

# **CJKI Arabic Romanization System CARS**

**by Jack Halpern**

**Presented by Dr. Aziz Khalil**

# Transcription and Transliteration

**Transliteration:** representing the source script letters (graphemes *not* phonemes) with the characters of another script: محمد → \mHmd\

**Transcription:** representing the source script of a language in the target script in a manner that reflects pronunciation. This includes:

**Phonetic:** muħëmmëd (IPA)

**Phonemic:** /muHammad

**Popular:** *Mohammed, Muhammad, Moohammad, Moohamad...+200*

# CARS Example

الْحُكُومَةُ الْيَابَانِيَّةُ

*'alḥukūmatu lyabānīyyatu*

*the Japanese government*

1. The **Standard CARS** symbol set aims to (1) unambiguously represent the Arabic phonemes or *phonemicization*, (2) accurately indicate word stress or *accentuation*, and (3) explicitly indicate which vowels are *neutralized*.

2. **Extended CARS** has three optional features: the indication of case endings or *declension*, the indication of velarized /a/ or *velarization*, and the indication of syllable boundaries or *syllabification*.

3. **Proxy CARS** is a set of substitute symbols to facilitate input from a standard ASCII keyboard.

# Standard CARS Symbols

Symbol	Unicode	Name	Proxy Symbol	Description
ˉ	U+0304	macron	vv	The <b>macron</b> , as in <i>ā ī ū</i> ( <i>aa ii uu</i> ), indicates a long vowel, as in سَافَرْتُ <i>sāfartu</i> .
´	U+0301	acute accent	/	The <b>acute accent</b> , as in <i>á í ú</i> , ( <i>a/ i/ u/</i> ) represents a stressed syllable, as in أَنَا <i>'ána</i> .
ˆ	U+0304 + U+0301	macron + acute accent	vv/	Indicates a long stressed vowel, as in اَلصَّرَافُ <i>'aṣṣarrāfu</i> .
˘	U+0331	macron below	x_	Represents neutralization, as in أَنَا <i>ána</i> ,
·		underdot	caps	Modifies certain consonants to represent others, i.e., ص <i>ṣ</i> , ض <i>ḍ</i> , ط <i>ṭ</i> , ظ <i>ẓ</i> and ح <i>ḥ</i> .
◌ِ	U+203F	undertie	~	Liaison between words resulting from the omission of 'alif waṣla (i) word initially, as in فِي الصَّيْنِ <i>fī ṣṣīni</i> .
'	U+0027	apostrophe	none	<i>Hamza</i> in all its written forms, i.e. 'a, 'i and 'u.
ε	U+0025	epsilon	E	The Greek letter <b>epsilon</b> represents the letter <i>ḥayin</i> (ع), as in سَعِيدٍ <i>saʿīd</i> .
◌ْ	U+02D8	breve	^	Disjoins the letters <i>t, s, d</i> and <i>k</i> from a following <i>h</i> to show that they are distinct letters ( <i>tḥ</i> تھ)

# Vowels and Diphthongs

## Short Vowels

CARS	Proxy	Description
a	a	short /a/
i	i	short /i/
u	u	short /u/
á	a/	stressed short /a/
í	i/	stressed short /i/
ú	u/	stressed short /u/

## Long Vowels

CARS	Proxy	Description
ā	aa	long /a/
ī	ii	long /i/
ū	uu	long /u/
ǎ	aa/	stressed long /a/
ǐ	ii/	stressed long /i/
ǔ	uu/	stressed long /u/

Short vowels are simply represented by *a*, *i* and *u*, as in كُتِبَ *kútiba* ‘was written’. Long vowels are represented by placing a macron over the vowel, as for example in الصَّرَافُ *’aṣṣarrāfu* and بُيُوتٌ *buyūtun*.

Nunation, *tā' marbūṭa* (ة), 'alif maqṣūra (ى) and various other special vowels are represented strictly as they are pronounced, as in *بُيُوتٌ buyūṭun*, *عَرَبِيَّةٌ ʿarabīyya* and *عَلَى ʿala*, with no attempt to indicate the original orthography.



# Plain Consonants

ب	b	bā'
ت	t	tā'
ج	j	jīm
د	d	dāl
ر	r	rā'
ز	z	zāy
س	s	sīn
ع	ε	εayin
ف	f	fā'

ق	q	qāf
ك	k	kāf
ل	l	lām
م	m	mīm
ن	n	nūn
ه	h	hā'
و	w	wāw
ي	y	yā'
ء	'	hamza

# Hamza, ʿayin and Diagraphs

	Arabic	CARS	Others
Assimilation	الصَّغِيرُ	'aṣṣaghîr	al-saghiir
Hamza	أَخَذَ	'ákhadha	akhadha
ʿayin	سَعِيد	saʿīd	sa`īd
Shadda	مُحَمَّد	muḥammad	muḥammad
Diagraphs	أَسْهَل	'ás`hal	ashal

## Short Stressed Vowels

CARS	Proxy	Description	Example
á	a/	stressed short /a/	<i>kátaba</i>
í	i/	stressed short /i/	<i>kalímati</i>
ú	u/	stressed short /u/	<i>kútab</i>

## Long Stressed Vowels

CARS	Proxy	Description	Example
â	aa/	stressed short /a/	<i>kâtib</i>
î	ii/	stressed short /i/	<i>jadîd</i>
û	uu/	stressed short /u/	<i>yaḅaniyũna</i>

# Neutralized Vowels

CARS	Description	Example
<u>a</u>	neutralized /a/	<i>hãdh<u>a</u></i>
<u>i</u>	neutralized /i/	<i>i<u>i</u></i>
<u>u</u>	neutralized /u/	<i>u<u>u</u></i>
<u>ɔ</u>	neutralized velarized /a/	<i>A<u>A</u></i>

# Undertie Example

في الصِّينِ *fi ṣṣīni* 'in China'

# Extended CARS Symbols

Symbol	Unicode	Symbol Name	Proxy Symbol	Description
	U+007C	vertical bar	()	Case endings, as in الصَّرَافُ <i>aṣṣarrāfu</i>
ɒ	U+0252	turned alpha	A	Represents a velarized /a/, as in الصَّرَافُ <i>vṣṣarrɒfu</i> .
.	U+00B7	middle dot	-	Represents a velarized /a/, as in الصَّرَافُ <i>vṣṣarrɒfu</i> .

# Extended CARS Example

الصَّرَافُ الْآلِيُّ 'vʃ·ʃpr·rɔ̃·fu\_l·'ā·líy·y|u ATM

# Proxy Symbols

Symbol Example	Symbol Name	Proxy Example	Function
$\bar{a}$	macron	$aa$	long vowels
$\acute{a}$	acute accent	$a/$	stressed vowels
$\tilde{a}$	macron + acute accent	$aa/$	long stressed vowels
$\underline{a}$	macron below	$a_$	neutralization
$\text{\textasciix}$	underdot	$S$	emphatics etc.
$a_b$	undertie	$a\sim b$	liaison
$d^h$	breve	$d^h$	digraph disjunction
$\upsilon$	turned alpha	$A$	velarization
$\epsilon$	epsilon	$E$	$\epsilon ayin$
$la\cdot ka$	middle dot	$la-ka$	syllabification



# Proxy CARS Example

Arabic:	أَخَذَ عَلَاءُ الدِّينِ الْقِرْدَ الصَّغِيرَ
English:	Aladdin took the little monkey
Standard CARS:	<i>'ákhadha ealā'u ddīni lqírda ṣṣaghîra</i>
Proxy CARS:	<i>'a/khadha Ealaa/'u~ddii/ni~lqi/rda~SSaghii/ra</i>

# CARS Tools

- 1.CARS IME:** This is an input method editor that converts proxy symbols to CARS symbols in real time, e.g. *'a/khadha Ealaa/'u~ddii/ni* is converted to *'ákhadha εalā'u ddīni*.
- 2.CARS Converter:** This batch conversion utility accepts text files written in proxy symbols as input and converts them to CARS symbols encoded in UTF-8. It can also convert CARS symbols to proxy symbols.
- 3.CARS Macros:** Macros for Microsoft Word and Excel, as well as for Open Office Writer and Calc, can convert from CARS symbols to proxy symbols and vice versa.
- 4.A2C Converter:** Converts fully vocalized Arabict CARS symbols or proxy symbols in batch mode. This powerful tool automatically determines the stressed and neutralized syllables.

# Sample Text

## Vocalized Arabic

مُنْذُ مِئَاتِ السِّنِينَ، كَانَتْ تَعِيشُ فِي الصِّينِ أَرْمَلَةً فَقِيرَةً مَعَ ابْنِهَا الصَّغِيرِ عَلَاءِ الدِّينِ. كَانَ عَلَاءُ الدِّينِ يَبْلُغُ  
مِنَ الْعُمُرِ اثْنَيْ عَشَرَ عَامًا فَقَطْ، وَلَكِنَّهُ كَانَ يَعْمَلُ فِي مَحَلِّ خِيَّاطٍ لِيَكْتَسِبَ مَا يَعِيشُ مِنْهُ هُوَ وَأُمَّهُ.

## Standard CARS

múndhu mi'āti ssinína, kánat taēīshu fī ṣṣīni 'armálatun faqíratun  
máea bnihā ṣṣaghîri ealā'i ddīni. kána ealā'u ddīni yáblughu mína leúmri 'ithnáy  
eáshara eáman fáqaṭ, walakínnahu kána yáemalu fī maḥállī khayyāṭin liyaktásiba  
ma yaēīshu mínhu húwa wa'úmmahu.

## Extended CARS (without syllabification)

múndhu mi'āti ssinín|a, kánat taēīshu fī ṣṣīn|i 'armála|tun foqíra|tun  
máea bnihā ṣṣoghîr|i ealā'i ddín|i. kána ealā'u ddín|i yáblughu mína leúmr|i  
'ithnáy eáshara eám|an fóqṭ, walakínnahu kána yáemalu fī maḥáll|i khayyōṭ|in  
liyaktásiba ma yaēīshu mínhu húwa wa'úmmahu.

شُكْرًا جَزِيلًا

The CJK Dictionary Institute, Inc.  
<http://www.cjk.org>