

Bustani: an introduction

The Bustani™ typeface emulates Naskh, one of the classical calligraphic styles of Arabic script. Naskh has long been considered the primary style for literary text. Naskh has a very rich set of alternative letter shapes with which a calligrapher can choose compose words and lines of text. In a similar manner, the internal intelligence of Bustani chooses the appropriate letter forms for a given text according to a complex set of rules.

Besides entering the text, the user need not make any decisions to get well-formed, Naskh-style words. In some contexts, and depending on the particular application, the user can choose alternative letter forms. In any sentence rendered in Bustani, we can observe that the words are set on a slight, left-leaning incline in a manner similar to a calligrapher's writing.

Unlike many other Arabic fonts, the Bustani™ typeface is designed to handle fully vocalized text (i.e. containing many marks—Arabic harakāt) reliably and gracefully as in the following sample:

عَنْ كُلِّ مَا عَرَفْتُهُ

As the above sample demonstrates, the vowel marks (harakāt) follow the contour of the words and are precisely placed.

Not only are the built-in capabilities of Bustani more than adequate for everyday Arabic text, they also meet the requirements of scholarly manuscripts.

Why a User Guide?

Once the Bustani font is installed, it can be used immediately like any other Arabic font. Since some of the features supported by Bustani are not immediately obvious and require some explanation. This brief User Guide aims to explain these features. You can continue reading each section in order, or scan ahead to a specific section that might interest you. Here are quick links to the topics covered:

[Character Set](#)

[Stylistic Sets](#)

[Numeric Notation](#)

[Language-specific Features](#)

[Quranic Marks](#)

[Transliteration/Romanization](#)

[Honorific Marks](#)

Where to look if you...?

Would like to know whether Bustani supports a particular character such as **ث** (u+06A4), have a look at the [Character Set](#)

Would you like to create a fraction such as $\frac{40}{12}$ or $\frac{12}{45}$ or how to write **٣١٢**, consult the section on [Numeric Notation](#).

Want to change **صخر** into **صخر**, or **كاف** into **كاف**, or **٢** into **٢**, visit the [Stylistic Sets](#) section. Bustani offers 17 ways to change the visual representation of text without disturbing the underlying text.

Want to write **صخر** as **ṣaḥr**, go read about [Transliteration](#).

Are wondering why the quotation marks in **«اهلاً»** look different from those in **«Hello»**, have a look at [Language-Specific Features](#).

Want to use symbols such as **ٲ**, **ٲ**, or **ٲ**, explore the details of [Quranic Marks](#).

Want to write something such as **خالد**, explore [Honorific Marks](#).

Would like to see various pieces of text composed in Bustani

[Universal Declaration of Human Rights \(an extract\) Arabic](#)

[Universal Declaration of Human Rights \(an extract\) Farsi](#)

[Poem by Nizār al-Qabbāni](#)

[Poem by Hāfez \(Farsi\)](#)

[Al-Ayyām by Tāha Ḥussein \(an extract\)](#)

[Sūrat al Qiyāma](#)

[A'māl al Rusul \(an extract\)](#)

[Qaṣīda by al al-Andalusi](#)

[Discourse by Qess ibn Sa'da](#)

Comments? Write to us at bustani@monotype.com

Character Set

Bustani's character set covers a broad range of languages including Classical and Modern Arabic, in addition to Farsi, Urdu, Sindhi, Panjabi, Pashto, Kashmiri, Jawi/Malay, and Kurdish, Uyghur, among others.

More precisely, Bustani's character set conforms to the Unicode Standard and supports all the Arabic-script characters in the Standard, Version 6.0, except for some minor marks and symbols.[‡] To see the full chart of characters, along with their corresponding Unicode values, click [here](#).

The character set consists primarily of alphabetic letters, combining marks such as vowels and other aids in pronunciation, numeric figures, punctuation marks, along with a range of other symbols. In addition to the Arabic-script repertoire, Bustani also includes an ample Roman-script repertoire of the Sabon™ typeface that supports all major European languages, as well as scholarly characters for transliterating Arabic text. The Latin-character repertoire can be seen [here](#), while samples of transliteration are shown on this page.

Note on the Tatweel character (u+0640): while this character is commonly used to stretch the width of an alphabetic letter, Bustani treats it as a neutral character that has no effect on width.

[‡]The following characters are not included: u+0606, u+0607, u+0608, u+0616, u+0617, u+0618, u+0619, u+061A

u+0600	ا	u+060F	ع	u+0620	ي	u+062A	ت	u+0634	ش
u+0601	س	u+0610	هـ	u+0621	ء	u+062B	ث	u+0635	ص
u+0602	م	u+0611	و	u+0622	آ	u+062C	ج	u+0636	ض
u+0603	ط	u+0612	ة	u+0623	أ	u+062D	ح	u+0637	ظ
u+0609	٪	u+0613	ة	u+0624	ؤ	u+062E	خ	u+0638	ظ
u+060A	٪...	u+0614	ة	u+0625	إ	u+062F	د	u+0639	ع
u+060B	ف	u+0615	ة	u+0626	ئ	u+0630	ذ	u+063A	غ
u+060C	،	u+061B	؛	u+0627	ا	u+0631	ر	u+063B	تي
u+060D	،	u+061E	؛	u+0628	ب	u+0632	ز	u+063C	كي
u+060E	هـ	u+061F	؟	u+0629	ة	u+0633	س	u+063D	ئ



u+063E	ی	u+0657	َ	u+066F	ف	u+0687	چ	u+069F	ظ
u+063F	ی	u+0658	َ	u+0670	ا	u+0688	ٹ	u+06A0	ع
u+0641	ف	u+0659	َ	u+0671	ا	u+0689	د	u+06A1	و
u+0642	ق	u+065A	َ	u+0672	ا	u+068A	د	u+06A2	ب
u+0643	ک	u+065B	َ	u+0673	ا	u+068B	ط	u+06A3	ف
u+0644	ل	u+065C	َ	u+0674	ا	u+068C	ت	u+06A4	ق
u+0645	م	u+065D	َ	u+0675	ا	u+068D	د	u+06A5	پ
u+0646	ن	u+065E	َ	u+0676	و	u+068E	ث	u+06A6	ق
u+0647	ہ	u+065F	َ	u+0677	و	u+068F	ت	u+06A7	ف
u+0648	و	u+0660	َ	u+0678	ئی	u+0690	ت	u+06A8	ق
u+0649	ی	u+0661	ا	u+0679	ٹ	u+0691	ڑ	u+06A9	کی
u+064A	ی	u+0662	ا	u+067A	ن	u+0692	ز	u+06AA	کا
u+064B	َ	u+0663	ا	u+067B	ب	u+0693	ر	u+06AB	می
u+064C	َ	u+0664	ا	u+067C	ت	u+0694	ر	u+06AC	ن
u+064D	َ	u+0665	ا	u+067D	ت	u+0695	ز	u+06AD	ن
u+064E	َ	u+0666	ا	u+067E	پ	u+0696	ر	u+06AE	پ
u+064F	َ	u+0667	ا	u+067F	ت	u+0697	ز	u+06AF	گی
u+0650	َ	u+0668	ا	u+0680	پ	u+0698	ز	u+06B0	گی
u+0651	َ	u+0669	ا	u+0681	خ	u+0699	ز	u+06B1	تی
u+0652	َ	u+066A	ا	u+0682	خ	u+069A	نن	u+06B2	گی
u+0653	ا	u+066B	ا	u+0683	چ	u+069B	پی	u+06B3	بی
u+0654	َ	u+066C	ا	u+0684	چ	u+069C	ش	u+06B4	شی
u+0655	َ	u+066D	ا	u+0685	خ	u+069D	پی	u+06B5	ن
u+0656	َ	u+066E	ا	u+0686	چ	u+069E	شی	u+06B6	ن

u+06B7	ش	u+06CF	ف	u+06E7	َ	u+06FF	ھ	u+0767	ن
u+06B8	پ	u+06D0	پی	u+06E8	ِ	u+0750	ی	u+0768	ن
u+06B9	ب	u+06D1	پی	u+06E9	ا	u+0751	ب	u+0769	ن
u+06BA	ر	u+06D2	ل	u+06EA	و	u+0752	پ	u+076A	ل
u+06BB	ر	u+06D3	ل	u+06EB	و	u+0753	پ	u+076B	ر
u+06BC	ن	u+06D4	ا	u+06EC	و	u+0754	ن	u+076C	ر
u+06BD	ش	u+06D5	ہ	u+06ED	و	u+0755	ر	u+076D	ش
u+06BE	ھ	u+06D6	ا	u+06EE	د	u+0756	ر	u+076E	ح
u+06BF	چ	u+06D7	ا	u+06EF	ر	u+0757	خ	u+076F	چ
u+06C0	ہ	u+06D8	و	u+06F0	و	u+0758	چ	u+0770	ش
u+06C1	ہ	u+06D9	ا	u+06F1	ا	u+0759	پ	u+0771	ر
u+06C2	ہ	u+06DA	ا	u+06F2	ب	u+075A	ر	u+0772	ط
u+06C3	ة	u+06DB	و	u+06F3	ب	u+075B	ر	u+0773	ا
u+06C4	و	u+06DC	و	u+06F4	۴/۲	u+075C	ش	u+0774	ا
u+06C5	و	u+06DD	ا	u+06F5	۵	u+075D	غ	u+0775	ی
u+06C6	و	u+06DE	ا	u+06F6	۶/۶	u+075E	غ	u+0776	ی
u+06C7	و	u+06DF	و	u+06F7	۷/۷	u+075F	غ	u+0777	ی
u+06C8	و	u+06E0	و	u+06F8	۸	u+0760	ف	u+0778	و
u+06C9	و	u+06E1	و	u+06F9	۹	u+0761	ف	u+0779	و
u+06CA	ق	u+06E2	و	u+06FA	ش	u+0762	نی	u+077A	مے
u+06CB	ق	u+06E3	و	u+06FB	ض	u+0763	شی	u+077B	مے
u+06CC	ی	u+06E4	و	u+06FC	غ	u+0764	کی	u+077C	ح
u+06CD	ی	u+06E5	و	u+06FD	۴	u+0765	م	u+077D	س
u+06CE	ی	u+06E6	ے	u+06FE	م	u+0766	م	u+077E	س

u+077F	ق	u+FDF1	قل	u+002B	+ / +	u+0043	C	u+005B	[/ [
u+08F0	َ	u+FDF2	الله	u+002C	,	u+0044	D	u+005C	\
u+08F1	و	u+FDF3	اكبر	u+002D	- / -	u+0045	E	u+005D] /]
u+08F2	ِ	u+FDF4	مُحَمَّد	u+002E	. / .	u+0046	F	u+005E	^
u+08F3	ُ	u+FDF5	صلعم	u+002F	/	u+0047	G	u+005F	_
u+FBB2	.	u+FDF6	رسول	u+0030	0	u+0048	H	u+0060	`
u+FBB3	.	u+FDF7	عليه	u+0031	1	u+0049	I	u+0061	a
u+FBB4	..	u+FDF8	وسلم	u+0032	2	u+004A	J	u+0062	b
u+FBB5	..	u+FDF9	صلى	u+0033	3	u+004B	K	u+0063	c
u+FBB6	..	u+FDFA	صَلَّى اللّٰهُ عَلَيْهِ وَسَلَّمَ	u+0034	4	u+004C	L	u+0064	d
u+FBB7	..	u+FDFB	جَلَّالَهُ	u+0035	5	u+004D	M	u+0065	e
u+FBB8	..	u+FDFC	يَاي	u+0036	6	u+004E	N	u+0066	f
u+FBB9	..	u+FDFD	بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ	u+0037	7	u+004F	O	u+0067	g
u+FBBA	..	u+0020		u+0038	8	u+0050	P	u+0068	h
u+FBBB	..	u+0021	! / !	u+0039	9	u+0051	Q	u+0069	i
u+FBBC	..	u+0022	"	u+003A	: / :	u+0052	R	u+006A	j
u+FBBD	..	u+0023	#	u+003B	;	u+0053	S	u+006B	k
u+FBBE	..	u+0024	\$	u+003C	<	u+0054	T	u+006C	l
u+FBBF	..	u+0025	%	u+003D	= / =	u+0055	U	u+006D	m
u+FBC0	ط	u+0026	&	u+003E	>	u+0056	V	u+006E	n
u+FBC1	ط	u+0027	'	u+003F	?	u+0057	W	u+006F	o
u+FD3E	ط	u+0028	(/ (u+0040	@	u+0058	X	u+0070	p
u+FD3F	ط	u+0029) / (u+0041	A	u+0059	Y	u+0071	q
u+FDF0	ط	u+002A	* / *	u+0042	B	u+005A	Z	u+0072	r

u+0073	s	u+00AD		u+00C5	Å	u+00DD	Ý	u+00F5	õ
u+0074	t	u+00AE	®	u+00C6	Æ	u+00DE	Þ	u+00F6	ö
u+0075	u	u+00AF	ˉ	u+00C7	Ç	u+00DF	ß	u+00F7	÷ / ÷
u+0076	v	u+00B0	°	u+00C8	È	u+00E0	à	u+00F8	ø
u+0077	w	u+00B1	±	u+00C9	É	u+00E1	á	u+00F9	ù
u+0078	x	u+00B2	²	u+00CA	Ê	u+00E2	â	u+00FA	ú
u+0079	y	u+00B3	³	u+00CB	Ë	u+00E3	ã	u+00FB	û
u+007A	z	u+00B4	´	u+00CC	Ì	u+00E4	ä	u+00FC	ü
u+007B	{	u+00B5	µ	u+00CD	Í	u+00E5	å	u+00FD	ý
u+007C		u+00B6	¶	u+00CE	Î	u+00E6	æ	u+00FE	þ
u+007D	}	u+00B7	· / ·	u+00CF	Ï	u+00E7	ç	u+00FF	ÿ
u+007E	~	u+00B8	,	u+00D0	Ð	u+00E8	è	u+0100	Ā
u+00A1	¡	u+00B9	¹	u+00D1	Ñ	u+00E9	é	u+0101	ā
u+00A2	¢	u+00BA	º	u+00D2	Ò	u+00EA	ê	u+0112	Ē
u+00A3	£	u+00BB	» / »	u+00D3	Ó	u+00EB	ë	u+0113	ē
u+00A4	¤	u+00BC	¼	u+00D4	Ô	u+00EC	ì	u+011E	Ĝ
u+00A5	¥	u+00BD	½	u+00D5	Õ	u+00ED	í	u+011F	ğ
u+00A6	¦	u+00BE	¾	u+00D6	Ö	u+00EE	î	u+0120	Ġ
u+00A7	§	u+00BF	¿	u+00D7	× / ×	u+00EF	ï	u+0121	ġ
u+00A8	¨	u+00C0	À	u+00D8	Ø	u+00F0	ð	u+0126	Ĥ
u+00A9	©	u+00C1	Á	u+00D9	Ù	u+00F1	ñ	u+0127	ĥ
u+00AA	ª	u+00C2	Â	u+00DA	Ú	u+00F2	ò	u+012A	Ī
u+00AB	« / «	u+00C3	Ã	u+00DB	Û	u+00F3	ó	u+012B	ī
u+00AC	¬	u+00C4	Ä	u+00DC	Ü	u+00F4	ô	u+0130	İ

u+0131	ı	u+01E6	Ğ	u+0304	-	u+1E6C	Ṭ	u+201E	„
u+0141	Ł	u+01E7	ğ	u+0306	˘	u+1E6D	ṭ	u+2020	†
u+0142	ł	u+01F4	Ġ	u+0307	˙	u+1E6E	Ṯ	u+2021	‡
u+014C	Ō	u+01F5	ġ	u+0308	¨	u+1E6F	ṯ	u+2022	•
u+014D	ō	u+0218	Ş	u+030C	ˇ	u+1E92	Ẑ	u+2025	..
u+0152	Œ	u+0219	ş	u+0323	.	u+1E93	ẑ	u+2026	.../...
u+0153	œ	u+021A	Ṭ	u+0326	,	u+1E96	ḥ	u+2030	‰
u+015E	Ş	u+021B	ṭ	u+0327	,	u+1E97	ḥ̇	u+2039	< / <
u+015F	ş	u+02C6	ˆ	u+0331	_	u+200C	ZWNJ	u+203A	> / >
u+0160	Š	u+02C7	ˇ	u+1E0C	Ḑ	u+200D	ZWJ	u+2044	/
u+0161	š	u+02D8	˘	u+1E0D	ḑ	u+200E	LRM	u+20AC	€
u+0162	Ṭ	u+02D9	˙	u+1E0E	Ḓ	u+200F	RLM	u+2122	™
u+0163	ṭ	u+02DA	˚	u+1E0F	ḓ	u+2013	— / —	u+2212	—
u+016A	Ū	u+02DB	˘	u+1E24	Ḥ	u+2014	— / —	u+2215	/
u+016B	ū	u+02DC	˜	u+1E25	ḥ	u+2018	‘ / ’	u+2219	•
u+0178	Ỳ	u+02DD	¨	u+1E2A	Ḧ	u+2019	’ / ’	u+25CC	◦
u+017D	Ž	u+0300	˘	u+1E2B	ḧ	u+201A	,	u+FB01	fi
u+017E	ž	u+0301	˙	u+1E62	Ş	u+201C	“ / ”	u+FB02	fl
u+0192	f	u+0302	ˆ	u+1E63	ş	u+201D	” / ”		

Stylistic Sets

The Bustani™ typeface automatically selects all the necessary shape variations of Arabic letters as the user inputs text. In certain contexts, the user can use the OpenType Stylistic Sets feature to override the automatically selected form with another. Of course, access to the Stylistic Sets varies from one application to another. Adobe® InDesign® allows access to all the listed stylistic sets; in addition, multiple stylistic sets can be applied to the same piece of text simultaneously. Note that in some specific text patterns, it might not always be possible to apply the requested change.

Stylistic Set 1: Rounded Reh

Transforms the sharp-tipped form of Reh into the rounded variant

في ركنه الذي اضطر إليه، وقد أخذ لنهار ينصرم
في ركنه الذي اضطر إليه، وقد أخذ لنهار ينصرم

Stylistic Set 2: Stretch Kaf (Zanadi)

Wherever possible, replaces the common form of the letter Kaf into the stretched form (ك), also known as Zanadi Kaf.

عاد إلى سكونه وجموده في ركنه
عاد إلى سكونه وجموده في ركنه

Stylistic Set 3: Flat-top numeral ٢

Changes the common shape of the numeral two into the flat-top form

١٩٧٢ < ١٩٧٢

٢/١ < ٢/١

٣١٢ < ٣١٢

عاد إلى سكونه وجموده^٢ في ركنه < عاد إلى سكونه وجموده^٢ في ركنه

Stylistic Set 4: Farsi numeral 0: round form

In Persian text, sometimes the round form of the numeral zero is preferred.

٢٠١٥ < ٢٠١٥

Stylistic Set 5: Yeh forms without two dots

As an editorial style, some publications prefer not to show the lower two dots on the final and separate forms of Yeh (u+064A).

فيعرف الصبي أن الليل قد أقبل، ويقدر في نفسه «
فيعرف الصبي أن الليل قد أقبل، ويقدر في نفسه

Stylistic Set 6: Tail-shaped Yeh

Replaces the bowl-shaped Yeh with the tail-shaped form (in single form or 2-letter combinations only)

فيعرف الصبي أن الليل قد أقبل، ويقدر في نفسه «
فيعرف الصبي أن الليل قد أقبل، ويقدر في نفسه
ي ي ع ي بي ثي كي في جي هي لي مي ضي طي سي شي «
ي ي ع ي بي ثي كي في جي هي لي مي ضي طي سي شي

Stylistic Set 7: Sukun: variant form for Urdu language

سکون « سکون

Stylistic Set 8: Noon Ghunna mark: variant form for Kashmiri language

Although common to both Urdu and Kashmiri text, the form of the Noon-Ghunna mark differs in Kashmiri-language context.

بند < بند

Stylistic Set 9: Feh & Qaf: Magrebi variants

Without changing the underlying character representation, this set changes the common dots for Feh and Qaf (ق ف) to conform to traditional Magrebi style (ف ب).

فيعرف الصبي أن الليل قد أقبل، ويقدر في نفسه «
يعرف الصبي أن الليل قد أقبل، ويقدر في نفسه

Stylistic Set 10: Final Heh: tail-like form

To be used wherever the lower, tail-like form of the final Heh is preferred.

ويعتمد بمرفقيه على ركبتيه ويخفي رأسه بيديه «
ويعتمد بمرفقيه على ركبتيه ويخفي رأسه بيديه
علامة واضحة « علامة واضحة

Stylistic Set 11: Meem: sloping form of final & separate shape

Replaces the vertical final and separate shapes of Meem with the more horizontal, sloping form.

ويخفي رأسه بيديه ويسلم نفسه لهذا الصوت ... «
ويخفي رأسه بيديه ويسلم نفسه لهذا الصوت ...
عام « عام

Stylistic Set 12: Heh: stacked knotted form

For certain 2-, 3-, and 4-letter sequences, replaces the bowtie-shape of the medial Heh with its stacked form.

جهد طهى نهب بهه مهى « جهد طهى نهب بهه مهى
ينهي مهند فهد سهه مهد « ينهي مهند فهد سهه مهد

Stylistic Set 13: Shadda & Kasra(-tan): uncombined form of the diacritics

To be used wherever the editorial style prefers not to allow the Kasra and Kasratan marks to combine with the Shadda.

سُلم الشيء إليه « سُلم الشيء إليه
ألم يك نطفة من مني تمنى « ألم يك نطفة من مني تمنى

Stylistic Set 14: Religious symbols: forms without marks

Provides an alternative, tashkeel-free form for the two symbols, u+FDF2 & u+FDF4.

الله « الله
محمد « محمد

Stylistic Set 15: Calendar indicators

Wherever the Western or Islamic calendar is explicitly designated by using the separate forms of Heh and Meem (هـ-م), allows the use of their respective preferred forms.

في عام ١٤٣٦ هـ
في سنة ١٩٩٢ م
سنة ١٤٣٦ هـ

Stylistic Set 16: Thousands Separator, alternative form

Replaces the common, top-aligned thousands-separator mark (u+066C) with an alternative, low-set variant.

١٢٠,٤٧٣ << ١٢٠'٤٧٣
١٢٠,٤٧٣,٩٩ << ١٢٠'٤٧٣,٩٩

Stylistic Set 17: Tanween marks: alternative forms for Koranic notation

Replaces the common forms of the Tanween marks (u+064B, 064C, 064D) with the corresponding open forms.

On systems that support the character set of Unicode 6.0, the three open-form Tanween characters can be accessed directly as u+07F0, u+07F1, and u+07F2, respectively.

بَلِ الْإِنْسَانُ عَلَىٰ نَفْسِهِ بَصِيرَةٌ << بَلِ الْإِنْسَانُ عَلَىٰ نَفْسِهِ بَصِيرَةٌ
ثُمَّ كَانَ عِلْقَةً فُخِّقَ فَسَوَّىٰ << ثُمَّ كَانَ عِلْقَةً فُخِّقَ فَسَوَّىٰ

Numeric Notation

Bustani supports a broad variety of numeric notation for an Arabic-script context, as well as for biscriptal setting (Latin & Arabic).

The basic numeral characters (and their variant forms) are included in the following styles:

Arabic (u+0660–0669)	٠١٢٣٤٥٦٧٨٩ (٢)
Indic-Arabic (u+06F0–06F9)	٠١٢٣٤٥٦٧٨٩ (٠١٢٣٤٥٦٧٨٩)
European (u+0030–0039)	0123456789

All the above numerals shown in their standard size are also available in the following variant forms:

Superscript	٠١٢٣٤٥٦٧٨٩ ٠١٢٣ 0123456789	Subscript	٠١٢٣٤٥٦٧٨٩ ٠١٢٣ 0123456789
Numerator	٠١٢٣٤٥٦٧٨٩ ٠١٢٣ 123456789	Denominator	٠١٢٣٤٥٦٧٨٩ ٠١٢٣ 123456789

Fractions

Bustani provides the means to create any fraction in either Arabic or European style. To create a fraction in European-style numerals, enter the numerator, a slash, followed by the denominator. So, for instance, if we wish to create the fraction $\frac{12}{45}$, we enter 12, then /, then 45. Once we can see the whole expression as 12/45, we select 12/45 before activating the OpenType fraction feature. When we do so, 12/45 instantly becomes $\frac{12}{45}$. For applications that do not provide access to the fraction feature, we can resort to the numerator and denominator features instead. Once we have entered 12/45, we first select the numerator as in 12/45 and activate the numerator feature. The result will appear as $\frac{12}{45}$. We then proceed to select the slash and the denominator as in $\frac{12}{45}$, then immediately activate the denominator feature, yielding the desired result of $\frac{12}{45}$.

Ideally, fractions of Arabic-style numerals should function the same way in applications because numbers in Arabic script also move from left to right. Unfortunately, most applications have not yet reached this level of sophistication. To create the $\frac{12}{45}$ fraction in Arabic-style numerals, we first enter the numerator ١٢, followed by slash (\ or /) then ٤٥, as in ١٢\٤٥. We then select the ١٢ and activate the numerator feature, yielding ١٢\٤٥; then we select slash along with the denominator ١٢\٤٥ and activate the denominator feature, giving the desired result ١٢\٤٥.

Number-bearing symbols

Bustani supports a gamut of number-bearing signs that allow the user to create the combined result of a particular symbol along with a specific set of numerals.

Each of the listed symbols can carry one or more figures.

Number sign (u+0600)	ٓ
Year sign (u+0601)	ٔ
Footnote marker (u+0602)	ٕ
Page marker (u+0603)	ٖ
End-of-Ayah (u+06DD)	◌۞

Most of the symbols carry up to 3 figures, while the year sign can carry up to 4. The internal logic of the font chooses the appropriate size of the figures for each of the symbols. Following are some samples:



To create a combined symbol as in the above samples, enter the code for the desired symbol (e.g. 0601 for ٔ) followed by the desired numerals.

Please note that some word processors and applications may not fully support all the numeric features listed above.

Script- and Language-specific features

In applications that allow text to be tagged as belonging to a particular language, a certain set of symbols in Bustani can appear in one of two styles, one compatible with the Roman (Latin) script, the other with Arabic. Some applications can automatically infer the language and/or script from the characters present in the text. In mixed-script contexts, it is often necessary to remove ambiguity by tagging the various portions of text manually with the appropriate language.

Latin vs. Arabic variants

In an Arabic-script context such as Arabic, Farsi, or Urdu, the following characters appear as

* [] () / : — — = ! « » () “ ” ‘ ’ - + - × ÷ . • ...

...while in a Latin-script context such as English, French, or German, they appear as

* [] () / : — — = ! « » < > “ ” ‘ ’ - + - × ÷ . • ...

In the character-set chart, symbols that have two potential representations are shown with the two forms separated by a slash /.

Farsi

Decimal point $\setminus / ٬ \bullet$ (u+066B) instead of $\setminus , ٬ \bullet$ Arabic


For the round zero \circ instead of \bullet , select Stylistic Set 4.

Urdu, Pashto

Eastern Arabic-Indic Numerals (u+06F0...) are rendered as $\bullet ٠ ١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩$ instead of the Persian forms of $\bullet ٠ ١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩$

Noon Ghunna \cup (u+06BA) in initial and medial position is rendered as پنجاب instead of the default style as پنجاب

Urdu & other South Asian languages

When written in Nastaliq style, South-Asian languages that utilize the Barree Yeh character (ے u+06D2), it is common to see all the lower dots and marks of a word transposed below the Barree Yeh. In deference to this tradition, Bustani follows this same model in rendering words such as 

Sindhi

Eastern Arabic-Indic Numerals (u+06F0–06F9) are rendered as ٠١٢٣٤٥٦٧٨٩, distinct from the Persian form ۰۱۲۳۴۵۶۷۸۹ or Urdu form ۰۱۲۳۴۵۶۷۸۹

Kashmiri

The Noon Ghunna mark (u+0658) is rendered as پَنجاب instead of Urdu style as پَنجاب

This rendering can also be achieved by applying Stylistic Set 8.

Quranic Marks

While most Quranic marks act as aids in pronunciation or recitation, some are simply organizational. The marks supported by Bustani lie in the Unicode ranges u+0615–u+06ED and u+08F0–u+08F3. Immediately below, the full set is shown. Each entry indicates the Unicode value along with its graphic representation. Yet further below, we show at least one use of each mark in Quranic text.

u+0615	ط	u+06DB	۞	u+06E1	َ	u+06E7	ے	u+06ED	◌ٓ
u+06D6	صل	u+06DC	س	u+06E2	م	u+06E8	ن	u+08F0	◌ٓ
u+06D7	قل	u+06DD	۞	u+06E3	س	u+06E9	۞	u+08F1	◌ٓ
u+06D8	م	u+06DE	۞	u+06E4	َ	u+06EA	◌ٓ	u+08F2	◌ٓ
u+06D9	لا	u+06DF	و	u+06E5	و	u+06EB	◌ٓ	u+08F3	◌ٓ
u+06DA	ج	u+06E0	و	u+06E6	ے	u+06EC	◌ٓ		

The characters shown above fall into several categories: non-spacing marks, spacing marks, and independent symbols as follows:

Similar in behavior to Arabic vowel marks (a.k.a. *harakāt*), non-spacing marks have no width of their own, and consequently, need to attach themselves to a character. Most commonly, a non-spacing mark will attach itself to an alphabetic letter, or even to one of the *harakāt*, at a predetermined point. The presence of a non-spacing mark in the middle of a word does not interrupt the flow of the word in any way. In the table above, each non-spacing mark appears above or below a dotted circle (◌ٓ) to indicate its status as a dependent symbol. In the table above, there are two pairs of marks that have the same shape, but differ only in their spacing behavior; compare u+06E6 with u+06E7, as well as u+06E5 with u+08F3. As made evident by the samples that follow, a non-spacing mark can combine with (above or below) another spacing or non-spacing mark.

On the other hand, each spacing mark has an intrinsic width and appears only before or after a word, but never in its midst. In fact, if a spacing mark is inserted into the middle of a word, the word will break in two.

Finally, independent symbols act as punctuation or place indicators within the text. The most common of these signs ۞ is the end of verse mark that follows each *āyah* (Arabic for verse) and indicates its sequential number within the current *sūrah*.

Sample 1: u+06DD ﴿ & u+06D7 ٓ

﴿سورة البقرة﴾

وَإِذَا قِيلَ لَهُمْ ءَامِنُوا كَمَا ءَامَنَ النَّاسُ قَالُوا أَنُؤْمِنُ كَمَا ءَامَنَ السُّفَهَاءُ ۖ أَلَا إِنْتُمْ هُمُ السُّفَهَاءُ وَلَكِنْ لَا يَعْلَمُونَ ﴿١٣﴾

To create a specific verse number, enter the code for the symbol followed by the number. The sample above was created by inputting ﴿ u+06DD followed by the number ١٣.

Sample 2: u+06D8 ُ & u+06D6 ُ & u+08F0 ُ

﴿إِنْ أَلَّهَ لَا يَسْتَحْيَ أَنْ يَضْرِبَ مَثَلًا مَّا بَعُوضَةً فَمَا فَوْقَهَا فَأَمَّا الَّذِينَ ءَامَنُوا فَيَعْلَمُونَ أَنَّهُ الْحَقُّ مِنْ رَبِّهِمْ ۖ وَأَمَّا الَّذِينَ كَفَرُوا فَيَقُولُونَ مَا ذَا أَرَادَ اللَّهُ بِهَذَا مَثَلًا ۖ يُضِلُّ بِهِ كَثِيرًا وَيَهْدِي بِهِ كَثِيرًا وَمَا يُضِلُّ بِهِ إِلَّا الْفَاسِقِينَ ﴿٢٦﴾﴾

Sample 3: u+06D9 ِ & u+08F1 ِ & u+08F2 ِ

وَبَشِّرِ الَّذِينَ ءَامَنُوا وَعَمِلُوا الصَّالِحَاتِ أَنَّ لَهُمْ جَنَّاتٍ تَجْرَى مِنْ تَحْتِهَا الْأَنْهَارُ ۖ كُلَّمَا رُزِقُوا مِنْهَا مِنْ ثَمَرَةٍ رِزْقًا قَالُوا هَذَا الَّذِي رُزِقْنَا مِنْ قَبْلُ وَأَتُوا بِهِ ۖ مُتَشَبِهًا وَلَهُمْ فِيهَا أَزْوَاجٌ مُطَهَّرَةٌ ۖ وَهُمْ فِيهَا خَالِدُونَ ﴿٢٥﴾﴾

Special Note on the open nunation (Arabic tanwīn) marks u+08F0-u+08F2 (ُ ِ ِ):

If these combining marks are not available on a particular system or application, one can use the tanwīn marks at u+064B-u+064D and then change their appearance to the open form through the stylistic set 17 (see Stylistic Sets).

Sample 4: u+06DA ِ

أَوْ كَصَيْبٍ مِّنَ السَّمَاءِ فِيهِ ظُلُمٌ وَرَعْدٌ وَبَرْقٌ يَجْعَلُونَ أَصْبَعَهُمْ فِي ۖءَاذَانِهِم مِّنَ الصَّوَاعِقِ حَذَرَ الْمَوْتِ ۗ وَاللَّهُ مُحِيطٌ بِالْكَافِرِينَ ﴿١٩﴾﴾

Sample 5: u+06DB ٥

ذٰلِكَ الْكِتٰبُ لَا رَيْبَ فِيْهِ هُدًى لِّلْمُتَّقِيْنَ ﴿٢﴾

Sample 6: u+06DC ٥ & u+06E5 و

مِّنْ ذَا الَّذِي يُّقْرِضُ اللّٰهَ قَرْضًا حَسَنًا فَيُضِعُّهُ لَهُ ۖ اَضْعَافًا كَثِيْرَةً ۗ وَاللّٰهُ يَقْبِضُ وَيَبْصُطُ ۗ وَاِلَيْهِ تُرْجَعُوْنَ ﴿٢٤٥﴾

Sample 7: u+06DF ٥ & u+06ED ٥

اُولٰٓئِكَ عَلٰى هُدًى مِّنْ رَّبِّهِمْ ۖ وَاُولٰٓئِكَ هُمُ الْمُفْلِحُوْنَ ﴿٥﴾

Sample 8: u+06E0 ٥ & u+06E7 ٥

اَلَمْ تَرَ اِلَى الَّذِي حَاجَّ اِبْرٰهِيْمَ فِيْ رَبِّهٖ ۚ اَنْ عَاتٰهُ اللّٰهُ الْمُلْكَ ۖ اِذْ قَالَ اِبْرٰهِيْمُ رَبِّى الَّذِي يُحْيِىْ وَيُمِيْتُ قَالَ اَنَا اُحْيِىْ ۚ وَاُمِيْتُ قَالَ اِبْرٰهِيْمُ فَاِنَّ اللّٰهَ يَآتِى بِالسَّمْسِ مِنْ الْمَشْرِقِ ۖ فَاَتِ بِهَا مِنَ الْمَغْرِبِ فَبُهِتَ الَّذِي كَفَرَ ۗ وَاللّٰهُ لَا يَهْدِى الْقَوْمَ الظّٰلِمِيْنَ ﴿٢٥٨﴾

Sample 9: u+06E1 ٥

الَّذِيْنَ يُؤْمِنُوْنَ بِالْغَيْبِ وَيُقِيْمُوْنَ الصَّلٰوةَ وَمِمَّا رَزَقْنٰهُمْ يُنْفِقُوْنَ ﴿٢﴾

Sample 10: u+06E2 ٥

فِيْ قُلُوْبِهِمْ مَّرَضٌ ۖ فَزَادَهُمُ اللّٰهُ مَرَضًا وَلَهُمْ عَذَابٌ اَلِيْمٌ ۖ بِمَا كَانُوْا يَكْذِبُوْنَ ﴿١٠﴾

Sample 11: u+06DE ٥ & u+06E4 ٥ & u+06E6 ٥

﴿٢٦﴾ اِنَّ اللّٰهَ لَا يَسْتَحْيِى ۚ اَنْ يُّضْرَبَ مَثَلًا مَّا بَعُوْضَةٌ مُّتَمِّتًا فَوْقَهَا ۖ فَاَمَّا الَّذِيْنَ ءَامَنُوْا فَيَعْلَمُوْنَ اَنَّهُ الْحَقُّ ۖ مِنْ رَّبِّهِمْ ۖ وَاَمَّا الَّذِيْنَ كَفَرُوْا فَيَقُوْلُوْنَ مَاذَا اَرَادَ اللّٰهُ بِهٰذَا مَثَلًا ۖ يُضِلُّ بِهِ ۚ كَثِيْرًا وَيَهْدِيْ بِهِ ۚ كَثِيْرًا ۖ وَمَا يُضِلُّ بِهِ ۚ اِلَّا الْفٰسِقِيْنَ ﴿٢٦﴾

Sample 12: u+06E3 ٲ

﴿سورة الطور﴾

أَمْ عِنْدَهُمْ خَزَائِنُ رَبِّكَ أَمْ هُمُ الْمُصِيطِرُونَ ﴿٢٧﴾

Sample 13: u+0615 ط

(Urdu tradition only)

﴿سورة المجادلة﴾

قَدْ سَمِعَ اللَّهُ قَوْلَ الَّتِي تُجَادِلُكَ فِي زَوْجِهَا وَتَشْتَكِي إِلَى اللَّهِ وَاللَّهُ يَسْمَعُ تَحَاوُرَكُمَا إِنَّ اللَّهَ سَمِيعٌ بَصِيرٌ ﴿١﴾

Sample 14: u+06E8 ٲ

﴿سورة الأنبياء﴾

فَأَسْتَجِبْنَا لَهُ وَنَجِّنَاهُ مِنَ الْغَمِّ وَكَذَلِكَ نُخَيِّضُ الْمُؤْمِنِينَ ﴿٨٨﴾

Sample 15: u+06E9 ٲ

﴿سورة الإعراف﴾

ٲ إِنَّ الَّذِينَ عِنْدَ رَبِّكَ لَا يَسْتَكْبِرُونَ عَنْ عِبَادَتِهِ وَيُسَبِّحُونَهُ وَلَهُ يَسْجُدُونَ ﴿٢٠٦﴾

Sample 16: u+06EA ٲ

﴿سورة هود﴾

وَقَالَ أَرْكَبُوا فِيهَا بِسْمِ اللَّهِ مَجْرِبَهَا وَمُرسَهَا إِنَّ رَبِّي لَغَفُورٌ رَحِيمٌ ﴿٤١﴾

Sample 17: u+06EB ٲ

﴿سورة يوسف﴾

قَالُوا يَا أَبَانَا مَا لَكَ لَا تَأْمَنَّا عَلَى يُوسُفَ وَإِنَّا لَهُ لَنَصِحُونَ ﴿١١﴾

﴿سورة فصلت﴾

وَلَوْ جَعَلْنَاهُ قُرْءَانًا أَعْجَمِيًّا لَقَالُوا لَوْلَا فُصِّلَتْ آيَاتُهُ ۖ أَأَعْجَمِيٌّ وَعَرَبِيٌّ ۚ قُلْ هُوَ لِلَّذِينَ آمَنُوا هُدًى وَشِفَاءٌ ۚ
وَالَّذِينَ لَا يُؤْمِنُونَ فِي آذَانِهِمْ وَقْرٌ وَهُوَ عَلَيْهِمْ عَمًى ۚ أُولَٰئِكَ يُنَادُونَ مِنْ مَّكَانٍ بَعِيدٍ ﴿٤٤﴾

﴿سورة الإسراء﴾

إِنْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنفُسِكُمْ ۖ وَإِنْ أَسَأْتُمْ فَلَهَا ۚ فَإِذَا جَاءَ وَعْدُ الْآخِرَةِ لِيَسُوءَ وُجُوهَكُمْ وَلِيَدْخُلُوا
الْمَسْجِدَ كَمَا دَخَلُوهُ أَوَّلَ مَرَّةٍ وَلِيُتَبِّرُوا مَا عَلَوْا تَتْبِيرًا ﴿٧﴾

Since this character (u+08F3) is a relatively recent addition to the Unicode Standard, it may not be accessible in certain applications and environments.

Transliteration

To accommodate the various traditions for transliteration and transcription—i.e. the representation of Arabic letters and sounds in the Roman alphabet, the character repertoire of Bustani includes all of the following characters, in addition to the standard set for Western, Northern, and Southern European languages:

' Āā Ēē Īī Ōō Ūū Đđ Ġġ Ģģ Ħħ Hh Ss Šš Tt Žž Žž

Each of the following examples represents an independent library record chosen for its representative quality. Each example demonstrates a different system of transliteration.

Example 1:

Manuscript: Description: Recueil factice. Papier. 200 feuillets. 225×170 mm.

Contenu: 1^o f. 16–27: Urğūza sur les relations par Ah. b. Sulh Tadağī...

2^o f. 28–52: Sentences de M. b. ‘Abd Karīm b. ‘Atā Allāh

(من علامات الاعتمادات على العمل incipit)

3^o f. 53–60: commentaire par ‘All. b. Ibrahīm b. ‘Alawī sur

une urğūza énumérant quarante maîtres mystiques

(الحمد لله التي قامت بامرہ الارضون incipit)

4° f. 61–89 [incomplet de la fin]: traité d'arithmétique par

Yah. b. 'All.... b. Zakariya Ansārī Awsī

(الحمد لله الواحد الفرد (incipit)

Example 2:

Ibn Hajar al-‘Asqalānī, Ahmad ibn ‘Alī, 1372–1449.

ابن حجر العسقلاني، أحمد بن علي، ١٣٧٢-١٤٤٩

L.C.

العسقلاني، شهاب الدين أبي الفضل أحمد بن علي

شهاب الدين ابي الفضل احمد بن علي بن محمد بن علي بن احمد

الكناني العسقلاني المصري

Honorific marks

The abbreviated honorific marks supported by Bustani lie in the Unicode range u+0610–u+0614 as shown immediately below. Each entry indicates the Unicode value along with its graphic representation.

u+0610	◌ۥ
u+0611	◌ۦ
u+0612	◌ۧ
u+0613	◌ۨ
u+0614	◌۩

The characters shown above are all non-spacing marks, which means they have no width of their own and will usually attach to an alphabetic letter or a diacritic mark such as َ, for instance. To add an honorific mark to a word, enter its code value (u+0610–0614) after the letter to which it should attach.

The following examples show examples of these marks in use:

اسْمَعِيلَ خَالِدٌ مَازِنٌ

خَالِدٌ

In the top line, all the honorific are well placed, while on the second line the mark is overlaying the letter Lam of the word. Since the honorific marks take up some space to their right and left, they need to be judiciously placed so as not to collide with neighboring letters.