

The arabluatex package

v1.2 – 2016/05/09

Robert Alessi
alessi@robertalessi.net

Contents

License and disclaimer	2	4.5 The “pipe” character ()	19
1 Introduction	2	4.6 Stretching characters: the <i>taṭwīl</i> . . .	20
1.1 arabluatex is for Lua \TeX	3	4.7 Digits	20
2 The basics of arabluatex	4	4.7.1 Numerical figures	20
2.1 Activating arabluatex	4	4.7.2 The <i>abjad</i>	20
2.2 Options	5	4.8 Additional characters	21
2.2.1 Classic contrasted with modern typesetting of Arabic . . .	5	4.9 Arabic emphasis	21
2.3 Typing Arabic	6	5 Special applications	21
2.3.1 Local options	7	6 Transliteration	22
3 Standard Arab\TeX input	7	6.1 Examples	23
3.1 Consonants	7	7 \TeX Commands in Arabic environments	23
3.2 Vowels	8	7.1 csquotes	26
3.2.1 Long vowels	8	7.2 reledmac	26
3.2.2 Short vowels	9	8 Future work	26
4 arabluatex in action	10	9 Implementation	27
4.1 The vowels and diphthongs	10	References	31
4.2 Other orthographic signs	11	Change History	31
4.3 Special orthographies	16	Index	32
4.4 Quoting	17		
4.4.1 Quoting the <i>hamza</i>	19		

List of Tables

1 Arab \TeX consonants	7	4 “Quoted” <i>hamza</i>	19
2 Arab \TeX long vowels	8	5 Additional Arabic codings	21
3 Arab \TeX short vowels	9		

Abstract

This package provides for Lua \TeX an Arab \TeX -like interface to generate Arabic writing from an ASCII transliteration. It is particularly well-suited for complex documents such as technical documents or critical editions where a lot of left-to-right commands intertwine with Arabic writing. `arabluatex` is able to process any Arab \TeX input notation. Its output can be set in the same modes of vocalization as Arab \TeX , or in different roman transliterations. It further allows many typographical refinements. It will eventually interact with some other packages yet to come to produce from `.tex` source files, in addition to printed books, TEI xml compliant critical editions and/or lexicons that can be searched, analyzed and correlated in various ways.

License and disclaimer

Copyright © 2016 Robert Alessi

Please send error reports and suggestions for improvements to Robert Alessi:

- email: [<alessi@robertalessi.net>](mailto:alessi@robertalessi.net)
- website: <http://www.robertalessi.net/arabluatex>
- development: <http://git.robertalessi.net/arabluatex>
- comments, feature requests, bug reports: <https://notabug.org/ralessi/arabluatex>

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see [<http://www.gnu.org/licenses/>](http://www.gnu.org/licenses/).

This release of `arabluatex` consists of the following source files:

- `arabluatex.ins`
- `arabluatex.dtx`
- `arabluatex.lua`
- `arabluatex_voc.lua`
- `arabluatex_fullvoc.lua`
- `arabluatex_novoc.lua`
- `arabluatex_trans.lua`

1 Introduction

In comparison to Prof. Lagally's outstanding Arab \TeX ,¹ ArabLua \TeX is at present nothing more than a modest piece of software. Hopefully—if I may say so—it will eventually provide all of its valuable qualities to the Lua \TeX users.

¹See <http://ctan.org/pkg/arabtex>

arabtex dates back to 1992. As far as I know, it was then the first and only way to typeset Arabic texts with \TeX and \LaTeX . To achieve that, arabtex provided—and still does—an Arabic font in *Nashī* style and a macro package that defined its own input notation which was, as the author stated, “both machine, and human, readable, and suited for electronic transmission and e-mail communication”.² Even if the same can be said about Unicode, Arab \TeX ASCII input notation still surpasses Unicode input, in my opinion, when it comes to typesetting complex documents, such as scientific documents or critical editions where footnotes and other kind of annotations can be particularly abundant. It must also be said that most text editors have trouble in displaying Arabic script connected with preceding or following \LaTeX commands: it often happens that commands seem misplaced, not to mention punctuation marks, or opening or closing braces, brackets or parentheses that are unexpectedly displayed in the wrong direction. Of course, some text editors provide ways to get around such difficulties by inserting invisible Unicode characters, such as LEFT-TO-RIGHT or RIGHT-TO-LEFT MARKS (U+200E, U+200F), RTL/LTR “embed” characters (U+202B, U+202A) and RLO/LRO “bidi-override” characters (U+202E, U+202D).³ Nonetheless, it remains that inserting all the time these invisible characters in complex documents rapidly becomes confusing and cumbersome.

The great advantage of Arab \TeX notation is that it is immune from all these difficulties, let alone its being clear and straightforward. One also must remember that computers are designed to process code. Arab \TeX notation is a way of encoding Arabic language, just as \TeX “mathematics mode” is a way of processing code to display mathematics. As such, not only does it allow greater control over typographical features, but it also can be processed in several different ways: so without going into details, depending on one’s wishes, Arab \TeX input can be full vocalized Arabic (*scriptio plena*), vocalized Arabic or non-vocalized Arabic (*scriptio defectiva*); it further can be transliterated into whichever romanization standard the user may choose.

But there may be more to be said on that point, as encoding Arabic also naturally encourages the coder to vocalize the texts—without compelling him to do so, of course. Accurate coding may even have other virtuous effects. For instance, hyphens may be used for tying particles or prefixes to words, or to mark inflectional endings, and so forth. In other words, accurate coding produces accurate texts that can stand to close grammatical scrutiny and to complex textual searches as well.

Having that in mind, I started arabluatex. With the help of Lua, it will eventually interact with some other packages yet to come to produce from `.tex` source files, in addition to printed books, \TeX \XML compliant critical editions and/or lexicons that can be searched, analyzed and correlated in various ways.

1.1 arabluatex is for Lua \LaTeX

It goes without saying that arabluatex requires Lua \LaTeX . \TeX and \LaTeX have arabtex, and \XeLaTeX has arabxetex. Both of them are much more advanced than arabluatex,

²Lagally (2004, p. 2).

³Gáspár Sinai’s Yudit probably has the best Unicode support. See <http://www.yudit.org>.

as they can process a number of different languages,⁴ whereas arabluatex can process only Arabic for the time being. More languages will be included in future releases of arabluatex.

In comparison to arabxetex, arabluatex works in a very different way. The former relies on the [TECKit](#) engine which converts ArabT_EX input on the fly into Unicode Arabic script, whereas the latter passes ArabT_EX input on to a set of Lua functions. At first, L^AT_EX commands are taken care of in different ways: some, as `\emph`, `\textbf` and the like are expected to have Arabic text as arguments, while others, as `\LR`, for “left-to-right text”, are not. Then, once what is Arabic is carefully separated from what is not, it is processed by other Lua functions which rely on different sets of correspondence tables to do the actual conversion in accordance with one’s wishes. Finally, Lua returns to T_EX the converted strings—which may in turn contain some other ArabT_EX input yet to be processed—for further processing.

2 The basics of arabluatex

2.1 Activating arabluatex

As usual put in your preamble:

```
\usepackage{arabluatex}
```

The only requirement of arabluatex is LuaL^AT_EX; it will complain if you try to compile your document with another engine. That aside, arabluatex does not load packages such as polyglossia or luabidi. It can work with polyglossia though, but does not require it.

Font setup If you wish to use your own Arabic font, you can define it before loading arabluatex. Assuming that fontspec is loaded, put this in your preamble just above the line that loads arabluatex:

```
\newfontfamily\arabicfont[Script=Arabic]{\fontname}
```

where `\fontname` is the standard name of the Arabic font you wish to use.

By default, if no Arabic font is selected, arabluatex will issue a warning message and attempt to load the Amiri font⁵ like so:—

```
\newfontfamily\arabicfont[Script=Arabic]{Amiri}
```

REM. By default Amiri places the *kasra* in combination with the *tašdid* below the consonant, like so: `كـ`. That is correct, as at least in the oldest manuscripts `كـ` may stand for `ك` as well as `كـ`. See Wright (1896, i.14.C–D). The placement of the *kasra* above the consonant may be obtained by selecting the `ss05` feature of the Amiri font, like so:—⁶

```
\newfontfamily\arabicfont[Script=Arabic,RawFeature={+ss05}]{Amiri}
```

Other Arabic fonts may behave differently.

⁴To date, both packages support Arabic, Maghribi, Urdu, Pashto, Sindhi, Kashmiri, Uighuric and Old Malay; in addition to these, arabxetex also has a Hebrew mode, including Judeo-Arabic and Yiddish.

⁵Hosny (2015).

⁶See the documentation of amiri, Hosny (2015, p. 5).

2.2 Options

arabluatex may be loaded with four mutually exclusive global options, each of which may be overridden at any point of the document (see below section 2.3.1 on page 7):

voc

default

In this mode, which is the one selected by default, every short vowel written generates its corresponding diacritical mark: *damma* (ُ), *fatha* (َ) and *kasra* (ِ). If a vowel is followed by $\langle uN, aN, iN \rangle$, then the corresponding *tanwīn* (ـِ, ـُ, ـِ or ـٍ) is generated. Finally, $\langle u, a, i \rangle$ at the commencement of a word indicate a “connective ‘alif” (‘*alifu* ‘*l-waṣli*), but voc mode does not show the *waṣla* above the ‘alif; instead, the accompanying vowel is expressed (اِ, اُ, اِ).

fullvoc

In addition to what the voc mode does, fullvoc expresses the *sukūn* and the *waṣla*.

novoc

None of the diacritics is showed in novoc mode, unless otherwise specified (see “quoting” technique below section 4.4 on page 17).

trans

This mode transliterates the ArabTeX input into one of the accepted standards. At present, two standards are supported (see below section 6 on page 22 for more details):

dmg *Deutsche Morgenländische Gesellschaft*, which is selected by default;

loc *Library of Congress*.

More standards will be included in future releases of arabluatex.

2.2.1 Classic contrasted with modern typesetting of Arabic

By default, arabluatex typesets Arabic in a classic, traditional style the most prominent features of which are the following:

- ‘Classic’ *madda*: when ‘alif and *hamza* accompanied by a simple vowel or *tanwīn* is preceded by an ‘alif of prolongation (ـِ), then a mere *hamza* is written on the line, and a *madda* is placed over the ‘alif, like so:—

samA'uN سَمَاءُ *samā'un*, jA'a جَاءَ *ǧā'a*, yaTasA'alUna يَتَسَاءَلُونَ *yatasā'alūna*⁷
(see on page 13 for further details).

- The euphonic *tašdīd* is generated (see on page 13).
- Assimilation rules laid on item b on page 13 are applied.
- In fullvoc mode, the *sukūn* is expressed.

\SetArbEasy

\SetArbDflt

Such refinements may be discarded by the command \SetArbEasy, either globally in the preamble or locally at any point of the document. Default ‘classic’ rules may be set back at any point of the document with the command \SetArbDflt. Examples follow:—

(a) \SetArbDflt:

- voc وَمَاتَ اسْتِسْقَاءَ قَبْلَ أَنْ يَتِمَّ كِتَابُهُ فِي نُجُومِ السَّمَاءِ
- fullvoc وَمَاتَ اسْتِسْقَاءَ قَبْلَ أَنْ يَتِمَّ كِتَابُهُ فِي نُجُومِ السَّمَاءِ

⁷Note that in old mss. such forms as سَمَاءُ, جَاءُ are also found; see Wright (1896, i. 24 D).

- iii. trans *wa-māta 'šisqā'an qabla 'ay yutimma kitāba-hu fī nuḡūmi 's-samā'i*
- (b) \SetArbEasy:
 - i. voc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نَجْمِ السَّمَاءِ
 - ii. fullvoc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نَجْمِ السَّمَاءِ
 - iii. trans *wa-māta 'šisqā'an qabla 'an yutimma kitāba-hu fī nuḡūmi 's-samā'i*

Please note that this document is typeset with \SetArbDflt throughout.

2.3 Typing Arabic

\arb Once arabluatex is loaded, a command \arb{*Arabic text*} is available for inserting Arabic text in paragraphs, like so:—

```

1 From \textcite[i. 1 A]{Wright}:--- Arabic, like Hebrew and
2 Syriac, is written and read from right to left. The letters
3 of the alphabet (\arb{.hurUfu 'l-hijA'i}, \arb{.hurUfu
4 'l-tahajjI}, \arb{al-.hurUfu 'l-hijA'iyyaTu}, or \arb{.hurUfu
5 'l-mu`jami}) are twenty-eight in number and are all
6 consonants, though three of them are also used as vowels
7 (see §3).
```

From Wright (1896, i. 1 A):— Arabic, like Hebrew and Syriac, is written and read from right to left. The letters of the alphabet (أَلِفٌ, حُرُوفُ التَّهْجِي, حُرُوفُ الْهِجَاءِ, or حُرُوفُ الْمُعْجَمِ) are twenty-eight in number and are all consonants, though three of them are also used as vowels (see § 3).

arab Running paragraphs of Arabic text should rather be placed inside an *Arabic environment*

```

1 \begin{arab}
2 [...]
3 \end{arab}
```

like so:—

```

1 \begin{arab}
2 'at_A .sadIquN 'il_A ju.hA ya.tlubu min-hu .himAra-hu
3 li-yarkaba-hu fI safraTiN qa.sIraTiN. wa-qAla la-hu:
4 \enquote{sawfa 'u`Idu-hu 'ilay-ka fI 'l-masA'i
5 wa-'adfa`u la-ka 'ujraTaN.} fa-qAla ju.hA:
6 \enquote{'anA 'AsifuN jiddaN 'annI lA 'asta.tI`u 'an
7 'u.haqqiqa la-ka .garbata-ka fa-'l-.himAru laysa hunA
8 'l-yawma.} wa-qabla 'an yutimma ju.hA kalAma-hu bada'a
9 'l-.himAru yanhaqu fI 'i.s.tabili-hi. fa-qAla la-hu
```

```

10 .sadIqu-hu: \enquote{'innI 'asma`u .himAra-ka yA ju.hA
11 yanhaqu.} fa-qAla la-hu ju.hA: \enquote{.garIbuN
12 'amru-ka yA .sadIqI 'a-tu.saddiqu 'l-.himAra
13 wa-tuka_d_dibu-nI?}
14 \end{arab}

```

أَتَى صَدِيقٌ إِلَى جَحَا يَطْلُبُ مِنْهُ حِمَارَهُ لِيَرْكَبَهُ فِي سَفَرَةٍ قَصِيرَةٍ. وَقَالَ لَهُ: "سَوْفَ أُعِيدُهُ إِلَيْكَ فِي الْمَسَاءِ
وَأَدْفَعُ لَكَ أُجْرَةً." فَقَالَ جَحَا: "أَنَا آسَفٌ جِدًّا أَنِّي لَا أَصْطَلِيعُ أَنْ أُحَقِّقَ لَكَ غَرَبَتَكَ فَالْحِمَارُ لَيْسَ هُنَا
الْيَوْمَ." وَقَبْلَ أَنْ يُتِمَّ جَحَا كَلَامَهُ بَدَأَ الْحِمَارُ يَنْهَقُ فِي إِصْطِلِيلِهِ. فَقَالَ لَهُ صَدِيقُهُ: "إِنِّي أَسْمَعُ حِمَارَكَ يَا جَحَا
يَنْهَقُ." فَقَالَ لَهُ جَحَا: "غَرِيبُ أَمْرِكَ يَا صَدِيقِي أَتَصَدِّقُ الْحِمَارَ وَتُكَذِّبُنِي؟"

2.3.1 Local options

As seen above in section 2.2 on page 5, arabluatex may be loaded with four mutually exclusive global options: `voc` (which is the default option), `fullvoc`, `novoc` and `trans`. Whatever choice has been made globally, it may be overridden at any point of the document, as the `\arb` command may take any of the `voc`, `fullvoc`, `novoc` or `trans` modes as optional arguments, like so:—

```

voc      - \arb[voc]{\langle Arabic text \rangle};
fullvoc  - \arb[fullvoc]{\langle Arabic text \rangle};
novoc    - \arb[novoc]{\langle Arabic text \rangle};
trans    - \arb[trans]{\langle Arabic text \rangle}.

```

The same optional arguments may be passed to the environment `arab`: one may have `\begin{arab}[\langle mode \rangle] \dots \end{arab}`, where `\langle mode \rangle` may be any of `voc`, `fullvoc`, `novoc` or `trans`.

3 Standard ArabTeX input

3.1 Consonants

Table 1 gives the ArabTeX equivalents for all of the Arabic consonants.

Letter	Transliteration ⁸		ArabTeX notation
	dmg	loc	
ا ⁹	<i>a</i>	<i>a</i>	<code>a</code>
ب	<i>b</i>	<i>b</i>	<code>b</code>
ت	<i>t</i>	<i>t</i>	<code>t</code>
ث	<i>t̤</i>	<i>th</i>	<code>_t</code>

⁸See below section 6 on page 22.

⁹For 'alif as a consonant, see Wright (1896, i. 16 D).

Letter	Transliteration		ArabTeX notation
	dmg	loc	
ج	ǧ	j	^g or j
ح	ḥ	ḥ	.h
خ	ḫ	kh	_h or x
د	d	d	d
ذ	ḏ	dh	_d
ر	r	r	r
ز	z	z	z
س	s	s	s
ش	š	sh	^s
ص	ṣ	ṣ	.s
ض	ḍ	ḍ	.d
ط	ṭ	ṭ	.t
ظ	ẓ	ẓ	.z
ع	‘	‘	^
غ	ġ	gh	.g
ف	f	f	f
ق	q	q	q
ك	k	k	k
ل	l	l	l
م	m	m	m
ن	n	n	n
ه	h	h	h
و	w	w	w
ي	y	y	y
ة	a	ah	T

Table 1: Standard ArabTeX (consonants)

3.2 Vowels

3.2.1 Long vowels

Table 2 gives the ArabTeX equivalents for the Arabic long vowels.

Letter	Transliteration ¹⁰		ArabTeX notation
	dmg	loc	
ا	ā	ā	A
و	ū	ū	U
ي	ī	ī	I
ى ¹¹	ā	á	_A or Y

¹⁰See below section 6 on page 22.

¹¹= *al-‘alifu ‘l-maqṣūratu*.

Letter	Transliteration		ArabTeX notation
	dmg	loc	
اَ	\bar{a}	\bar{a}	<code>_a</code>
وَ	\bar{u}	\bar{u}	<code>_u</code>
يَ	\bar{i}	\bar{i}	<code>_i</code>

Table 2: Standard ArabTeX (long vowels)

REM. *a*. The long vowels \bar{a} , \bar{u} , \bar{i} , otherwise called *hurūfu 'l-maddi*, the letters of prolongation, involve the placing of the short vowels *a*, *u*, *i* before the letters ا, و, ي respectively. `arabluatex` does that automatically in case any from `voc`, `fullvoc` or `trans` modes is selected e.g. قَالَ $qāla$, قِيلَ $qīla$, يَقُولُ $yaqūlu$.

REM. *b*. Defective writings, such as اَ, *al-'alifu 'l-maḥḍūfatu*, or defective writings of \bar{u} and \bar{i} are encoded `_a _u` and `_i` respectively, e.g. `_d_alika` ذَلِكَ, `al-mal_a'ikaTu` 'l-ra.hm_anu الْمَلِكَةُ الرَّحْمَنُ, `.hu_dayfaTu bnu 'l-yamAn_i` حَذِيفَةُ بْنُ الْيَمَانِ for *Ḥudayfatu bnu 'l-Yamānī*, etc.

3.2.2 Short vowels

Table 3 gives the ArabTeX equivalents for the Arabic short vowels.

Letter	Transliteration ¹²		ArabTeX notation
	dmg	loc	
اَ	<i>a</i>	<i>a</i>	<code>a</code>
وَ	<i>u</i>	<i>u</i>	<code>u</code>
يَ	<i>i</i>	<i>i</i>	<code>i</code>
اَ	<i>an</i>	<i>an</i>	<code>aN</code>
وَ	<i>un</i>	<i>un</i>	<code>uN</code>
يَ	<i>in</i>	<i>in</i>	<code>iN</code>

Table 3: Standard ArabTeX (short vowels)

Whether Arabic texts be vocalized or not is essentially a matter of personal choice. So one may use `voc` mode and decide not to write vowels except at some particular places for disambiguation purposes, or use `novoc` mode, not write vowels—as `novoc` normally does not show them—except, again, where disambiguation is needed.¹³

However, it may be wise to always write the vowels, leaving to the various modes provided by `arabluatex` to take care of showing or not showing the vowels.

¹²See below section 6 on page 22.

¹³See below section 4.4 on page 17.

That said, there is no need to write the short vowels *fatha*, *damma* or *kasra* except in the following cases:

- at the commencement of a word, to indicate that a connective *ʿalif* is needed, with the exception of the article (see below section 4.4 on page 17);
- when arablumatex needs to perform a contextual analysis to determine the carrier of the *hamza*;
- in the various transliteration modes, as vowels are always expressed in romanized Arabic.

4 arablumatex in action

4.1 The vowels and diphthongs

Short vowels As said above, they are written ⟨a, u, i⟩:

_halaqa (or xalaqa) خَلَقَ *halaqa*, ^samsuN شَمْسُ *šamsun*, karImuN كَرِيمُ *Karīmun*.
 bi-hi بِهِ *bi-hi*, 'aqi.tuN أَقِيطُ *'aqiṭun*.
 la-hu لَهُ *la-hu*, .hujjaTuN حُجَّةُ *ḥuḡḡatun*.

Long vowels They are written ⟨U, A, I⟩:

qAla قَالَ *qāla*, bI`a بَيْعَ *bī'a*, .tUruN طُورُ *tūrun*, .tInuN طِينُ *ṭīnun*,
 murU'aTuN مُرُوَّةُ *murū'atun*.

ʿalif maqṣūra It is written ⟨_A⟩ or ⟨Y⟩:

a1-fat_A الْفَاتَى *al-fatā*, a1-maqh_A الْمَقْهَى *al-maqhā*, 'i1_A إِلَى *'ilā*.

ʿalif otiosum Said *ʿalifu l-wiqāyati*, “the guarding *ʿalif*”, after و at the end of a word, both when preceded by *damma* and by *fatha* is written ⟨UA⟩ or ⟨aW, aWA⟩:

na.sarUA نَصَرُوا *naṣarū*, katabUA كَتَبُوا *katabū*, ya.gzUA يَغْزُوا *yaḡzū*,
 ramaW رَمَوْا *ramaw*, banaWA بَنَوْا *banaw*.

ʿalif maḥdūfa and defective ū, ī They are written ⟨_a, _i _u⟩:

a1-l_ahu اللَّهُ *al-lāhu*, 'i1_ahuN إِلَهٌ *'ilāhun*.
 a1-ra.hm_anu الرَّحْمَنُ *ar-raḥmānu*, l_akin لَكِنْ *lākin*, h_ahuna هَاهُنَا *hāhunā*,
 .hunaynu bnu 'is.h_aqa حَنِينُ بْنُ إِسْحَاقَ *Ḥunaynu bnu 'Ishāqa*, rabb_i رَبِّ *rabbī*, a1-`A.s_i الْعَاصِ *al-Āṣī*.

Silent ي/و Some words ending with *ة* are usually written *ة* or *ة* instead of *ة*: see Wright (1896, i. 12 A). arabluatex preserves that particular writing; the same applies to words ending in *ة* for *ة*. Long vowels ⟨U, I⟩ shall receive no *sukūn* after a ‘*alif* *maḥdūfa* and are discarded in *trans* mode:

.hay_aUTuN حَيَوة *hayātun*, .sal_aUTuN صَلَوة *ṣalātun*, mi^sk_aUTuN
مَشْكُوة *miškātun*, tawr_aITuN تَوْرِيَّة *tawrātun*.

And so also: al-rib_aITu الرِّبِيَّة *ar-ribātu*

‘Amrun, and the silent و To that name a silent و is added to distinguish it from ‘*Umarun*: see Wright (1896, i. 12 C). In no way this affects the sound of the *tanwīn*, so it has to be discarded in *trans* mode:

`amruNU عَمْرُو *‘amrun*, `amraNU عَمْرَوا *‘amran*, `amriNU عَمْرٍو *‘amrin*.

When the *tanwīn* falls away (Wright 1896, i. 249 B): `amruU bnu mu.hammaDIN عَمْرُو بْنُ مُحَمَّدٍ *‘Amru bnu Muḥammadin*, mu.hammadu bnu `amriU bni _hAlidIN عَمْرٍو بْنُ خَالِدٍ *Muḥammadu bnu ‘Amri bni Ḥālidin*.

And so also: al-rib_aUA الرِّبَا *ar-ribā*, ribaNU رِبَا *riban*.

tanwīn The marks of doubled short vowels, َ, ِ, ُ are written ⟨uN, aN, iN⟩ respectively. arabluatex deals with special cases, such as َ taking an | after all consonants except ة, and *tanwīn* preceding ى as in هُدًى, which is written ⟨aN_A⟩ or ⟨aNY⟩:

mAluN مَال *mālun*, bAbaN أَبَا *bāban*, madInaTaN مَدِينَة *madīnatan*, bintIN بنت *bintin* maqhaN_A مَقْهَى *maqḥan*, faTaNY فَتًى *fatān*.

arabluatex is aware of special orthographies: ^say'uN شَيْءٌ *šay’un*, ^say'aN شَيْءٍ *šay’in*, ^say'an شَيْئاً *šay’an*, ^say'in شَيْءٍ *šay’in*.

4.2 Other orthographic signs

tā’ marbūṭa It is written ⟨T⟩:

madInaTuN مَدِينَة *madīnatun*, madInaTaN مَدِينَة *madīnatan*, madInaTiN مَدِينَة *madīnatin*.

hamza It is written ⟨'⟩, its carrier being determined by contextual analysis. In case one wishes to bypass this mechanism, he can use the “quoting” feature that is described below in section 4.4 on page 17.

Initial hamza: 'asaduN اَسَدٌ 'asadun, 'u_htuN أُخْتُ 'uhtun, 'iqlIduN اِقْلِيدُ 'iqliḏun, 'anna اِنَّ 'anna, 'inna اِنَّ 'inna.

hamza followed by the long vowel و is encoded '_U: '_U1_A اُولَى 'ulā,
'_U1U اُولُو 'ulū, '_U1A اُولَاكَ 'ulā'ika.

hamza followed by the long vowel ي is encoded '_I: '_ImAnuN اِيْمَانٌ 'imā-nun.

Middle hamza: xA.ti'Ina خَاطِئِينَ ḥāṭi'ina, ru'UsuN رُوُسُ ru'usun,
xa.ti'aTuN خَاطِئَةً ḥaṭi'atun, su'ila سُئِلَ su'ila, 'as'ilaTuN اَسْئَلَةً 'as'i-latun, mas'alaTuN مَسْأَلَةً mas'alatun, 'as'alu اَسْأَلَ 'as'alu, yataSA'a'lUna يَتَسَاءَلُونَ yatasā'alūna, murU'aTuN مُرُوَّةٌ murū'atun, ta'xIruN تَأْخِيرٌ ta'hīrun,
ta'axxara تَأَخَّرَ ta'aḥḥara, ji'tu-ka جِئْتُكَ ġi'tu-ka, qA'iluN قَاتِلٌ qā'ilun.

From Wright (1896, i. 14 B):— All consonants, whatsoever, not even 'alif hēmzatum excepted, admit of being doubled and take *tašdid*. Hence we speak and write ra''AsuN رَأْسٌ ra''āsun, sa''AluN سَأَلَ sa''ālun, na''AjuN نَأَجٌ na''āḡun.

Final hamza: xa.ta'uN خَطَأٌ ḥaṭa'un, xa.ta'aN خَطَأً ḥaṭa'an, xa.ta'in خَطِئًا ḥaṭa'in, 'aqra'u اَقْرَأُ 'aqra'u, taqra'Ina تَقْرَأِينَ taqra'ina, taqra'Una تَقْرَأُونَ taqra'ūna, yaqra'na يَقْرَأْنَ yaqra'na, yaxba'Ani يَخْبَأْنَ yaḥba'āni, xaba'A خَبَأَ ḥaba'a, xubi'a خُبِيَ ḥubi'a, xubi'UA خُبُوا ḥubi'ū, jA'a جَاءَ ġā'a, ridA'uN رَدَأَ ridā'un, ridA'aN رَدَأَ ridā'an, jI'a جِئَ ġī'a, radI'in رَدِيَ radī'in, sU'uN سُوءٌ sū'un, .daw'uN دَوَّءٌ daw'un, qay'in قَيَّ qay'in.

^say'uN سَيَّ say'un, ^say'aN سَيَّ say'an, ^say'in سَيَّ say'in, al-^say'u أَشْيَاءُ aš-šay'u, 'a^syA'u أَشْيَاءُ ašyā'u, 'a^syA'a أَشْيَاءُ ašyā'a, .zim'aN ظِمًّا zim'an.

taḥfīfu 'l-hamzati: if the hamza has ḡazma and is preceded by 'alifhamzatum, it must be changed into the letter of prolongation that is homogeneous with the preceding vowel; hence: 'a'mana آمَنَ 'āmana, 'u'minu أُؤْمِنُ 'ūminu, 'i'mAnuN اِيْمَانٌ 'imānun. For other possible ways of encod-

ing such sequences, see on page 12 (*hamza* followed by و and ي) and the *madda* on this page.

madda At the beginning of a syllable, 'alif with *hamza* and *fatḥa* (أ) followed by 'alifu 'l-maddi ('alif of prolongation) or 'alif with *hamza* and *ḡazma* (إ) are both represented in writing 'alif with *madda*: آ (see Wright 1896, i. 25 A–B).

Hence one should keep to this distinction and encode 'a 'ku lu أَكُلُ 'ākulu and 'Ak iluN أَكِلُ 'ākilun respectively.

arabluatex otherwise determines *al-'alifu 'l-mamdūdatu* by context analysis.

'is 'AduN إِسَادُ 'is'ādun, 'Ak ilUna أَكِلُونُ 'ākilūna, 'a 'mannA آمَنَّا 'āmannā,
al-qur' Anu الْقُرْآنُ al-qur'ānu.
jA 'a جَاءَ ḡā'a, yataSA 'alUna يَتَسَاءَلُونَ yatasā'alūna, riDA 'uN رِدَاءُ ridā'un,
xaba 'A خَبَا ḥaba'ā, yaxba 'Ani يَخْبَانُ yaḥba'āni.

šadda *tašdīd* is either *necessary* or *euphonic*.

The necessary *tašdīd* always follows a vowel, whether short or long (see Wright 1896, i. 15 A–B). It is encoded in writing the consonant that carries it twice:

`allaqa عَقَّى 'allaqa, mAdduN مَادَّ māddun, 'ammara أَمَّرَ ammara, murrUN مَرَّرَ murrun.

The euphonic *tašdīd* always follows a vowelless consonant which is passed over in pronunciation and assimilated to a following consonant. It may be found (Wright 1896, i. 15 B–16 C):—

(a) With the *solar* letters ت, ث, د, ذ, ر, ز, س, ش, ص, ض, ط, ظ, ل, ن, after the article اَلْ:—

Unlike arabtex and arabxetex, arabluatex *never requires the solar letter to be written twice*, as it automatically generates the euphonic *tašdīd* above the letter that carries it, whether the article be written in the assimilated form or not, e.g. al-^samsu الشَّمْسُ aš-šamsu, or a^s-^samsu الشَّمْسُ aš-šamsu.

al-tamru اَلْتَمَرُ at-tamru, al-ra . hm_anu الرَّحْمَنُ ar-raḥmānu, al- . zulmu اَلْظُلْمُ az-ẓulmu, al-lu . gaTu اَللُّغَةُ al-luḡatu.

(b) With the letters ر, ل, م, و, ي after ن with *ḡazma*, and also after the *tanwīn*:—

Note the absence of *sukūn* above the passed over ن in the following examples, each of which is accompanied by a consistent transliteration:
 min rabbi-hi مِنْ رَبِّهِ, *mir rabbi-hi*, min laylin مِنْ لَيْلٍ *mil laylin*, 'an yaqtula أَنْ يَقْتُلَ 'ay yaqtula.

With *tanwīn*: kitābun mubīn كِتَابٌ مُبِينٌ *kitābum mubīnun*.

- (c) With the letter ت after the dentals ث, د, ذ, ض, ط, ظ in certain parts of the verb: this kind of assimilation, e.g. لَبِثْتُ for لَبِثْتُ *labittu*, will be discarded here, as it is largely condemned by the grammarians (see Wright 1896, i. 16 B–C).

The definite article and the 'alifu 'l-waṣli At the beginning of a sentence, ا is never written, as اَلْحَمْدُ لِلّٰهِ; instead, to indicate that the 'alif is a connective 'alif ('alifu 'l-waṣli), the *hamza* is omitted and only its accompanying vowel is expressed:

al-.hamdu li-l-l_ahi اَلْحَمْدُ لِلّٰهِ *al-ḥamdu li-llāhi*.

As said above on page 5, *fullvoc* is the mode in which *arabluatex* expresses the *sukūn* and the *waṣla*. *arabluatex* will take care of doing that automatically provided that the vowel which is to be absorbed by the final vowel of the preceding word is properly encoded, like so:—

- (a) Definite article at the beginning of a sentence is encoded
 al- , or a<solar letter>-
 if one wishes to mark the assimilation—which is in no way required, as *arabluatex* will detect all cases of assimilation.
- (b) Definite article inside sentences is encoded
 'l- or '<solar letter>- .
- (c) In all remaining cases of elision, the 'alifu 'l-waṣli is expressed by the vowel that accompanies the omitted *hamza*: ⟨u, a, i⟩.

Article: bAbu 'l-madrasaTi بَابُ الْمَدْرَسَةِ *bābu 'l-madrasati*, al-maqālaTu 'l-'_Ul_A الْمَقَالَةُ الْأُولَى *al-maqālatu 'l-'ūlā*, al-lu.gaTu 'l-'arabiyyaTu فِي صِنَاعَةِ اللُّغَةِ الْعَرَبِيَّةِ *al-luḡatu 'l-'arabiyyatu*, fI .sinā`aTi 'l-.tibbi فِي صِنَاعَةِ الْإِنْتِقَاضِ إِلَى الْإِبْتِدَاءِ *fi 'l-ibtidā'i*, 'abU 'l-wazIri أَبُو الْوَزِيرِ *abu 'l-wazīri*, fa-lamma ra'aW 'l-najma فَلَمَّا رَأَوْا النَّجْمَ *fa-lammā ra'awu 'l-najma*.

Particles:—

- (a) *li-*: 'alifu 'l-waṣli is omitted in the article اَلْ when it is preceded by the preposition لِ: li-l-rajuli لِلرَّجُلِ *li-r-raḡuli*.
 If the first letter of the noun be ل, then the ل of the article also falls away, but *arabluatex* is aware of that: li-l-laylaTi لِلَّيْلِ *li-l-laylati*.

- (b) *la*:- the same applies to the affirmative particle لَ: 1a-1- . haqqu لَحَقَّ *la-l-haqqu*.
- (c) With the other particles, 'alifu 'l-wašli is expressed: fI 'l-madInaTi فِي الْمَدِينَةِ *fi 'l-madīnati*, wa- 'l-rajulu وَالرَّجُلُ *wa-'r-raġulu*, bi- 'l-qalami بِالْقَلَمِ *bi-'l-qalami*, bi- 'l-ru`bi بِالرُّعْبِ *bi-'r-ru'bi*.

Perfect active, imperative, nomen actionis: qAla isma قَالَ اسْمَعْ *qāla 'sma*, qAla uqtul قَالَ أَقْتُلْ *qāla 'qtul*, huwa inhazama هُوَ أَنهَزَمَ *huwa 'nhazama*, wa-ustu`mila وَأَسْتَعْمِلَ *wa-'štu'mila*, qad-i in.sarafa قَدْ أَنْصَرَفَ *qadi 'nšarafa*, al-iqtidAru الْأَقْتِدَارُ *al-iqtidāru*, 'il_A 'l-intiqA.di إِلَى الْإِنْتِقَاضِ *'ila 'l-intiqādi*, lawi istaqbala لَوْ اسْتَقْبَلَ *lawi 'štaqbala*.

Other cases: 'awi ismu-hu أَوْ اسْمُهُ *'awi 'smu-hu*, zayduN ibnu `amriNU زَيْدُ بْنُ عَمْرٍو *Zaydun ibnu 'Amrin*,¹⁴ imru'u 'l-qaysi إِمْرُؤُ الْقَيْسِ *Imru'u 'l-Qaysi*, la-aymunu 'l-l_ahi لَا يُؤْمِنُ اللَّهُ *la-'ymunu 'l-lāhi*.

'alifu 'l-wašli preceded by a long vowel The long vowel preceding the connective 'alif is shortened in pronunciation (Wright 1896, i. 21 B–D). This does not appear in the Arabic script, but arabluatex takes it into account in some transliteration standards:—

fI 'l-nAsi فِي النَّاسِ *fi 'n-nāsi*, 'abU 'l-wazIri أَبُو الْوَزِيرِ *'abu 'l-wazīri*, fI 'l-ibtidA 'i الْإِبْتِدَاءُ *fi 'l-ibtidā'i*, _dU 'l-i`lAlI ذُو الْأَعْلَالِ *du 'l-i'lāli*.

'alifu 'l-wašli preceded by a diphthong The diphthong is resolved into two simple vowels (Wright 1896, i. 21 D–22 A) viz. *ay* → *āi* and *aw* → *āū*. arabluatex detects the cases in which this rule applies:—

fI `aynay 'l-maliki فِي عَيْنِي الْمَلِكِ *fi 'aynaya 'l-maliki*, ix^say 'l-qawma إِيْخْسَى الْقَوْمِ *iḥšayi 'l-qawma*, mu.s.tafaw 'l-l_ahi مُصْطَفَوُ اللَّهِ *muṣṭafawu 'l-lāhi*.
ramaW 'l-.hiJAraTa رَمَوْا الْحِجَارَةَ *ramawu 'l-ḥiġārata*, fa-lamma ra'aW 'l-najma فَلَمَّا رَأَوْا النَّجْمَ *fa-lammā ra'awu 'n-naġma*.

'alifu 'l-wašli preceded by a consonant with sukūn The vowel which the consonant takes is either its original vowel, or that which belongs to the connective 'alif or the *kasra*; in most of the cases (Wright 1896, i. 22 A–C), it is encoded explicitly, like so:—

¹⁴“Zayd is the son of ‘Amr”: the second noun is not in apposition to the first, but forms part of the predicate. Hence زَيْدُ بْنُ عَمْرٍو and not زَيْدُ بْنُ عَمْرٍو, “Zayd, son of ‘Amr”.

'antumū 'l-kA_dibUnā اَنْتُمْ الْكَادِبُونَ 'antumū 'l-kāḍibūna, ra'aytumū
'l-rajula رَاَيْتُمُ الرَّجُلَ ra'aytumū 'r-raḡula, mani 'l-ka_d_dAbu مَنِ
الْكَذَّابُ mani 'l-kaddābu, qatalati 'l-rUmu قَتَلَتِ الرُّومَ qatalati 'r-Rūmu.

However, the Arabic script does not show the *kasra* which is taken by the nouns having *tanwīn* although it is explicit in pronunciation and must appear in some transliteration standards. arabluatex takes care of that automatically:—

mu.hammaduN 'l-nabI مُحَمَّدُ النَّبِيّ Muḥammaduni 'n-nabī.

4.3 Special orthographies

The name of God The name of God, اَللّٰهُ, is compounded of the article اَلْ, and اِلَٰهٌ (noted اِلهٌ with the defective 'alif) so that it becomes اِلَٰلَٰه; then the *hamza* is suppressed, its vowel being transferred to the ل before it, so that there remains اَلله (I refer to Lane, *Lexicon*, I. 83 col. 1). Finally, the first ل is made quiescent and incorporated into the other, hence the *tašdīd* above it. As arabluatex never requires a solar letter to be written twice (see above, on page 13), the name of God is therefore encoded a1-l_ahu or 'l-l_ahu:—

a1-l_ahu اَلله al-lāhu, yA¹⁵ a1-l_ahu يَ اَلله yā al-lāhu, 'a-fa¹⁶-a1-l_ahi
la-ta.g`alanna اَلله لَتَعْلَنَّ 'a-fa-al-lāhi la-tag`alanna, bi-'l-l_ahi
بِ اَلله bi-'l-lāhi, wa-'l-l_ahi وَاَلله wa-'l-lāhi, bi-smi 'l-l_ahi بِسْمِ اَلله bi-
smi 'l-lāhi, a1-. hamdu li-l-l_ahi اَلْحَمْدُ لِلّٰهِ al-ḥamdu li-llāhi, li-l-l_ahi
'l-qA' ilu اَللّٰهُ الْقَائِلُ li-llāhi 'l-qā'ilu.

The conjunctive اَلَّذِي Although it is compounded of the article اَلْ, the demonstrative letter ل and the demonstrative pronoun ذَا, both masculine and feminine forms that are written defectively are encoded alla_dI and allatI respectively. Forms starting with the connective 'alif are encoded 'lla_dI and 'llatI:—

'a_hAfu mina 'l-maliki 'lla_dI ya.zlimu 'l-nAsa مِنْ اَخَافُ
اَهَافُو مِنا اَلْمَلِكِ الَّذِي يَظْلِمُ النَّاسَ 'llaḍī yaẓlimu 'n-nāsa, `udtu
'l-^say_ha 'lladI huwa marI.duN اَلَّذِي هُوَ مَرِيضٌ udtu
'š-šayḥa 'lladī huwa marīḍun, mA 'anA bi-'lla_dI qA'iluN لا-ka
^say'an اَنَا الَّذِي قَائِلٌ لَكَ شَيْئًا mā 'anā bi-'llaḍī qā'ilul la-ka šay'an.

¹⁵Note the “pipe” character ‘|’ here after yA and below after fa before footnote mark 16: it is needed by the dmG transliteration mode as in this mode any vowel at the commencement of a word preceded by a word that ends with a vowel, either short or long, is absorbed by this vowel viz. ‘ala ‘t-ṭarīqi. See section 4.5 on page 19 on the “pipe” and section 6 on page 22 on dmG mode.

¹⁶See footnote 15.

'ari-na 'lla_dayni 'a.dalla-nAmina 'l-jinni wa-'l-'insi
 أَرِنَا الَّذِينَ أَضَلَّانَا مِنَ الْجِنَّ وَالْإِنْسِ
 'insi.

The other forms are encoded regularly as a1-l or 'l-l:—

fa-'inna na_dkuru 'l-. sawtayni 'l-la_dayni rawayna-humA
 ۞an ja.h.zaTa فَإِنَّا نَذْكُرُ الصَّوْتَيْنِ اللَّذَيْنِ رَوَيْنَاهُمَا عَنْ حِطَّةٍ
 sawtayni 'l-la_dayni rawaynā-humā ۞an Ġaḥḥata.

And also: al-la_dAni اللَّذَانِ al-la_dāni, al-la_dayni اللَّذَيْنِ al-la_dayni,
 al-latAni اللَّتَانِ al-latāni, al-latayni اللَّتَيْنِ al-latayni, al-latI اللَّاتِي al-
 lātī, al-lA' | Ati¹⁷ اللَّاءَاتِ al-lā'āti, al-lA' I اللَّاءِي al-lā'i, and so forth.

4.4 Quoting

It is here referred to “quoting” after the package arabtex.¹⁸ The “quoting” mechanism of arabluatex is designed to be very similar in effect to the one of arabtex.

To start with an example, suppose one types the following in novoc mode: علم علم
 الهيئة; is it ۞لَّم, he was taught the science of astronomy, or ۞لَّم, he taught the science of as-
 tronomy? In order to disambiguate this clause, it may be sensible to put a damma above
 the first ۞لَّم, which is achieved by “quoting” the vowel u, like so: ۞"ullima,
 or, with no other vowel than the required u: ۞"ullm.

This is how the “quoting” mechanism works: metaphorically speaking, it acts as a
 toggle switch. If something, in a given mode, is supposed to be visible, “quoting” hides
 it; conversely, if it is supposed not to, it makes it visible.

As shown above, “quoting” means inserting one straight double quote (") before the
 letter that is to be acted upon. Its effects depend on the mode which is currently selected,
 either novoc, voc or fullvoc:—

novoc In this mode, “quoting” essentially means make visible something that ought
 not to be so.

(a) Quoting a vowel, either short or long, makes the damma, fatha or kasra appear above
 the appropriate consonant:—

۞"ullima ۞ilma 'l-hay'aTi ۞لَّم ۞لَّم الهيئة ۞ullima ۞ilma 'l-hay'ati,
 ya.gz"UA يَغْزُوا yağzū.

(b) The same applies when “quoting” the tanwīn:—

¹⁷Note here the “pipe” character '|’: as already stated on page 13, the sequence 'A usually encodes 'alif with
 hamza followed by 'alif of prolongation, which is represented in writing 'alif with madda: ۞. The “pipe” charac-
 ter prevents this rule from being applied. See section 4.5 on page 19.

¹⁸See Lagally (2004, p. 22)

wa-'inna sawfa tudriku-na 'l-manāyā muqadd"araT"an وإنا سوف ندرِكُ المنايا مقدِّرةً
wa-'innā sawfa tudriku-na 'l-manāyā muqaddaratan.

- (c) If no vowel follows the straight double quote, then a *sukūn* is put above the preceding consonant:—

qAla isma`" قال اسمعُ qāla 'sma', jA'at" hinduN جَاءَتْ هِنْدُ gā'at
Hindun, ^sabIhuN bi-man q"u.ti`at" qadamA-hu شَبِيهَ بِنِ قُطِعَتْ
šābihun bi-man quṭi'at qadamā-hu.

- (d) At the commencement of a word, the straight double quote is interpreted as 'alifu 'l-waṣli:—

wa-"ust"u`mila وأَسْتَعْمَلُ wa-'stu'mila, huwa "inhazama هو أَنَهَزَمَ huwa
'nhazama, al-"intiqA.du الِإِنْتِقَاضُ al-intiqādu.

voc In accordance with the general rule, in this mode, “quoting” makes the vowels and the *tanwīn* disappear, should this feature be required for some reason:—

- (a) Short and long vowels:—

q"Ala q"A' iluN قَالَ قَائِلٌ qāla qā'ilun, ibnu 'abI 'u.saybi`aT"a
إِبْنُ أَبِي أُصَيْبَةَ Ibnu 'Abī 'Uṣaybi'ata.

- (b) *tanwīn*:—

madInaT"aN مَدِينَةٌ madīnatan, bAb"aN أَبَا bāban, hud"aN_A هُدًى hudan,
^say'"iN شَيْءٌ šay'in.

One may more usefully “quote” the initial vowels to write the *waṣla* above the 'alif or insert a straight double quote after a consonant not followed by a vowel to make the *sukūn* appear:—

- (a) 'alifu 'l-waṣli:—

fI "istisqA'a أَسْتِسْقَاءُ فِي fi 'stisqā'a, wa-"istisqA'u وَأَسْتِسْقَاءُ wa-'stisqā'u,
qAla "uhrub fa-lan tuqtala قَالَ أَهْرُبُ فَلَنْ تُقْتَلَ qāla 'hrub fa-lan tuq-
tala.

- (b) *sukūn*:—

qAla "uqtul" fa-lan tuqtala قَالَ أَقْتُلْ فَلَنْ تُقْتَلَ qāla 'qtul fa-lan tuq-
tala, mA jA'at" mini imra'aTIN مِمَّا جَاءَتْ مِنْ إِمْرَأَةٍ mā gā'at mini
'mra'atin, kam" qad"ma.dat" min" laylaTiN كَمْ قَدْ مَضَتْ مِنْ لَيْلَةٍ kam
qad maḍat min laylatin.

fullvoc In this mode, “quoting” may be used to take away any short vowel (or *tanwīn*, as seen above) or any *sukūn*:—

al-jamru 'l-.sayfiyyu 'lla_dI kAna bi-q"rAn" | nUna الجمر الصّيفيّ
 al-ḡamru 'š-šayfiyyu 'llaḍi kāna bi-Qrānnūna. كَان بَقْرَانُون

4.4.1 Quoting the *hamza*

As said above in section 4.2 on page 12, the *hamza* is always written ('), its carrier being determined by contextual analysis. “Quoting” that straight single quote character like so: (' ') allows to determine the carrier of the *hamza* freely, without any consideration for the context. Table 4 gives the equivalents for all the possible carriers the *hamza* may take:

Letter	Transliteration ¹⁹		ArabTeX notation
	dmg	loc	
آ	ʾ	ʾ	" ' "
ا	ʾā	ʾā	A " ' "
إ	ʾ	ʾ	a " ' "
أ	ʾ	ʾ	u " ' "
ؤ	ʾ	ʾ	w " ' "
إِ	ʾ	ʾ	i " ' "
ئ	ʾ	ʾ	y " ' "

Table 4: “Quoted” *hamza*

As one can see from table 4, the carrier of the *hamza* is inferred from the letter that precedes the straight double quote ("). Of course, any "quoted" *hamza* may take a short vowel, which is to be written *after* the ArabTeX equivalent for the *hamza* itself, namely ('). For example, أَ is encoded (w" ' a), while أُ is encoded (w" ' '). In the latter example, the second straight double quote encodes the *sukūn* in voc mode in accordance with the rule laid above on page 18.

'a`dA'ukum اَعْدَاؤُكُمْ 'a`dā'ukum, 'a`dA|"'ukum اَعْدَاءُكُمْ 'a`dā'ukum, 'a`dA'ikum اَعْدَائِكُمْ 'a`dā'ikum, 'a`dA|"'"ikum اَعْدَاءُكُمْ 'a`dā'ikum.

4.5 The “pipe” character (|)

In the terminology of Arab_TE_X, the “pipe” character ‘|’ is referred to as the “invisible consonant”. Hence, as already seen above in section 4.4.1, its usage to encode the *hamza* alone, with no carrier: | " ' ء.

Aside from that usage, the “pipe” character is used to prevent almost any of the contextual analysis rules that are described above from being applied. Two examples have already been given to demonstrate how that particular mechanism works in footnote 15 on page 16 and in footnote 17 on page 17. One more example follows:—

¹⁹See below section 6 on page 22.

bi-qraN|nUna بِقْرَانُونِ *bi-Qrānnūna*, “in Crannon” (Thessaly, Greece).²⁰

As one can see, the “pipe” character between the two $\langle n \rangle$ prevents the necessary *tašdīd* rule (page 13) from being applied.

4.6 Stretching characters: the *tatwīl*

A double hyphen $\langle -- \rangle$ stretches the ligature in which one letter is bound to another. Although it is always better to rely on automatic stretching, this technique may be used to a modest extent, especially to increase legibility of letters and diacritics which stand one above the other:—

.hunaynu bnu 'is.h--_aqa حُنَيْنُ بْنُ إِسْحَاقَ *Hunaynu bnu 'Ishāqa*

4.7 Digits

4.7.1 Numerical figures

The *Indian numbers*, *ar-raqamu 'l-hindiyyu*, are ten in number, and they are compounded in exactly the same way as our numerals:—

1874 ١٨٧٤, 123-456, 789 ١٢٣-٤٥٦, ٧٨٩, fI sanaTi 1024 ١٠٢٤ فِي سَنَةِ

4.7.2 The *abjad*

The numbers may also be expressed with letters from right to left arranged in accordance with the order of the Hebrew and Aramaic alphabets (see Wright 1896, i. 28 B–C). The *'abjad* numbers are usually distinguished from the surrounding words by a stroke placed over them.

`\abjad` *'abjad* numbers are inserted with the command `\abjad{<number>}` in any of the `voc`, `fullvoc` and `novoc` modes, where $\langle number \rangle$ may be any number between 1 and 1999, like so:—

`\abjad{45}` kitAbu-hu fI 'l-`AdAt i مَكَّابُهُ فِي الْعَادَاتِ 45 *kitābu-hu fi 'l-`ādāti*.

REM. a. As can be seen in the above given example, *arabuatex* expresses the *'abjad* numbers in Roman numerals if it finds the command `\abjad` in any of the transliteration modes.

REM. b. `\abjad` may also be found outside Arabic environments. In that case, *arabuatex* does not print the stroke as a distinctive mark over the number for it is not surrounded by other Arabic words. In case one nonetheless wishes to print the stroke, he can use the `\aemph` command that is described below in section 4.9 on the following page:—

The `\arb[trans]{ 'abjad}` number for 1874 is `\abjad{1874}` The *'abjad* number for 1874 is غَضْعَد.

The `\arb[trans]{ 'abjad}` number for 1874 is `\aemph{\abjad{1874}}` The *'abjad* number for 1874 is غَضْعَد.

²⁰See more context on page 19.

4.8 Additional characters

In the manuscripts, the unpointed letters, *al-ḥurūfu 'l-muḥmalatu*, are sometimes further distinguished from the pointed by various contrivances, as explained in Wright (1896, i. 4 B–C). One may find these letters written in a smaller size below the line, or with a dot or another mark below. As representing all the possible contrivances leads to much complexity and also needs to be agreed among scholars, new ways of encoding them will be proposed and gradually included as arabluatex will mature.

For the time being, the following is included:—

Letter	Transliteration ²¹		ArabTeX notation
	dmg	loc	
ب	<i>b</i>	<i>b</i>	.b
د	<i>d</i>	<i>d</i>	^d
ف	<i>f</i>	<i>f</i>	.f
ق	<i>q</i>	<i>q</i>	.q
ك	<i>k</i>	<i>k</i>	.k
ن	<i>n</i>	<i>n</i>	.n
﴿	((((
﴾))))

Table 5: Additional Arabic codings

'afAman . tUs Gal.(M) . fmn . n . ts (sic) Gal.(E1), أفامنطوس Gal.(M) فنطس (sic) Gal.(E1), 'afāmanṭūs Gal.(M) fmnṭs (sic) Gal.(E1).

4.9 Arabic emphasis

As already seen in section 4.7.2 on page 20, the 'abḡad numbers are distinguished from the surrounding words by a stroke placed over them. This technique is used to distinguish further words that are proper names or book titles.

\aemph One may use the command \aemph{⟨Arabic text⟩} to use the same technique to emphasize words, like so:—

\abjad{45} : kitAbu-hu \aemph{fI 'l-`AdAt i} كِتَابُهُ فِي الْعَادَاتِ 45:م
kitābu-hu fi 'l-`Ādāti.

5 Special applications

Linguistics The same horizontal stroke as the *tatwīl* (see section 4.6 on page 20) may be encoded ⟨B⟩; ⟨BB⟩ will receive the *tašdīd*. This is useful to make linguistic annotations and comments on vowels:—

²¹See below section 6 on the next page.

Bu Ba Bi BuN BaN BiN----- u a i un an in, BBu BBa BBi----- u a i, B--aN
 ُ an, B ..

6 Transliteration

It may be more appropriate to speak of “romanization” than “transliteration” of Arabic. As seen above in section 2.2 on pages 5–7, the “transliteration mode” may be selected globally or locally.

This mode transliterates the ArabTeX input into one of the accepted standards. As said above on page 5, two standards are supported at present:

dmg *Deutsche Morgenländische Gesellschaft*, which was adopted by the International Convention of Orientalist Scholars in Rome in 1935.²² dmg transliteration convention is selected by default;

loc *Library of Congress*: this standard is part of a large set of standards for romanization of non-roman scripts adopted by the American Library Association and the Library of Congress.²³

More standards will be included in future releases of arabluatex.

`\SetTranslitConvention` **Convention** The transliteration mode, which is set to dmg by default, may be changed at any point of the document by the command `\SetTranslitConvention{<mode>}`, where `<mode>` may be either dmg or loc. This command is also accepted in the preamble should one wish to set the transliteration mode globally, eg.:-

```
1 \usepackage{arabluatex}
2 \SetTranslitConvention{loc}
```

`\SetTranslitStyle` **Style** Any transliterated Arabic text is printed in italics by default. This also can be changed either globally in the preamble or locally at any point of the document by the command `\SetTranslitStyle{<style>}`, where `<style>` may be any font shape selection command, eg. `\upshape`, `\itshape`, `\slshape`, and so forth. Any specific font may also be selected using the font-selecting commands of the fontspec package.

`\cap` **Proper names** Proper names or book titles that must have their first letters uppercased may be passed as arguments to the command `\cap{<word>}`. `\cap` is a clever command, for it will give the definite article *al-* in lower case in all positions. Moreover, if the initial letter, apart from the article, cannot be uppercased, viz. ‘ or ‘, the letter next to it will be uppercased:-

`\cap{.hunaynu}` bnu `\cap{'is.h_aqa}` حَنِينُ بْنُ إِسْحَقَ Hunaynu bnu
 ‘Ishāqa, `\cap{`u_tm_anu}` عُثْمَنُ ‘Utmānu, .daraba `\cap{zaydu}` bnu

²²See Brockelmann et al. (1935).

²³See <http://www.loc.gov/catdir/cpsd/roman.html> for the source document concerning Arabic language.

`\cap{_h_alidiN} \cap{sa`da} bna \cap{`awfi} bni \cap{`abdi}`
`\cap{'1-1_ahi} ضَرْبُ زَيْدُ بْنُ خَلْدٍ سَعْدُ بْنُ عَوْفِ بْنِ عَبْدِ اللَّهِ` *daraba Zaydu bnu*
Ḥālidin Sa'da bna 'Awfi bni 'Abdi 'l-Lāhi.

However, `\cap` must be used cautiously in some very particular cases, for the closing brace of its argument may prevent a rule from being applied. To take an example, as seen above on page 16, the transliteration of مُحَمَّدُ النَّبِيُّ must be *Muḥammaduni 'n-nabī*, as nouns having the *tanwīn* take a *kasra* in pronunciation before *'alifu 'l-waṣli*. In that case, encoding مُحَمَّد like so: `\cap{mu . hammaduN}` is wrong, because the closing brace would prevent arabuatex from detecting the sequence $\langle -uN \rangle$ immediately followed by $\langle 'l \rangle$. Fortunately, this can be circumvented in a straightforward way by inserting only part of the noun in the argument of `\cap` vz. up to the first letter that is to be uppercased, like so: `\cap{m}u . hammaduN`.

6.1 Examples

Here follows in transliteration the story of Ġuḥā and his donkey (جُحَا وَحَمَارُهُ). See the code on page 6:—

'dmg' standard: `'atā ṣadīqun 'ilā Ġuḥā yaṭlubu min-hu ḥimāra-hu li-yarkaba-hu fī safratin qaṣīratin. wa-qāla la-hu: "sawfa 'u'īdu-hu 'ilay-ka fī 'l-masā'i wa-'adfa'u la-ka 'uḡratan."` *fa-qāla Ġuḥā: "anā 'āsifun ḡiddan 'annī lā 'aṣṭaṭī'u 'an 'uḥaqqiqa la-ka ḡarbata-ka fa-'l-ḥimāru laysa huna 'l-yawma."* `wa-qabla 'ay yutimma Ġuḥā kalāma-hu bada'a 'l-ḥimāru yanhaqu fī 'iṣṭabili-hi. fa-qāla la-hu ṣadīqu-hu: "innī 'asma'u ḥimāra-ka yā Ġuḥā yanhaqu."` *fa-qāla la-hu Ġuḥā: "ḡarībun 'amru-ka yā ṣadīqī 'a-tuṣaddiqu 'l-ḥimāra wa-tukadḏibu-nī?"*

'loc' standard: `atā ṣadīqun ilā Juḥā yaṭlubu min-hu ḥimāra-hu li-yarkaba-hu fī safratin qaṣīratin. wa-qāla la-hu: "sawfa u'īdu-hu ilay-ka fī al-masā'i wa-'adfa'u la-ka ujratan."` *fa-qāla Juḥā: "anā 'āsifun jiddan annī lā aṣṭaṭī'u an uḥaqqiqa la-ka gharbata-ka fa-al-ḥimāru laysa hunā al-yawma."* `wa-qabla an yutimma Juḥā kalāma-hu bada'a al-ḥimāru yanhaqu fī iṣṭabili-hi. fa-qāla la-hu ṣadīqu-hu: "innī asma'u ḥimāra-ka yā Juḥā yanhaqu."` *fa-qāla la-hu Juḥā: "gharībun amru-ka yā ṣadīqī a-tuṣaddiqu al-ḥimāra wa-tukadhdhibu-nī?"*

7 L^AT_EX Commands in Arabic environments

General principle L^AT_EX commands are accepted in Arabic environments. The general principle which applies is that single-argument commands (`\command{\langle arg \rangle}`) such as `\emph{\langle text \rangle}`, `\textbf{\langle text \rangle}` and the like, are assumed to have Arabic text as their arguments:—

`\abjad{45} kitAbu-hu \emph{fī 'l-\cap{`AdAti}}` مَهْ كِتَابُهُ فِي لِـعَادَاتِ
 45 kitābu-hu fī 'l-Ādāti.²⁴

²⁴This is odd in Arabic script, but using such features as `\emph` or `\textbf` is a matter of personal taste.

The same applies to footnotes:—

```

1 \renewcommand{\footnoterule}%
2   {\hfill\noindent\rule[1mm]{.4\textwidth}{.15mm}}
3 \begin{arab}
4 'inna 'abI kAna mina 'l-muqAtilaTi\footnote{al-muqAtilaTi:
5 al-muqAtilIna.}, wa-kAnat 'ummI min `u.zamA'i buyUti
6 'l-zamAzimaTi\footnote{al-zamAzimaTu: .tA'ifaTu mina
7 'l-fursi.}.
8 \end{arab}

```

إِنَّ أَيَّ كَانَ مِنَ الْمُقَاتِلَةِ، وَكَانَتْ أُمِّي مِنْ عُظَمَاءِ بُيُوتِ الزَّمَاةِ -

«الْمُقَاتِلَةُ: الْمُقَاتِلِينَ -
الزَّمَاةُ: طَائِفَةٌ مِنَ الْفُرْسِ -

Some commands, however, do not expect running text in their arguments, or one may wish to insert English text eg. in footnotes or in marginal notes. arabluatex provides a set of commands to handle such cases.

\LR `\LR{<arg>}` is designed to typeset its argument from left to right. It may be used in an Arabic environment, either `\arb{<Arabic text>}` or `\begin{arab} <Arabic text> \end{arab}`, for short insertions of left-to-right text, or to insert any \TeX command that would otherwise be rejected by arabluatex, such as commands the argument of which is expected to be a dimension or a unit of measurement.

\RL `\RL{<arg>}` does the same as `\LR{<arg>}`, but typesets its argument from right to left. Even in an Arabic environment, this command may be useful.

For example, to distinguish words with a different color, one may proceed like so:—

```

1 \begin{arab}
2 _tumma "intalaqa _dU 'l-qarnayni 'il_A 'ummaTiN 'u_hr_A fI
3 \LR{\textcolor{red}{\arb[fullvoc]{((ma.tli`i 'l-^samsi)}}}}
4 wa-lA binA'a la-hum yu'amminu-hum mina 'l-^samsi.
5 \end{arab}

```

ثُمَّ اتَّخَذَ ذُو الْقَرْنَيْنِ إِلَى أُمَّةٍ أُخْرَى فِي «مَطْلَعِ الشَّمْسِ» وَلَا بِنَاءَ لَهُمْ يَوْمَئِذٍ مِنَ الشَّمْسِ -

\LRfootnote `\LRfootnote{<text>}` and **\RLfootnote** `\RLfootnote{<text>}` typeset left-to-right and right-to-left footnotes respectively in Arabic environments. Unlike `\footnote{<text>}`, the arguments of both `\LRfootnote` and `\RLfootnote` are not expected to be Arabic text. For example, `\LRfootnote` may be used to insert English footnotes in running Arabic text:—


```

1 \arb[fullvoc]{\cap{z}ayduN\LRfootnote{\enquote{Zayd
2 is the son of Amr}: the second noun is not in
3 apposition to the first, but forms part of
4 the predicate\ldots} "ibnu \cap{`amriNU}}

```

زَيْدُ ابْنِ عَمْرٍو

[#]Zayd is the son of ‘Amr’: the second noun is not in apposition to the first, but forms part of the predicate...

When footnotes are typeset from right to left, it may happen that the numbers of the footnotes that are at the bottom of the page be typeset in the wrong direction. For example, instead of an expected number 18, one may get 81. arabluatex is not responsible for that, but should it happen, it may be necessary to redefine in the preamble the \LaTeX macro `\thefootnote` like so:—

```
\renewcommand*{\thefootnote}{\textsuperscript{\LR{\arabic{footnote}}}}
```

`\FixArbFtnmk`

Another solution is to put in the preamble, below the line that loads arabluatex, the command `\FixArbFtnmk`. However, for more control over the layout of footnotes marks, it is advisable to use the package `scrextend`.²⁵

`\LRmarginpar`

The command `\LRmarginpar` does for marginal notes the same as `\LRfootnote` does for footnotes. Of course, it is supposed to be used in Arabic environments. Note that `\marginpar` also works in Arabic environments, but it acts as any other single-argument command inserted in Arabic environments. The general principle laid on page 23 applies.

`\setRL`

`\setRL` and `\setLR` may be used to change the direction of paragraphs, either from left to right or from right to left. As an example, an easy way to typeset a right-to-left sectional title follows:—

`\setLR`

```

1 \setRL
2 \section*{\arb{barzawayhi li-buzurjumihra bni 'l-buxtikAni}}
3 \setLR
4 \begin{arab}
5 qAla barzawayhi bnu 'azhara, ra'su 'a.tibba'i fArisa...
6 \end{arab}

```

بَرْزَوِيهِ لِبُزْرَجُمِهْرَ بْنِ الْبُخْتِكَانِ
قَالَ بَرْزَوِيهِ بْنُ أَزْهَرَ، رَأْسُ أَطِبَّاءِ فَارِسٍ---

²⁵See <http://ctan.org/pkg/koma-script>; read the documentation of KOMA-script for details about the `\deffootnotemark` and `\deffootnote` commands.

7.1 csquotes

The recommended way of inserting quotation marks in running Arabic text is to use `csquotes`. With the help of the `\DeclareQuoteStyle` command, one can define an Arabic style, like so:—

```
1 \usepackage{csquotes}
2 \DeclareQuoteStyle{arabic}
3 {\rmfamily\textquotedblright}{\rmfamily\textquotedblleft}
4 {\rmfamily\textquoteright}{\rmfamily\textquoteleft}
```

Then, use this newly defined style with `\setquotestyle`, like so:—

```
1 \setquotestyle{arabic}
2 \begin{arab}
3   fa-qAla la-hu ju.hA: \enquote{.garIbuN 'amru-ka yA .sadIqI
4     'a-tu.saddiqu 'l-.himAra wa-tuka_d_dibu-nI?}
5 \end{arab}
6 \setquotestyle{english}
```

فَقَالَ لَهُ جُحَا: "غَرِيبٌ أَمْرُكَ يَا صَدِيقِي أَتَصَدِّقُ الْجَمَّارَ وَتَكْذِبُنِي؟"

REM. Do not forget to set back the quoting style to its initial state once the Arabic environment is closed. See the last line in the code above.

7.2 reledmac

The two-arguments command `\edtext{<lemma>}{<commands>}` is supported inside `\begin{arab} ... \end{arab}`. As an example, one may get `arabluatex` and `reledmac` to work together like so:—

```
1 \beginnumbering
2 \pstart
3 \begin{arab}
4   wa-ya.sIru ta.hta 'l-jildi
5   \edtext{\arb{.sadIduN}}{\Afootnote{M: \arb{.sadIdaN} E1}}
6   \end{arab}
7 \pend
8 \endnumbering
```

8 Future work

A short, uncommented, list of what is planned in the versions of `arabluatex` to come follows:

- (a) Short-term:
 - i. Support for typesetting Arabic poetry.
 - ii. The *Qur'ān*: support for typesetting the *Qur'ān*.
 - iii. TEI xml support: arabluatex will interoperate with TEI xml through new global and local options that will output Arabic in a TEI xml compliant file in addition to the usual PDF output: see on page 3.
- (b) Medium-term:
 - i. More languages: the list of supported languages will eventually be the same as arabtex: see footnote 4 on page 4.
 - ii. Formulate propositions for extending the ArabTeX notation and the transliteration tables. Include them in arabluatex. See section 4.8 on page 21.

9 Implementation

The most important part of arabluatex relies on Lua functions and tables. Read the .lua files that accompany arabluatex for more information.

```

1 \NeedsTeXFormat{LaTeX2e}
2 \ProvidesPackage{arabluatex}%
3 [2016/05/09 v1.2 ArabTeX-like interface for LuaLaTeX]
4 \RequirePackage{ifluatex}

```

arabluatex requires Lua⁴TeX of course. Issue a warning if the document is processed with another engine.

```

5 \ifluatex\else
6 \PackageError{arabluatex}{lualatex needed}{%
7 Package `arabluatex' needs LuaTeX.\MessageBreak
8 So you should use `lualatex' to process your document.\MessageBreak
9 See documentation of `arabluatex' for further information.}%
10 \expandafter\expandafter\expandafter\csname endinput\endcsname
11 \fi

```

Declare the global options, and define them:

```

12 \DeclareOption{voc}{\def\al@mode{voc}}
13 \DeclareOption{fullvoc}{\def\al@mode{fullvoc}}
14 \DeclareOption{novoc}{\def\al@mode{novoc}}
15 \DeclareOption{trans}{\def\al@mode{trans}}
16 \ExecuteOptions{voc}
17 \ProcessOptions\relax
18 \def\al@mode@voc{voc}
19 \def\al@mode@fullvoc{fullvoc}
20 \def\al@mode@novoc{novoc}
21 \def\al@mode@trans{trans}

```

The following line will be used in the next release of arabluatex:

```

22 % \newif\ifal@mode@defined

```

Packages that are required by arabluatex:

```

23 \RequirePackage{fontspec}
24 \RequirePackage{amsmath}

```

```

25 \RequirePackage{etoolbox}
26 \RequirePackage{luacode}
27 \RequirePackage{xparse}
28 \RequirePackage{environ}

```

Here begins the real work: load `arabluatex.lua`:

```

29 \luadirect{do file(kpse.find_file("arabluatex.lua"))}

```

This is needed by the current versions of `polyglossia` and `luabidi`. `luabidi` provides a `\Footnote` command. Use it as well if it is loaded.

```

30 \luadirect{tex.enableprimitives("luatex",tex.extraprimitives("omega"))}

```

Font setup. If no Arabic font is selected, issue a warning message and attempt to load the Amiri font which is included in `TEXlive`:

```

31 \AtBeginDocument{\ifdefined\arabicfont\relax\else
32 \PackageWarning{arabluatex}{\string\arabicfont\ is not defined.^}
33 \message{will try to load Amiri}%
34 \newfontfamily\arabicfont[Script=Arabic]{Amiri}\fi}%

```

`\setRL` This neutralizes what is defined by the same command in `luabidi`:

```

35 \AtBeginDocument{\def\setRL{\pardir TRT\textdir TRT}}

```

`\setLR` The same applies to `\setLR`:

```

36 \AtBeginDocument{\def\setLR{\pardir TLT\textdir TLT}}

```

`\LR` This command typesets its argument from left to right. As `\LR` may be already defined, we need to redefine for it to suit our purpose:

```

37 \AtBeginDocument{\ifdef\LR}%
38 {\RenewDocumentCommand\LR}{m}{\bgroup\textdir TLT\rmfamily#1\egroup}}
39 {\NewDocumentCommand\LR}{m}{\bgroup\textdir TLT\rmfamily#1\egroup}}

```

`\RL` This one typesets its argument from right to left. Same remark as above regarding the need of redefinition.

```

40 \AtBeginDocument{\ifdef\RL}%
41 {\RenewDocumentCommand\RL}{m}{\bgroup\textdir TRT\rmfamily#1\egroup}}
42 {\NewDocumentCommand\RL}{m}{\bgroup\textdir TRT\rmfamily#1\egroup}}

```

`\aemph` Arabic emphasis. Needs to be redefined as well.

```

43 \AtBeginDocument{\ifdef\aemph}%
44 {\RenewDocumentCommand\aemph}{m}{\overline{\text{#1}}}}
45 {\NewDocumentCommand\aemph}{m}{\overline{\text{#1}}}}

```

`\SetArbEasy` By default, `arabluatex` applies complex rules to generate euphonic *tašdīd*, *ʿalif mamdūda* and *sukūn* depending on the modes which are selected, either `voc`, `fullvoc` or `trans`. Such refinements can be discarded with `\SetArbEasy`, either globally in the preamble or at any point of the document. Default complex rules can be set back at any point of the document with `\SetArbDflt`.

```

46 \def\al@arb@rules{dflt}
47 \NewDocumentCommand\SetArbEasy{}{\def\al@arb@rules{easy}}
48 \NewDocumentCommand\SetArbDflt{}{\def\al@arb@rules{dflt}}

```

`\SetTranslitStyle` By default any transliterated Arabic text is printed in italics. This can be changed globally in the preamble or at any point of the document:

```

49 \def\al@trans@style{\itshape}%
50 \NewDocumentCommand{\SetTranslitStyle}{m}{\def\al@trans@style{#1}}

```

`\SetTranslitConvention` `\SetTranslitConvention{<convention>}` may be used to change the transliteration convention, which is `dmg` by default:

```

51 \def\al@trans@convention{dmg}
52 \NewDocumentCommand{\SetTranslitConvention}{m}{\def\al@trans@convention{#1}}

```

`\cap` Proper Arabic names or book titles should be passed to the command `\cap` so that they have their first letters upcased. `\cap` is actually coded in Lua.

```

53 \DeclareDocumentCommand{\cap}{m}%
54   {\luairect{tex.sprint(cap(\luastringN{#1}))}}

```

`\txarb` `\txarb` sets the direction to right-to-left and selects the Arabic font. As it is supposed to be used internally by several Lua functions, this command is not documented, but available to the user should he wish to insert utf8 Arabic text in his document.

`\txtrans` `\txtrans` is used internally by several Lua functions to insert transliterated Arabic text.

```

55 \DeclareDocumentCommand{\txarb}{+m}{\bgroup\textdir
56   TRT\arabicfont#1\egroup}
57 \DeclareDocumentCommand{\txtrans}{+m}{\bgroup\textdir
58   TLT\rmfamily#1\egroup}

```

`\arb` The `\arb` command detects which Arabic mode is to be used, either globally if no option is set, or locally, then passes its argument to the appropriate Lua function.

```

59 \DeclareDocumentCommand{\arb}{O{\al@mode} +m}%
60 {\edef\@tempa{#1}%
61   \ifx\@tempa\al@mode@voc%
62     \bgroup\textdir TRT\arabicfont%
63     \luairect{tex.sprint(processvoc(\luastringN{#2}),
64       \luastringO{\al@arb@rules}))}\egroup%
65   \else%
66     \ifx\@tempa\al@mode@fullvoc%
67       \bgroup\textdir TRT\arabicfont%
68       \luairect{tex.sprint(processfullvoc(\luastringN{#2}),
69         \luastringO{\al@arb@rules}))}\egroup%
70     \else%
71       \ifx\@tempa\al@mode@novoc%
72         \bgroup\textdir TRT\arabicfont%
73         \luairect{tex.sprint(processnovoc(\luastringN{#2}))}\egroup%
74       \else%
75         \ifx\@tempa\al@mode@trans%
76           \bgroup\textdir TLT\al@trans@style%
77           \luairect{tex.sprint(processstrans(\luastringN{#2}),
78             \luastringO{\al@trans@convention}),
79             \luastringO{\al@arb@rules}))}\egroup%

```

```

80 \else%
81 \fi\fi\fi\fi}

```

arab The arab environment does for paragraphs the same as \arb does for short insertions of Arabic text.

```

82 \NewEnviron{arab}[1][\al@mode]%
83 {\par\edef\@tempa{#1}%
84 \ifx\@tempa\al@mode@voc%
85 \bgroup\pdir TRT\textdir TRT\arabicfont%
86 \luadirect{tex.sprint(processvoc(\luastringO{\BODY},
87 \luastringO{\al@arb@rules}))}\egroup%
88 \else%
89 \ifx\@tempa\al@mode@fullvoc%
90 \bgroup\pdir TRT\textdir TRT\arabicfont%
91 \luadirect{tex.sprint(processfullvoc(\luastringO{\BODY},
92 \luastringO{\al@arb@rules}))}\egroup%
93 \else%
94 \ifx\@tempa\al@mode@novoc%
95 \bgroup\pdir TRT\textdir TRT\arabicfont%
96 \luadirect{tex.sprint(processnovoc(\luastringO{\BODY}))}\egroup%
97 \else \ifx\@tempa\al@mode@trans%
98 \bgroup\pdir TLT\textdir TLT\al@trans@style%
99 \luadirect{tex.sprint(processtrans(\luastringO{\BODY},
100 \luastringO{\al@trans@convention},
101 \luastringO{\al@arb@rules}))}\egroup%
102 \else \fi\fi\fi\fi}\par}

```

\abjad \abjad{<number>} expresses its argument in Arabic letters in accordance with the 'abjad arrangement of the alphabet. <number> must be between 1 and 1999. It is now coded in Lua so that polyglossia is no longer needed. See arabluatex.lua for more information.

```

103 \AtBeginDocument{%
104 \ifdefined\abjad%
105 \RenewDocumentCommand{\abjad}{m}%
106 {\luadirect{tex.sprint(abjadify(#1))}}%
107 \else%
108 \NewDocumentCommand{\abjad}{m}%
109 {\luadirect{tex.sprint(abjadify(#1))}}
110 \fi}

```

\LRmarginpar \LRmarginpar is supposed to be inserted in an Arabic environment. It typesets his argument in a marginal note from left to right.

```

111 \DeclareDocumentCommand{\LRmarginpar}{m}{\marginpar{\textdir TLT #1}}

```

\LRfootnote \LRfootnote and \RLfootnote are supposed to be used in Arabic environments for insertions of non Arabic text. \LRfootnote typesets its argument left-to-right...

\RLfootnote while \RLfootnote typesets its argument left-to-right.

```

112 \DeclareDocumentCommand{\LRfootnote}{m}{\bgroup\pdir

```

```

113 TLT\LR{\footnote{#1}}\egroup}
114 \DeclareDocumentCommand{\RLfootnote}{m}{\bgroup\pdir
115 TRT\LR{\footnote{#1}}\egroup}

```

`\FixArbFtnmk` In the preamble, just below `\usepackage{arabluatex}`, `\FixArbFtnmk` may be of some help in case the footnote numbers at the bottom of the page are printed in the wrong direction. This quick fix uses and loads `scrextend` if it is not already loaded.

```

116 \NewDocumentCommand{\FixArbFtnmk}{}{%
117   \@ifpackageloaded{scrextend}%
118   {\AtBeginDocument{\deffootnote{2em}{1.6em}{\LR{\thefootnotemark}.\enskip}}}%
119   {\RequirePackage{scrextend}
120   \AtBeginDocument{\deffootnote{2em}{1.6em}{\LR{\thefootnotemark}.\enskip}}}%

```

That is it. Say goodbye before leaving.

```

121 \endinput

```

References

- Brockelmann, Carl et al. (1935). “Die Transliteration der arabischen Schrift in ihrer Anwendung auf die Hauptliteratursprachen der islamischen Welt”. In: *Denkschrift dem 19. internationalen Orientalistenkongreß in Rom*. In collab. with Ph. S. van Ronkel and Otto Spies. Deutschen Morgenländischen Gesellschaft. Leipzig: Deutschen Morgenländischen Gesellschaft, in Kommission bei F. A. Brockhaus. URL: http://www.naher-osten.uni-muenchen.de/studium_lehre/werkzeugkasten/dmgtransliteration.pdf.
- Hosny, Khaled (2015). *Amiri*. URL: <http://www.amirifont.org/>.
- Lagally, Klaus (2004). *ArabTeX. Typesetting Arabic and Hebrew*. User Manual Version 4.00. Version 4.00. URL: <http://mirrors.ctan.org/language/arabic/arabtex/doc/html/arabtex.htm>.
- Lane, Edward William (1863–1893). *An Arabic-English lexicon*. 8 vols. London – Edinburgh: Williams and Norgate.
- Wright, W. LL.D (1896). *A Grammar of the Arabic Language*. Rev. by W. Robertson Smith and M. J. de Goeje. With a forew. by Pierre Cachia. 3rd ed. 2 vols. Beirut: Librairie du Liban.

Change History

v1.0		v1.2
General: Initial release	1	\SetArbEasy: New
v1.0.1		\SetArbEasy/\SetArbDflt
General: Minor update of the		for ‘modern’ or ‘classic’ Arabic
documentation	1	styles. 28
v1.1		
\abjad: New and more flexible		
\abjad command.	30	

Index

Numbers written in *italic* refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in *roman* refer to the code lines where the entry is used.

Symbols		
\@ifpackageloaded	117	
\@tempa	60, 61, 66, 71, 75, 83, 84, 89, 94, 97	
_	32	
A		
\abjad	20, 103	
\aemph	21, 43	
\al@arb@rules	46, 47, 48, 64, 69, 79, 87, 92, 101	
\al@mode	12, 13, 14, 15, 59, 82	
\al@mode@fullvoc	19, 66, 89	
\al@mode@novoc	20, 71, 94	
\al@mode@trans	21, 75, 97	
\al@mode@voc	18, 61, 84	
\al@trans@convention	51, 52, 78, 100	
\al@trans@style	49, 50, 76, 98	
<i>Amiri</i>	4	
amiri (package)	4	
arab (environment)	6, 82	
<i>ArabTeX</i>	3, 17	
<i>Arabic-English Lexicon</i> , <i>An</i>	16	
\arabicfont	31, 32, 34, 56, 62, 67, 72, 85, 90, 95	
arabluatex (package)	2–7, 9–11, 13–17, 20–28	
arabtex (package)	3, 4, 13, 17, 27	
arabulatex (package)	14	
arabxetex (package)	3, 4, 13	
\arb	6, 59	
B		
\BODY	86, 91, 96, 99	
Brockelmann, Carl	22	
C		
\cap	22, 53	
csquotes (package)	26	
D		
\DeclareOption	12, 13, 14, 15	
\deffootnote	118, 120	
E		
\enskip	118, 120	
environments:		
arab	6, 82	
\ExecuteOptions	16	
F		
\FixArbFtnmk	25, 116	
fontspec (package)	4, 22	
\footnote	113, 115	
fullvoc (option)	5, 7	
G		
<i>Grammar of the Arabic</i> <i>Language, A</i>	4–7, 11–15, 20, 21	
H		
Hosny, Khaled	4	
I		
\ifal@mode@defined	22	
\ifluatex	5	
\itshape	49	
K		
KOMA-script (package)	25	
L		
Lagally, Klaus	3, 17	
Lane, Edward William	16	
\LR	24, 37, 113, 115, 118, 120	
\LRfootnote	24, 112	
\LRmarginpar	25, 111	
luabidi (package)	4, 28	
\luadirect	29, 30, 54, 63, 68, 73, 77, 86, 91, 96, 99, 106, 109	
\luastringN	54, 63, 68, 73, 77	
\luastringO	64, 69, 78, 79, 86, 87, 91, 92, 96, 99, 100, 101	
M		
\marginpar	111	
N		
\NewEnviron	82	
\newfontfamily	34	
novoc (option)	5, 7	
O		
options:		
fullvoc	5, 7	
novoc	5, 7	
trans	5, 7	
voc	5, 7	
\overline	44, 45	
P		
\par	83, 102	
\pardir	35, 36, 85, 90, 95, 98, 112, 114	
polyglossia (package)	4, 28, 30	
\ProcessOptions	17	
\ProvidesPackage	2	
R		
reledmac (package)	26	
\RequirePackage	4, 23, 24, 25, 26, 27, 28, 119	
\RL	24, 40	

<code>\RLfootnote</code>	24, 112	<code>\SetTranslitStyle</code>	<i>Transliteration der arabis-</i>
<code>\rmfamily</code> 38 , 39 , 41 , 42 , 58		<i>chen Schrift, Die</i> . . . 22
S		T	
<code>scrextend</code> (package)	25, 31	<code>\text</code>	44 , 45
<code>\SetArbDflt</code>	5, 46	<code>\textdir</code>	35 ,
<code>\SetArbEasy</code>	5, 46		36 , 38 , 39 , 41 , 42 ,
<code>\setLR</code>	25, 36		55 , 57 , 62 , 67 , 72 ,
<code>\setRL</code>	25, 35		76 , 85 , 90 , 95 , 98 , 111
<code>\SetTranslitConvention</code>		<code>\thefootnotemark</code> .	
.	22, 51	118 , 120
		<code>trans</code> (option)	5, 7
		V	
		<code>\txarb</code>	55
		<code>\txtrans</code>	55
		W	
		<code>voc</code> (option)	5, 7
		W	
		Wright, W. LL.D	
		4–7, 11–15, 20, 21