of stress in PA

 Word stress, along with vowel neutralization and vowel length, are of great importance to learners. They are essential to achieving good sentence rhythm and to achieving natural speech. This is true for both MSA and PA, but is even more important to sounding natural in PA. This is a complex subject fully described in Halpern [3]. The stress rules for PA are simpler than for MSA. The DARS system represents the stress by an accent mark. The Arabic script does not (cannot) represent stress.

Elihay [x], author of *Speaking Arabic* (arguably the best book on PA) says about stress "Stress is very important indeed! ….misplaced stress can, in somce cases, change the meaning of the entire sentnece...It's woth while learning how to stress wods properly from the outset""

7.2 Phonotactics [E]

PA syllables have the following structure: an obligatory *onset*, consisting of one consonant (C) or two consonants (*CC*), an obligatory *nucleus,* consisting of a short vowel (*v*) or a long vowel (*vv*), and an optional *coda,* consisting of one or two consonants. The syllables can be classified into three types, referred to as **syllable weight**:

[Unify fonts to Dejavu?]
[what exact are the {…..} and [….] below

Light

Cv	*{[fa]-taḥ}*

Heavy

CvC	*{fa-[taḥ]}*
Cvv	*{[kā]-tib}*

Superheavy

CvvCC	*{shābb}*
CvCC	*{[burd]-qān}*, *{kabb}*
CvvC	*{[msab]-be}*
CCvv	<[mghaa]-ra]
CCvvC	*{ktāb}*, *{mnīḥ}*

MSA phonotacticts is described in Halpern [x]. It should be noted that some of the syllable patterns in PA do not exist in MSA because PA allows syllable initial consonant clusters like CCvC and CCvvC, which do not occur in MSA. It is important to understand syllable weight in order to determine the correct word stress.

7.3 Stress rules [E]

The rules for word stress in MSA are fairly complex. Most Arabic grammar books give stress rules that are inadequate or incomplete. Actually, MSA stress rules are fairly similar (almost identical) to PA stress rules. PA stress is treated in the excellent textbooks by Sharon [x] and Elihay [x], but the rules given there are not 100% complete. Below are stress rules for PA based on the MSA stress rules in Halpern [x] and Elihay [y]

[italicize DARS below?]
1. *Stress always falls on the last syllable if that syllable is superheavy.*
 This rule takes precedence over all others.

ja-dī́d	'new'	جديد
da-ra-jā́t	'degrees'	درجات
ka-tábt	'you wrote'	كتبت

2. *In monosyllabic words, stress falls on the last (the only) syllable.*
 (In DARS, the accent mark is not shown in such cases.)

ma̱	'what'	ما

bēt   'house'   بيت

3. *In disyllabic words, stress falls on the penultimate*
(unless the final syllable is superheavy → Rule #1)

    wá-lad   'child'   ولد

    {3ín-di}   'child'   عندي

    ʾí-za   if, whether   إذ

    ká-tab   he wrote   كتب

4. *In polysyllabic words, stress falls on the penultimate if that syllable is heavy.*

    ka-táb-na   'we wrote'   كتبنا

    ja̲-bát-li   'she brought me'   جابتلي

5. *In polysyllabic words, stress falls on the antepenultimate if the penultimate is light.*

    zá-la-me   man, guy   زلمة

    ká-ta-bu̲   they wrote   كتبو

6. *Long open vowels at the end of a word are never stressed.*
(In general, unstressed long vowels are always pronounced short or neutralized).

    ʾá-na̲   I   أنا

It is important to note the occurrence of stress shift For eample the stress in *wa/ra* 'behind' is on the first syllable, but attaching the pronomial enclitic *ak* causes the stress to shift to the last syllable in *warãk*

The above rules are rather complicated. Here are simpler versions with only two rules (but <mark>probably</mark> with some sacrifice in precision). According to Rice et al in their excllent book *Eastern Arabic* [x], there are two rules. He defines "long syllable" as a "short or long vowel followed by two or more consonants".

- Stress falls on the long syllable nearest the end (*{ḥālak}*, *{mabṣūt}*, *{shúkran}*)
- Otherwise the stress falls on the first syllable (*{ána}*)

Yoda [x] gives the following two rules

- Stress falls on the vvC or vCC closest to word end (*{ʾus-ta-zãt}*, *{yabãni}*)
- In all other cases, stress falls on the first syllable *{dá-ra-bu}*

7.4 Stressed vowels

Table 5. Stressed vowels
[E] but add examples

DARS	DARS+	IPA	Arabic	Description	Examples	Elihay

á	a/	ˈa	-	stressed short /a/		**a**
ɑ́	A/	ˈɑ	-	stressed short /ɑ/		**ɑ**
é	e/	ˈe	-	stressed short /e/		**e**
í	i/	ˈɪ, ˈi	-	stressed short /i/	Can both be stressed?	**i**
ó	o/	ˈʊ	-	stressed short /o/		**o**
ú	u/	ˈʊ, ˈu	-	stressed short /u/	Can both be stressed?	**u**
ā́	aa/	ˈaː	-	stressed long /a/		**ā**
ɑ̄́	AA/	ˈɑː	-	stressed long /ɑ/		**ā**
ḗ	ee/	ˈeː	-	stressed long /e/		**ē**
ī́	ii/	ˈiː	-	stressed long /i/		**ī**
ṓ	oo/	ˈoː	-	stressed long /o/		**ō**
ū́	uu/	ˈuː	-	stressed long /u/		**ū**

8. Using DARS

8.1 Example sentences

The sentences below are written in DARS. Under each sentence is the DARS+ version, an ASCII representation that can be converted to DARS using conversion tool. DARS+ is easy to type without special font or OS support. (The version in parentheses is a more academically oriented alternative.)

[Reiew from acad/pedag point of view]

No	Palestinian Arabic	English
1	‏إمي إسا موجودة في أميركا.‏ **ʾímmi ʾíssa mawjū̆de fi ʾamḗrka.** 'i/mmi 'i/ssa mawjuu/de fi 'amee/rka.	My mother is in America at the moment.
2	‏كتبت مكتوب من هناك قبل جمعة.‏ **kátbat maktū̆b min hunā́k ʾábᵉl** (qábᵉl) **júm3a** (júmˈa). ka/tbat maktuu/b min hunaa/k 'a/be^l (qa/be^l) ju/m3a.	She wrote a letter from there a week ago.
	‏خدنا المكتوب اليوم الصبح.‏	And we got the letter this

20

	u-akhád-na_l-maktū́b_il-yṓm iṣ-ṣúbᵒḥ. u-akha/d-na_l-maktuu/b_il-yoo/m iS-Su/bo^H.	morning.
	أ	
4	.أنا كمان كتبت مكتوب **ʾána kamā́n katábᵉt maktū́b.** 'a/na kamaa/n kata/be^t maktuu/b.	I've written a letter too.
5	كيف كتبت المكتوب، عاى ماكينة ولا بالقلم؟ **kīf katábt_il-maktū́b, 3al-** (ʾal-) **mā́kina wílla bil-ʾálam** (qálam)**?** kiif kata/bt_il-maktuu/b, 3al-maa/kina wi/lla bil-'alam (qa/lam)?	How did your write the letter, on the typewriter or with a pen?
6	.بالقلم، ما عنديش ماكينة **b-il-ʾálam** (qálam)**, ma 3indī́-sh** (ʾindī́-sh) **mā́kina.** b-il-'a/lam (qa/lam), ma 3indii/-sh maa/kina.	With a pen. I haven't got a typewriter.

8.2 Two orientations

DARS is suitable for both pedagogical purposes and for linguistically analysis.. We refer to the first as **pedagogically oriented** and to second as **academically oriented.** To this end there are some alternative symbols for some phonemes, and some feature are optional. It is possible to to mix the two to various degrees.

Academically oriented
This is more difficult to read and write but more precise and better for linguistic analysis.

- ع is written as /ʿ/, common in academic transcriptions
- ق is written as /q/, even when pronounced as the glottal stop
- Neutralized vowels are shown by underlining
- Word stress is shown by the acute accent

Arabic	كيف كتبت المكتوب إلى يابان، عاى ماكينة ولا بالقلم؟
DARS+	kiif kata/bt il-maktuu/b 'ila_-l-ya_bAA/n, Eal-maa/kina wi/lla Bil<-qa/lam?
DARS	*kīf katábt il-maktū́b ʾilạ-l-yạbā́n, ʾal-mā́kina wílla bil-qálam?*

Pedagogically oriented
This enables quick typing and results in a cleaner transcription. It is much easier to convert DARS+ to DARS with our conversion tool than to type it directly.

- ع is written as /3/, more popular and simpler
- ق is written as ʾ when pronounced as the glottal stop

- Omit underlines for neutralized vowels
- Omit stress accents (though desirable to keep)

Arabic	كيف كتبت المكتوب إلى يابان، عاى ماكينة ولا بالقلم؟
DARS+	kiif katabt il-maktuub 'ila-l-yabAAn, 3al-maakina willa bil-'alam
DARS	*kīf katabt il-maktūb ʾila-l-yabān, 3al-mākina willa bil-ʾalam?*

8.3 CARS sample sentences

For reference, below are some MSA sentences transcribed in CARS, the system DARS is based on.

Sample Text for CARS MSA

Vocalized Arabic	مُنْذُ مِئَاتِ ٱلسِّنِينَ، كَانَتْ تَعِيشُ فِي ٱلصِّينِ أَرْمَلَةٌ فَقِيرَةٌ مَعَ ٱبْنِهَا ٱلصَّغِيرِ عَلَاءِ ٱلدِّينِ.
Standard CARS	múndhu miʾắti‿ssinína, kắnat taɛíshu fi‿ṣṣíni ʾarmálatun faqíratun máɛa‿bniha‿ṣṣaghíri ɛalắʾi‿ddíni.
Proxy CARS	mu/ndhu miʾaa/ti~ssinii/na, kaa/nat taEii/shu fi_~SSii/ni ʾarma/latun faqii/ratun ma/Ea~bniha_~SSaghii/ri Ealaa/ʾi~ddii/ni.
Extended CARS	múndhu miʾắti‿ssinín|a, kắnat taɛíshu fi‿ṣṣín|i ʾarmála|tun fɒqíra|tun máɛa‿bniha‿ṣṣɒghír|i ɛalắʾi‿ddín|i.
English	Hundreds of years ago, there lived in China a poor widow with her young son, Aladdin.

9. Conclusions

References

[This section needs to be organized and edited]

J. Elihay, *The Olive Tree Dictionary* (Minerva Instruction & Consultation, 2004)

Jack Halpern, *Word Stress and Vowel Neutralization in Modern Standard Arabic*
https://www.cjki.org/arabic/an_paper.pdf

Bruno Herin, *Do Jordanians really speak like Palestinians?*, Journal of Arabic and Islamic Studies (2017)
https://www.lancaster.ac.uk/jais/volume/docs/vol13/v13_06_herin_99-114.pdf

Maria Gouskova, Nancy Hall, *Levantine Arabic Vowel Epenthesis: Phonetics, Phonology, and Learning* (2007)
https://web.stanford.edu/dept/linguistics/linginst/nsf-workshop/Gouskova&Hall_abstract.pdf

Nizar Habash, Mustafa Jarrar, Faeq Alrimawi, Diyam Akra, Nasser Zalmout, Eric Bartolotti, and Mahdi Arar, *Palestinian Arabic Conventional Orthography Guidelines - Technical Report*
https://core.ac.uk/download/pdf/86433442.pdf

https://www.palestinianarabic.com/

Detailed phonemes
　　　　https://www.kau.edu.sa/Files/0001063/files/2826_phon5.pdf

https://www.palestinianarabic.com/coursebooks.html

EASTERN ARABIC
Frank A. Rice and Majed F. Sa'id
GUP, 2011

Palestinian Arabic Verbs: Conjugation Tables and Grammar
Matthew Aldrich, Ahmed Younis
Lingualism, 2021

https://www.palestinianarabic.com/pronunciation.html

Bernie, *Word stress in Levantine Arabic*, "From Fuṣḥā to ʿĀmmiyya" blog post (2012)
　　　　https://learnammiyya.wordpress.com/2012/10/05/word-stress-in-levantine-arabic/

Issam M.Abu-Salim, Vowel shortening in Palestinian Arabic: A metrical perspective, ScienceDirect (1986)
https://doi.org/10.1016/0024-3841(86)90005-7

See EasternArabic bottom of p xxxiv
See https://en.wikipedia.org/wiki/Levantine_Arabic_phonology#Vowels_and_diphthongs

Levantine Arabic vowel epenthesis: phonetics, phonology, and learning
(use for epenthesis, get book)
chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fweb.stanford.edu%2Fdept%2Flinguistics%2Flinginst%2Fnsf-workshop%2FGouskova%26Hall_abstract.pdf&clen=113303&chunk=true

REFERENCE AND RESEARCH

Study DARS format
Major issue with *{a}*. See PA_fonts.ods

خدنا المكتوب اليوم الصبح.

As sentence
u-akhád-na_l-maktū̆b il-yṓm iṣ-ṣúbᵒḥ.
u-akhád-na_l-maktū̆b il-yṓm iṣ-ṣúbᵒḥ.
/u-akhád-na_l-maktū̆b il-yṓm iṣ-ṣúbᵒḥ./
<u-akhád-na_l-maktū̆b il-yṓm iṣ-ṣúbᵒḥ.>
u-akhád-na_l-maktū̆b il-yṓm iṣ-ṣúbᵒḥ.

[M-san a +0061 ɑ +0251 look identical in italics. What to do. Use other font but ok in Arial Unicode
[M-san *a* +0061 *ɑ* +0251 look identical in italics. What to do. Use other font
[ắ doesn't work in Arial Unicode]
[above paragraph is experiment in Arial Unicode MS. Charsil best?]
Conclusion use Deja vu sans 9.5 points but A/ seems to scrw up

Embedded
1. DARS il-yṓm iṣ-ṣúbᵒḥ embedded in English
2. DARS il-yṓm iṣ-ṣúbᵒḥ embedded in English
3. DARS *il-yṓm iṣ-ṣúbᵒḥ* embedded in English
4. DARS **il-yṓm iṣ-ṣúbᵒḥ** embedded in English
5. DARS /il-yṓm iṣ-ṣúbᵒḥ/ embedded in English

[3. was selected as final by M-san's recommendation]

DARS＋	IPA	Example	Phonemic Description	Phonetic description
eˆ	ə̰ ĕ ə	lᵉktā̆b (اَلْكْتَاب)	extra short epenthetic /e/	mid-central extra short vowel tending to [ə] OR Extra short close-mid front unrounded vowel [ĕ]
oˆ	ə~ɵ ŏ	shugᵒl	extra short epenthetic /o/	mid-central extra short vowel tending to [o] OR

	ə ɘ θ	(شُغْل)		extra short close-mid back rounded vow [ŏ]

epenthetic /e/

ə	U+0259
ĕ	U+0015
ɘ	U+0258

epenthetic /o/

ə ~ θ	U+0259 ~ U+0275
ŏ	U+014F
ə	U+0259
ɘ	U+0258
θ	U+0275

From wiki

In phonology, epenthesis (/ɪˈpɛnθəsɪs, ɛ-/; Greek ἐπένθεσις) means the addition of one or more sounds to a word, especially to the interior of a word (at the beginning prothesis and at the end paragoge are commonly used). The word *epenthesis* comes from epi- "in addition to" and *en-* "in" and *thesis* "putting". Epenthesis may be divided into two types: excrescence for the addition of a consonant, and for the addition of a vowel, svarabhakti (in Hindi, Bengali and other North Indian languages, stemming from Sanskrit) or alternatively anaptyxis (/ˌænəpˈtɪksɪs/). The opposite process, where one or more sounds are removed, is referred to as elision.

NOTES

0. Which is best for representing hamza /'/ that is originally /q/\

<table>
<tr>
<td>

DARS

هناك قبل جمعة

1 hunā́k qábᵉl júmɛa.
2 hunā́k ʾábeᵉl júmɛa.
3 hunā́k �q́ábᵉl júmɛa.

</td>
<td>

Use q^ for expressin hamza that was originally /q/?
DARS+
1. qa/be^l
2. 'a/be^l
3. q^a/be^l

</td>
</tr>
</table>

1. CODA: /a:, i:, u:, e:, o:/ => (aA, iy, uw, ay, aw) (-) - وَ- , يَ- , يِ ُو , ا /aw, ay/ => (aw, ay) (-) - يَ وَ This is the same as MSA with the exception