

The arabluatex package

v1.12 – 2018/06/24

Robert Alessi

alessi@robertalessi.net

Contents

| | | | |
|--|-----------|--|-----------|
| License and disclaimer | 2 | 4.3 Special orthographies | 20 |
| 1 Introduction | 3 | The name of God 20 The | |
| 1.1 arabluatex is for Lua ^A T _E X . . | 4 | conjunctive ^{اَللّٰهِي} 21 | |
| 2 The basics of arabluatex | 5 | 4.4 Quoting | 22 |
| 2.1 Activating arabluatex | 5 | novoc 22 voc 22 fullvoc | |
| Font setup 5 | | 23 | |
| 2.2 Options | 5 | 4.4.1 Quoting the <i>hamzah</i> | 23 |
| 2.2.1 Classic contrasted | | 4.5 The “pipe” character () . . . | 24 |
| with modern typesetting of | | 4.6 Putting back on broken con- | |
| Arabic | 6 | textual analysis rules | 24 |
| 2.3 Typing Arabic | 7 | 4.7 Stretching characters: the | |
| Caveat 8 | | <i>taḥwīl</i> | 26 |
| 2.3.1 Local options | 9 | 4.8 Digits | 26 |
| 3 Standard Arab^TE_X input | 9 | 4.8.1 Numerical figures . . | 26 |
| 3.1 Consonants | 9 | 4.8.2 The <i>abjad</i> | 26 |
| 3.2 Additional characters | 11 | 4.9 Additional characters | 27 |
| 3.3 Vowels | 11 | 4.10 Arabic emphasis | 28 |
| 3.3.1 Long vowels | 11 | 5 Arabic poetry | 28 |
| 3.3.2 Short vowels | 12 | Scaling and distortion of | |
| 4 arabluatex in action | 13 | characters 31 Footnotes 31 | |
| 4.1 The vowels and diphthongs . . | 13 | Line numbering 31 | |
| Short vowels 13 Long vow- | | 5.1 Example | 31 |
| els 13 <i>‘alif maqṣūrah</i> 13 | | 6 Special applications | 33 |
| <i>‘alif otiosum</i> 13 <i>‘alif maḥdū-</i> | | Linguistics 33 Brackets 33 | |
| <i>fah</i> and defective <i>ū</i> , <i>ī</i> 14 | | Additional Arabic marks 33 | |
| Silent ^و / _ي 14 <i>‘Amr^{un}</i> , and | | 7 Color | 35 |
| the silent ^و / _ي 14 <i>tanwīn</i> 14 | | 7.1 Tricks of the trade | 36 |
| 4.2 Other orthographic signs . . . | 15 | Diacritics 36 <i>tanwīn</i> 37 | |
| <i>tā’ marbūṭah</i> 15 <i>hamzah</i> 15 | | <i>waṣlah</i> and <i>maddah</i> 38 <i>ṣad-</i> | |
| <i>maddah</i> 17 <i>ṣaddah</i> 17 The | | <i>dah</i> 39 The definite article | |
| definite article and the <i>‘alif^u</i> | | and the euphonic <i>taṣdīd</i> 39 | |
| <i>‘l-waṣlⁱ</i> 18 | | <i>hamzah</i> 39 | |

| | | | |
|--|-----------|---|-----------|
| 8 Transliteration | 39 | 11 L^AT_EX Commands in Arabic environments | 48 |
| Convention | 40 | General principle | 48 |
| Font | 40 | 11.1 New commands | 50 |
| Proper names | 40 | 11.2 Environments | 51 |
| Hyphenation | 41 | 11.2.1 Lists | 52 |
| ‘Long’ proper names | 41 | رموز الكتاب 52 Caveat | 54 |
| Proper names outside Arabic environments | 41 | 11.3 csquotes | 55 |
| 8.1 Additional note on <code>dmg</code> convention | 42 | 11.4 Two-argument special commands | 55 |
| <i>ʾirāb</i> boundaries | 43 | textcolor | 55 |
| Discarding the <i>ʾirāb</i> | 43 | 11.5 quran | 56 |
| Uncertain short vowels | 44 | 12 Future work | 57 |
| 8.2 Examples | 44 | 13 Implementation | 57 |
| 9 Buckwalter input scheme | 44 | References | 68 |
| ‘base’, ‘xml’ and ‘safe’ schemes | 45 | Change History | 68 |
| Transliteration | 47 | Index | 69 |
| 10 Unicode Arabic input | 47 | | |

List of Tables

| | | | |
|--|----|--|----|
| 1 ArabT _E X consonants | 10 | 6 Additional Arabic codings | 27 |
| 2 ArabT _E X additional characters | 11 | 7 Additional Arabic marks | 33 |
| 3 ArabT _E X long vowels | 12 | 8 ArabT _E X diacritics for <code>\arbcolor</code> | 36 |
| 4 ArabT _E X short vowels | 12 | 9 Buckwalter scheme | 45 |
| 5 “Quoted” <i>hamzah</i> | 23 | | |

Abstract

This package provides for LuaL^AT_EX an ArabT_EX-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 ArabT_EX input notation. Its output can be set in the same modes of vocalization as ArabT_EX, 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

License applicable to the software

arabluatex — Processing ArabT_EX notation under LuaL^AT_EX.
Copyright © 2016–2018 Robert Alessi

Please send error reports and suggestions for improvements to Robert Alessi:
– email: <alessi@robertalessi.net>

gpl3+

- website: <http://www.robtalessi.net/arabluatex>
- development: <http://git.robtalessi.net/arabluatex>
- comments, feature requests, bug reports: <https://notabug.org/ralessi/arabluatex/issues>

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/>.


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`
- `arabluatex.el`

License applicable to this document

Copyright © 2016–2018 Robert Alessi

CC BY-SA 4.0

 This document is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

1 Introduction

In comparison to Prof. Lagally’s outstanding `ArabTeX`,¹ `ArabLuaTeX` 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 `LuaLATeX` users.

`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, `ArabTeX` ASCII input notation still surpasses

¹See <http://ctan.org/pkg/arabtex>

²Lagally (2004, p. 2).

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 L^AT_EX 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 ArabT_EX 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. ArabT_EX notation is a way of encoding Arabic language, just as T_EX “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, ArabT_EX 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, TEI xml compliant critical editions and/or lexicons that can be searched, analyzed and correlated in various ways.

1.1 arabluatex is for LuaL^AT_EX

It goes without saying that `arabluatex` requires LuaL^AT_EX. T_EX and L^AT_EX have `arabtex`, and X_LL^AT_EX has `arabxetex`. Both of them are much more advanced than `arabluatex`, 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`.

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

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

In comparison to `arabxetex`, `arabluatex` works in a very different way. The former relies on the `TEckit` engine which converts ArabTeX input on the fly into Unicode Arabic script, whereas the latter passes ArabTeX 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 ArabTeX input yet to be processed—for further processing.

2 The basics of arabluatex

2.1 Activating arabluatex

`arabluatex` is loaded the usual way:

```
\usepackage{arabluatex}
```

The only requirement of `arabluatex` is LuaL^AT_EX; it will complain if the document is compiled with another engine. That aside, `arabluatex` does not load packages such as `polyglossia`. Although it can work with `polyglossia`, it does not require it.

Font setup Any Arabic font can be defined to be used with `arabluatex`. For example, assuming that `fontspec` is loaded, this line may be inserted in the preamble, just above the line that loads `arabluatex`:

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

where *<fontname>* is the standard name of the Arabic font to be used.

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{Amiri}[Script=Arabic]
```

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

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

Other Arabic fonts may behave differently.

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 9):

default

voc

⁵Hosny (2017).

⁶See the documentation of `amiri`, Hosny (2017, p. 6).

In this mode, which is the one selected by default, every short vowel written generates its corresponding diacritical mark: *dammah* (◌ُ), *fathah* (◌َ) and *kasrah* (◌ِ). If a vowel is followed by N, viz. $\langle uN, aN, iN \rangle$, then the corresponding *tanwīn* (◌ِ, ◌ُ, ◌ِ, ◌ِ or ◌ِ) 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 *waṣlah* above the ‘*alif*”; instead, the accompanying vowel may be expressed at the beginning of a sentence (أَ).

fullvoc

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

novoc

None of the diacritics is showed in **novoc** mode, unless otherwise specified (see “quoting” technique below section 4.4 on page 22).

trans

This mode transliterates the ArabTeX input into one of the accepted standards. At present, three standards are supported (see below section 8 on page 39 for more details):

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

loc *Library of Congress*;

arabica *Arabica*.

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’ *maddah*: when ‘*alif* and *hamzah* accompanied by a simple vowel or *tanwīn* is preceded by an ‘*alif* of prolongation (◌ِ), then a mere *hamzah* is written on the line, and a *maddah* 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 16 for further details).

- The euphonic *tašdīd* is generated (see on page 17).
- In **fullvoc** mode, the *sukūn* is expressed.
- In such words as ظَمًا, شَيْئًا and the like, the *hamzah* alone is not written over the letter *yā’* with no diacritical points below as in ظَمًا, شَيْئًا, but over a horizontal stroke placed in the continuation of the preceding letter.

Please note that only few Arabic fonts provide such contrivances. In case this feature is not supported by some Arabic font, it is advisable to use `\SetArbEasy`.

`\SetArbEasy`

Such refinements as ‘classic’ *maddah* may be discarded by the `\SetArbEasy`

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

New feature
v.1.2

New feature
v1.4.4

command, either globally in the preamble or locally at any point of the document.

`\SetArbEasy*` The difference between `\SetArbEasy` and its ‘starred’ version `\SetArbEasy*` is that the former keeps the *sukūn* that is generated by the `fullvoc` mode, while the latter further takes it away. Default ‘classic’ rules may be set back at any point of the document with the `\SetArbDflt` command. Assimilation rules laid on item (b) on page 17 may also be applied by the ‘starred’ version of this command `\SetArbDflt*` either in the preamble or at any point of the document.⁸ Examples follow:—

`\SetArbDflt*`

(a) `\SetArbDflt`:

- i. voc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- ii. fullvoc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- iii. trans *wa-māta 'stisqā^{an} qabla 'an yutimma kitāba-hu fī nuġūmⁱ 's-samāⁱ*

(b) `\SetArbDflt*`:

- i. voc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- ii. fullvoc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- iii. trans *wa-māta 'stisqā^{an} qabla 'ay yutimma kitāba-hu fī nuġūmⁱ 's-samāⁱ*

(c) `\SetArbEasy`:

- i. voc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- ii. fullvoc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- iii. trans *wa-māta 'stisqā^{an} qabla 'an yutimma kitāba-hu fī nuġūmⁱ 's-samāⁱ*

(d) `\SetArbEasy*`:

- i. voc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- ii. fullvoc وَمَاتَ اسْتِسْقَاءٌ قَبْلَ أَنْ يُتِمَّ كِتَابَهُ فِي نُجُومِ السَّمَاءِ
- iii. trans *wa-māta 'stisqā^{an} qabla 'an yutimma kitāba-hu fī nuġūmⁱ 's-samāⁱ*

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

2.3 Typing Arabic

`\arb` Once `arabluatex` is loaded, a `\arb{⟨Arabic text⟩}` command 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{.hurUf-u 'l-hijA'-i}`), `\arb{.hurUf-u`

⁸For an example, see section 5.1 on page 31.

```

4 'l-tahajjI}, \arb{al-.hurUf-u 'l-hijA'iyyaT-u}, or
5 \arb{.hurUf-u 'l-mu`jam-i}) are twenty-eight in number and
6 are all consonants, though three of them are also used as
7 vowels (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 (حُرُوفُ الهجاء, حُرُوفُ الحروف الهجائية, or حُرُوفُ المعجم) are twenty-eight in number and are all consonants, though three of them are also used as vowels (see § 3).

Caveat For some reason, left-to-right paragraphs that start with Arabic words lose their indentation. For the time being, this can be circumvented by appending the `\indent` command at the commencement of such paragraphs.

The same remark applies to left-to-right list environments: when items start with Arabic words, the `\arb` command must be prefixed with `\indent`. The following example comes from Wright (1896, i. 213 C):—

```

1 \begin{enumerate}[label=\Roman*., start=16]
2   \item \indent\arb{fawA`ilu}*.
3     \begin{enumerate}[label=\arabic*.]
4       \item \indent\arb{fA`aluN}; as \arb{_hAtamuN} \emph{a
5         signet-ring}, ...
6     \end{enumerate}
7 \end{enumerate}

```

XVI. فَوَاعِلُ*.

1. فَاعِلٌ; as خَاتَمٌ *a signet-ring*, ...

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 fa-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 ra.gbata-ka fa-'l-.himAr-u laysa hunA
8 'l-yawm-a.} wa-qabla 'an yutimma ju.hA kalAma-hu bada'a
9 'l-.himAr-u yanhaqu fI 'i.s.tabli-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-.himAr-a
13 wa-tuka_d_diba-nI?}
14 \end{arab}

```

أَتَى صَدِيقٌ إِلَى جَحَّا يَطْلُبُ مِنْهُ حِمَارَهُ لِيَرْكَبَهُ فِي سَفَرَةٍ قَصِيرَةٍ فَقَالَ لَهُ: "سَوْفَ أُعِيدُهُ إِلَيْكَ فِي الْمَسَاءِ
وَأَدْفَعُ لَكَ أُجْرَةً." فَقَالَ جَحَّا: "أَنَا آسَفٌ جِدًّا أَنِّي لَا أَسْتَطِيعُ أَنْ أُحَقِّقَ لَكَ رَغْبَتَكَ فَالْحِمَارُ لَيْسَ هُنَا
الْيَوْمَ." وَقَبْلَ أَنْ يَتِمَّ جَحَّا كَلَامَهُ بَدَأَ الْحِمَارُ يَنْهَقُ فِي إِصْطَبْلِهِ. فَقَالَ لَهُ صَدِيقُهُ: "إِنِّي أَسْمَعُ حِمَارَكَ يَا جَحَّا
يَنْهَقُ." فَقَالَ لَهُ جَحَّا: "غَرِيبُ أَمْرِكَ يَا صَدِيقِي أَتَصَدِّقُ الْحِمَارَ وَتَكْذِبُنِي؟"

2.3.1 Local options

As seen above in section 2.2 on page 5, arabuatex 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 argument, like so:—

| | |
|----------------------|--|
| <code>voc</code> | — <code>\arb[voc]{\langle Arabic text \rangle};</code> |
| <code>fullvoc</code> | — <code>\arb[fullvoc]{\langle Arabic text \rangle};</code> |
| <code>novoc</code> | — <code>\arb[novoc]{\langle Arabic text \rangle};</code> |
| <code>trans</code> | — <code>\arb[trans]{\langle Arabic text \rangle};</code> |

The same optional arguments may be passed to the environment `arab`: one may have `\begin{arab}[\langle mode \rangle] ... \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 ⁹ | | | ArabT _E X notation |
|-----------------|------------------------------|------------|------------|-------------------------------|
| | dmg | loc | arabica | |
| أ ¹⁰ | 'u 'a 'i | 'u, 'a, 'i | 'u, 'a, 'i | 'u or 'a or 'i |
| ب | <i>b</i> | <i>b</i> | <i>b</i> | b |
| ت | <i>t</i> | <i>t</i> | <i>t</i> | t |
| ث | <i>ṭ</i> | <i>th</i> | <i>ṭ</i> | _t |
| ج | <i>ǧ</i> | <i>j</i> | <i>ǧ</i> | ^g or j |
| ح | <i>ḥ</i> | <i>ḥ</i> | <i>ḥ</i> | .h |
| خ | <i>ḫ</i> | <i>kh</i> | <i>ḫ</i> | _h or x |
| د | <i>d</i> | <i>d</i> | <i>d</i> | d |
| ذ | <i>ḍ</i> | <i>dh</i> | <i>ḍ</i> | _d |
| ر | <i>r</i> | <i>r</i> | <i>r</i> | r |
| ز | <i>z</i> | <i>z</i> | <i>z</i> | z |
| س | <i>s</i> | <i>s</i> | <i>s</i> | s |
| ش | <i>š</i> | <i>sh</i> | <i>š</i> | ^s |
| ص | <i>ṣ</i> | <i>ṣ</i> | <i>ṣ</i> | .s |
| ض | <i>ḍ</i> | <i>ḍ</i> | <i>ḍ</i> | .d |
| ط | <i>ṭ</i> | <i>ṭ</i> | <i>ṭ</i> | .t |
| ظ | <i>ẓ</i> | <i>ẓ</i> | <i>ẓ</i> | .z |
| ع | <i>ʿ</i> | <i>ʿ</i> | <i>ʿ</i> | ` |
| غ | <i>ǧ</i> | <i>gh</i> | <i>ǧ</i> | .g |
| ف | <i>f</i> | <i>f</i> | <i>f</i> | f |
| ق | <i>q</i> | <i>q</i> | <i>q</i> | q |
| ك | <i>k</i> | <i>k</i> | <i>k</i> | k |
| ل | <i>l</i> | <i>l</i> | <i>l</i> | l |
| م | <i>m</i> | <i>m</i> | <i>m</i> | m |
| ن | <i>n</i> | <i>n</i> | <i>n</i> | n |
| ه | <i>h</i> | <i>h</i> | <i>h</i> | h |
| و | <i>w</i> | <i>w</i> | <i>w</i> | w |
| ي | <i>y</i> | <i>y</i> | <i>y</i> | y ¹¹ |
| ة | <i>ah</i> | <i>ah</i> | <i>a</i> | T |

Table 1: Standard ArabT_EX (consonants)

REM. *a*. Please note that in all cases of elision, the *ʿalifu ʿl-waṣli* is expressed only by the vowel that accompanies the omitted *hamzah*: ⟨*u, a, i*⟩ as in **wa-inhazama** وانْهَزَمَ *wa-ʿnhazama*. For more details on the definite article and the *ʿalifu ʿl-waṣli* see section 4.2 on page 18.

⁹See below section 8 on page 39.

¹⁰See below, Rem. *a*. For *ʿalif* as a consonant, see Wright (1896, i. 16 D). The *hamzah* itself is encoded <'> and may be followed by either ⟨*u, a*⟩ or ⟨*i*⟩. See below section 4.2 on page 15.

¹¹For the letter ع with no diacritical points below, see Rem. *b*. below.

That said, ʾ as a consonant is actually the *spiritus lenis* of the Greeks and is distinguished by the *hamzah* (ء) as it is shown in the above table. However, the bare *ʾalif* may also be encoded as .A whether it be followed by a vowel or not, like so: $\text{wa-}.\text{An}$ وَآن $wa-.n$ (where the dot symbolizes the absence of vowel), $\text{wa-}.\text{Aan}$ وَآن $wa-an$, $\text{wa-}.\text{Ain}$ وَآين $wa-in$.

REM. *b*. The letter ي with two points below, $\text{أَلِیَاءُ الْمُنَاةُ مِنْ تَحْتِیَا}$, may also be written without diacritical points as ی . When it is used as a consonant, it is encoded aY , where a recalls the *fathah* placed above the preceding letter in vocalized Arabic, like so: qaY'uN قَیْ qay^{un} , ^saY'uN سَیْ say^{un} , ^saY'aN سَیْ say^{an} .

The same result may be achieved by encoding this letter as .y , like so: qa.y'uN قَیْ qay^{un} , ^sa.y'uN سَیْ say^{un} , ^sa.y'aN سَیْ say^{an} .

3.2 Additional characters

Table 2 gives the ArabTeX equivalents for some additional Persian characters.

| Letter | Transliteration ¹² | | | ArabTeX notation |
|-----------------|-------------------------------|-----------|-----------------------|------------------|
| | dmg | loc | arabica ¹³ | |
| پ | <i>p</i> | <i>p</i> | <i>p</i> | p |
| چ | <i>č</i> | <i>ch</i> | <i>č</i> | ^c |
| ژ | <i>ž</i> | <i>zh</i> | <i>ž</i> | ^z |
| ف ¹⁴ | <i>v</i> | <i>v</i> | <i>v</i> | v |
| گ | <i>g</i> | <i>g</i> | <i>g</i> | g |
| ن ¹⁵ | <i>ñ</i> | <i>ñ</i> | <i>ñ</i> | ^n |

Table 2: Standard ArabTeX (additional characters)

REM. The alveolar consonants چ and ژ are processed as solar letters by *arabluatex*.

3.3 Vowels

3.3.1 Long vowels

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

¹²See below section 8 on page 39.

¹³The characters that are listed in this table are not included in this standard. However, as *arabica* is based on *dmg*, the *dmg* equivalents have been used here.

¹⁴This character is not found in Brockelmann et al. (1935, p. 2). It is taken from the DIN 31 635 (2011) standard.

¹⁵See footnote 14.

| Letter | Transliteration ¹⁶ | | | ArabTeX notation |
|-----------------|-------------------------------|-------------|-----------|------------------|
| | dmg | loc | arabica | |
| ا | \bar{a} | \bar{a} | \bar{a} | A |
| و | \bar{u} | \bar{u} | \bar{u} | U |
| ي | \bar{i} | \bar{i} | \bar{i} | I ¹⁷ |
| ى ¹⁸ | \bar{a} | \acute{a} | \bar{a} | _A or Y |
| أ | \bar{a} | \bar{a} | \bar{a} | _a |
| و | \bar{u} | \bar{u} | \bar{u} | _u |
| ي | \bar{i} | \bar{i} | \bar{i} | _i |

Table 3: Standard ArabTeX (long vowels)

REM. a. The long vowels \bar{a} , \bar{u} , \bar{i} , otherwise called *hurūf^u '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*, قَالَ يَقُولُ *yaqūlu*.

REM. b. Defective writings, such as ا, *al-'alif^u 'l-mahdūfat^u*, or defective writings of \bar{u} and \bar{i} are encoded _a_u and _i respectively, e.g. _d_alika ذَلِكَ, al-mal_a'ikaT-u 'l-ra.hm_an-u الْمَلِكُ الرَّحْمَنُ, hu_dayfaT-u bn-u 'l-yamAn_i حَذِيفَةُ بْنُ الْيَمَانِ for *Hudayfat^u bn^u 'l-Yamānī*, etc.

REM. c. The letter ي with two points below, أَلْيَاءُ الْمُتَأَنَّنِ مِنْ تَحْتِهَا, may also be written without diacritical points as ى. When it is used as a long vowel, it is encoded iY, where i recalls the *kasrah* placed below the preceding letter in vocalized Arabic, like so: liY لِي \bar{i} , yam[~]siY يَمَشِي *yamšī*.

3.3.2 Short vowels

Table 4 gives the ArabTeX equivalents for the Arabic short vowels.

| Letter | Transliteration ¹⁹ | | | ArabTeX notation |
|--------|-------------------------------|-----------|-----------|------------------|
| | dmg | loc | arabica | |
| ا | <i>a</i> | <i>a</i> | <i>a</i> | a |
| و | <i>u</i> | <i>u</i> | <i>u</i> | u |
| ي | <i>i</i> | <i>i</i> | <i>i</i> | i |
| اَ | <i>an</i> | <i>an</i> | <i>an</i> | aN |
| وَ | <i>un</i> | <i>un</i> | <i>un</i> | uN |
| يَ | <i>in</i> | <i>in</i> | <i>in</i> | iN |

Table 4: Standard ArabTeX (short vowels)

¹⁶See below section 8 on page 39.

¹⁷For the letter ي with no diacritical points, see Rem. c. below.

¹⁸= *al-'alif^u 'l-maḡṣūrat^u*.

¹⁹See below section 8 on page 39.

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.²⁰

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.

That said, there is no need to write the short vowels *fathah*, *dammah* or *kasrah* 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 22);
- when `arabluatex` needs to perform a contextual analysis to determine the carrier of the *hamzah*;
- in the various transliteration modes, as vowels are always expressed in romanized Arabic.

4 arabluatex in action

4.1 The vowels and diphthongs

Short vowels As said above, they are written $\langle a, u, i \rangle$:

_halaqa (or xalaqa) خَلَقَ *halaqa*, ^samsuN شَمْسٌ *šams^u*, karImuN كَرِيمٌ *Karīm^u*.

bi-hi بِهِ *bi-hi*, 'aqi.tuN أَقِطُ *'aqiṭ^u*.

la-hu لَهُ *la-hu*, .hujjaTuN هُجَّةٌ *huğğat^u*.

Long vowels They are written $\langle U, A, I \rangle$:

qAla قَالَ *qāla*, bI`a بَيْعٌ *bī`a*, .tUruN طُورٌ *ṭūr^u*, .tInuN طِينٌ *ṭīn^u*,
murU'aTuN مُرُوءَةٌ *murū`at^u*.

ʿalif maqṣūrah It is written $\langle _A \rangle$ or $\langle Y \rangle$:

al-fat_A الْفَتَى *al-fatā*, al-maqh_A الْمَقْهَى *al-maqhā*, 'il_A إِلَى *ʾilā*.

ʿalif otiosum Said *ʿalif^u ʿl-wiqāyatⁱ*, “the guarding *ʿalif*”, after و at the end of a word, both when preceded by *dammah* and by *fathah* is written $\langle UA \rangle$ or $\langle aW, aWA \rangle$:

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

²⁰See below section 4.4 on page 22.

‘alif maḥḍūfah and defective ū, ī They are written ⟨_a, _i _u⟩:

al-l_ah-u اللهُ *al-lāh^u*, 'il_ahNإله *ilāh^{un}*.
 al-ra.hm_an-uالرَّحْمَنُ *ar-rahmān^u*, l_akinلَكِنْ *lākin*, h_ahunAهَهْنَأُ *hāhunā*,
 .hunayn-u bn-u 'is.h_aq-aحُنَيْنُ بْنُ إِسْحَقَ *Hunayn^u bn^u 'Ishāq^a*, rabb_i
 رَبِّ *rabbī*, al-`A.s_iالْعَاصِ *al-Āṣī*.

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

.hay_aUTuNحَيَوَة *hayāt^{un}*, .sal_aUTuNصَلَوَة *ṣalāt^{un}*, mi^sk_aUTuNمَشْكُوَة *miš-kāt^{un}*, tawr_aITuNتَوْرِيَة *tawrāt^{un}*.
 And so also: al-rib_aIT-uالرِّيَبَة *ar-ribāt^u*.

‘Amr^{un}, and the silent و To that name a silent و is added to distinguish it from ‘Umar^u: 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:

`amruNUعَمْرُو *amr^{un}*, `amraNUعَمْرُوا *amr^{an}*, `amriNUعَمْرُو *amrⁱⁿ*.
 When the *tanwīn* falls away (Wright 1896, i. 249 B): `amr-uU bn-u mu.ammadiNعَمْرُو بْنُ مُحَمَّدٍ *Amr^u bn^u Muḥammadⁱⁿ*, mu.ammad-u bn-u `amr-iU bn-i_hAlidiNعَمْرُو بْنُ خَالِدٍ *Muḥammad^u bn^u Amrⁱ bnⁱ Ḥālidⁱⁿ*.
 And so also: al-rib_aUAالرِّيَوَا *ar-ribā*, ribaNUرِيَوَا *rib^{an}*.

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

mAluNمَالٌ *māl^{un}*, bAbaNبَابًا *bāb^{an}*, madInaTaNمَدِينَةٌ *madīnat^{an}*, bintiNبِنْتٌ *bintⁱⁿ* maqhaN_Aمَقْهَى *maqhaⁿ*, fataNYفَتًى *fataⁿ*.
 arablutex is aware of special orthographies: ^say'uNشَيْءٌ *šay^{un}*,
 ^say'aNشَيْئًا *šay^{an}*, ^say'iNشَيْءٍ *šayⁱⁿ*.

In some cases, it may be useful to mark the root form of defective words so as to produce a more accurate transliteration of ending *tanwīn*. As seen above, *tanwīn* preceding ى is written ⟨aN_A⟩ or ⟨aNY⟩. Such forms as قَاضٍ may likewise be written ⟨iNI⟩:—

al-qA.dI الْقَاضِي *al-qāḍī*, qA.diyaN قَاضِيًا *qāḍiy^{an}*, qA.diNI قَاضٍ *qāḍiⁿ*.

4.2 Other orthographic signs

tā' marbūṭah It is written ⟨T⟩:

madInaTuN مَدِينَةٌ *madīnat^{un}*, madInaTaN مَدِينَةً *madīnat^{an}*, madInaTiN مَدِينَةٍ *madīnatⁱⁿ*.

hamzah 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 22.

Initial hamzah: 'asaduN أَسَدٌ *'asad^{un}*, 'u_htuN أُخْتُ *'uht^{un}*, 'iqlIduN إِقْلِيدٌ *'iqlīd^{un}*, 'anna أَنْ *'anna*, 'inna إِنْ *'inna*.

hamzah followed by the long vowel و is encoded '_U: '_U1_A أُولَى *'ulā*, '_U1U أُولُو *'ulū*, '_U1A'ika أُولَئِكَ *'ulā'ika*.

hamzah followed by the long vowel ي is encoded '_I: '_ImAnuN إِيْمَانٌ *'īmā-n^{un}*²¹.

Middle hamzah: xA.ti'-Ina خَاطِنَ *hāṭi'ina*, ru'UsuN رُوُوسٌ *ru'ūs^{un}*, xa.tI'aTuN خَاطِنَةٌ *hāṭi'at^{un}*, su'ila سُئِلَ *su'ila*, 'as'ilaTuN أَسْئِلَةُ *'as'ilat^{un}*, mas'alaTuN مَسْأَلَةٌ *mas'alat^{un}*, 'as'alu أَسْأَلَ *'as'alu*, yataSA'alUna يَتَسَاءَلُونَ *yatasā'alūna*, murU'aTuN مَرْوَةٌ *taw'am^{un}*, murū'at^{un}, ta'xIruN تَأْخِرُ *ta'hīr^{un}*, ta'axxara تَأْخَرُ *ta'ahhara*, ji'tu-ka جِئْتُكَ *ǧi'tu-ka*, qA'iluN قَائِلٌ *qā'il^{un}*, .hIna'i_diN هِنَا *'idⁱⁿ* حِينَدَ, hay'aTuN هَيْةٌ *hay-'at^{un}*, hay'AtuN هَيَاتٌ *hay'āt^{un}*.

From Wright (1896, i. 14 B):— All consonants, whatsoever, not even *'alif hēmzatum* excepted, admit of being doubled and take *tašdīd*. Hence we speak and write ra'AsuN رَأْسٌ *ra'ās^{un}*, sa'AluN سَأَلَ *sa'al^{un}*, na'AjuN نَاجٌ *na'āǧ^{un}*.

²¹For another way of encoding the initial *hamzah* followed by a long vowel, see the *tahfīf^u* 'l-*hamzat*ⁱ on the next page.

Final hamzah: xa.ta'uN خَطَأُ *ḥaṭaʿun*, xa.ta'aN خَطَأٌ *ḥaṭaʿan*, xa.ta'iN خَطِئَ *ḥaṭaʿin*, 'aqra'u أَقْرَأُ *'aqra'u*, taqra'Inا تَقْرَأُ *taqra'ina*, taqra'Una تَقْرَأُونَ *taqra'una*, yaqra'na يَقْرَأُ *yaqra'na*, yaxba'Anي يَخْبَأُ *yahbaʿani*, xaba'A خَبَأَ *ḥabaʿa*, xubi'a خُبِيَ *ḥubiʿa*, xubi'UA خُبِئُوا *ḥubiʿu*, jA'a جَاءَ *ǧāʿa*, ridA'uN رَدَأَ *ridāʿun*, ridA'aN رَدَأَ *ridāʿan*, jI'a جِئَ *ǧīʿa*, radI'iN رَدِيَ *radīʿin*, sU'uN سُوءُ *sūʿun*, .daw'uN ضَوءٌ *ḍawʿun*, qay'iN قَيَّ *qayʿin*, ~sifA'I شِفَائِي *šifāʿi*, man~sa'I مَنْشِئِي *manšaʿi*, nisA'uN نِسَاءٌ *nisāʿun*, nisA'u-hu نِسَاؤُهُ *nisāʿu-hu*, nisA'i-hi نِسَائِهِ *nisāʿi-hi*, nisA'I نِسَائِي *nisāʿi*.

~say'uN سَيَّ *šayʿun*, ~say'aN شَيْئًا *šayʿan*, ~say'iN شَيْءٍ *šayʿin*, al-~say'-u أَلشَّيْءُ *aš-šayʿu*, 'a~syA'-u أَشْيَاءُ *ašyāʿu*, 'a~syA'-a أَشْيَاءُ *ašyāʿa*, .zim'aN ظِمًّا *ẓimʿan*, radI'aN رَدِيًّا *radīʿan*.

taḥfīf^u 'l-hamzatⁱ: if the *hamzah* has *ǧazmah* and is preceded by *'alif hamzatum*, it must be changed into the letter of prolongation that is homogeneous with the preceding vowel; hence: 'a'mana آمَنَ *āmana*, 'u'minu أَوْمِنُ *ūminu*, 'i'mAnuN إِيمَانٌ *īmān^{un}*. For other possible ways of encoding such sequences, see on the preceding page (*hamzah* followed by و and ي) and the *maddah* on the current page.

Imperatives of verbs that have the *hamzah* as the first radical are other cases of *taḥfīf^u 'l-hamzatⁱ:* i'sir إِيسِرْ *īsir*, i'_dan اِدْنْ *īdan*, u'mul اُوْمَلْ *ūmul*. arablutext also provides ways of encoding those words when the initial *'alif* comes into *waṣl*, so as to make the *'alif waṣl* fall away when preceded by وَ or فَ: wa-'sir وَأَسِرْ *wa-'sir*, fa-'_dan فَادْنْ *fa-'dan*, fa-'ti فَاتِ *fa-'ti*, wa-'tamirUA وَأَتَمِرُوا *wa-'tamirū*; or be retained outside the imperative, as in fa-i'tazarat فَاتَتَزَرَّتْ *fa-'tazarat*, ba`da i'tilAfIn بَعْدَ اِتِّلَافٍ *ba`da 'tilāfⁱⁿ*.

The strange spelling of mi'at^{un}: mi'aTuN مَائَةٌ *mi'at^{un}*, mi'atAnي مَائَتَانِ *mi'atāni*, mi'atayni مَائَتَيْنِ *mi'atayni*, mi'Una مِئُونَ *mi'una*, mi'AtuN مِئَاتٌ *mi'ātun*, mi'aN_A مَائِي *mi'aⁿ*. Of course, the 'pipe' character can be used to prevent this rule from being applied (see section 4.5 on page 24): mi'a|TuN مِئَةٌ *mi'at^{un}*.

maddah At the beginning of a syllable, *ʿalif* with *hamzah* and *fathah* (أ) followed by *ʿalifu ʿl-maddi* (*ʿalif* of prolongation) or *ʿalif* with *hamzah* and *ǧazmah* (إ) are both represented in writing *ʿalif* with *maddah*: آ (see Wright 1896, i. 25 A–B).

Hence one should keep to this distinction and encode 'a'kulu أَكُلُ *ʿākulu* and 'Akilun أَكِلُ *ʿākil^{un}* respectively.

arabluatex otherwise determines *al-ʿalif^u ʿl-mamdūdāt^u* by context analysis.

'is'AduN إِسَادُ *ʿisʿād^{un}*, 'AkilUna أَكِلُونُ *ʿākilūna*, 'a'mannA آمَنَّا *ʿāmannā*,
al-qurʿAn-u الْقُرْآنُ *al-qurʿān^u*.
jA'a جَاءَ *ǧāʿa*, yatasA'alUna يَتَسَاءَلُونَ *yatasāʿalūna*, ridA'uN رَدَاءُ *ridā^{un}*,
xaba'A خَبَا *ḫabāʿa*, yaxba'Ani يَخْبَانُ *yahbaʿāni*.

šaddah *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:

ʿallaqa عَلَّقَ *ʿallaqa*, mAdduN مَادَّ *mādd^{un}*, 'ammara أَمَّرَ *ammara*, murruN مَرَّرَ *murr^{un}*.

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-^ˆsams-u الشَّمْسُ *aš-šams^u*, or a^ˆs-^ˆsams-u الشَّمْسُ *aš-šams^u*.

al-tamr-u التَّمْرُ *at-tamr^u*, al-ra.hm_an-u الرَّحْمَنُ *ar-raḥmān^u*, al-
.zulm-u الظُّلْمُ *aẓ-ẓulm^u*, al-lu.gaT-u اللَّغَةُ *al-luǧat^u*.

- (b) With the letters ر, ل, م, و, ي after ن with *ǧazmah*, 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 مِلَّيْ *mil laylⁱⁿ*, 'an yaqtula أَنْ يَقْتُلَ 'ay yaqtula.

With *tanwīn*: kitAbuN mubInuN كِتَابٌ مُبِينٌ *kitāb^{um} mubīn^{un}*.

REM. This particular feature must be put into operation by the `\SetArbDflt*` command explicitly. See above section 2.2.1 on page 6 for further details. Other kinds of assimilations, including the various cases of *'idgām*, will be included in *arabluatex* gradually.

- (c) With the letter ت after the dentals ث, د, ذ, ط, ض, ظ in certain parts of the verb: this kind of assimilation, e.g. لَبِئْتُ for لَبِئْتُ *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 'alif^u 'l-waṣlⁱ At the beginning of a sentence, ^ā is never written, as اَلْحَمْدُ لِلَّهِ; instead, to indicate that the ^ā*alif* is a connective ^ā*alif* ('*alif^u 'l-waṣlⁱ*'), the *hamzah* is omitted and only its accompanying vowel is expressed:

al-.hamd-u li-l-l_ah-i اَلْحَمْدُ لِلّٰهِ *al-ḥamd^u li-l-lāhⁱ*.

As said above on page 6, *fullvoc* is the mode in which *arabluatex* expresses the *sukūn* and the *waṣlah*. *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 be 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 ^ā*alifu 'l-waṣlⁱ* is expressed by the vowel that accompanies the omitted *hamzah*: *<u, a, i>*.

Article: bAb-u 'l-madrasaT-i بَابُ الْمَدْرَسَةِ *bāb^u 'l-madrasatⁱ*, al-maqA laT-u 'l-'_U1_A الْمَقَالَةُ الْأُولَى *al-maqālat^u 'l-'ūlā*, al-lu.gaT-u 'l-'ara biyyaT-u اللُّغَةُ الْعَرَبِيَّةُ *al-luḡat^u 'l-'arabiyyat^u*, fI .sinA`aT-i 'l-.tibb-i إِلَى الْإِسْتِقْصَا فِي صِنَاعَةِ الطِّبِّ *il_ā 'l-intiqā.d-i 'l-ibtidā`i*, fI 'l-ibtidA'-i فِي الْإِبْتِدَاءِ *fi 'l-ibtidāⁱ*, 'abU 'l-wazIr-i أَبُو الْوَزِيرِ *abu 'l-wazīrⁱ*, fa-lamma ra'aW 'l-najm-a فَلَمَّا رَأَوْا النَّجْمَ *fa-lammā ra'awu 'n-nağm^a*.

Particles:—

- (a) *li-*: 'alif^u 'l-waṣlⁱ is omitted in the article أَلْ when it is preceded by the preposition لِ: li-l-rajul-i لِلرَّجُلِ *li-r-rağulⁱ*.
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-laylat-i لِلَّيْلَةِ *li-l-laylatⁱ*.
- (b) *la-*: the same applies to the affirmative particle لَ: la-l-.haqq-u لِلْحَقِّ *la-l-haqq^u*.
- (c) With the other particles, 'alif^u 'l-waṣlⁱ is expressed: fI 'l-madIna T-i فِي الْمَدِينَةِ *fi 'l-madīnatⁱ*, wa-'l-rajul-u وَالرَّجُلُ *wa-'r-rağul^u*, bi-'l-qalam-i بِالْقَلَمِ *bi-'l-qalamⁱ*, bi-'l-ru`b-i بِالرُّعْبِ *bi-'r-ru`bⁱ*.

Perfect active, imperative, nomen actionis: qAla isma قَالَ أَسَمَ *qāla 'sma*, qAla uqtul قَالَ أَقْتُلْ *qāla 'qtul*, huwa inhazama هُوَ أَنهَزَمَ *huwa 'nhazama*, wa-ustu`mila وَأَسْتَعْمَلَ *wa-'stu`mila*, qad-i in.sarafa قَدْ أَنْصَرَفَ *qadi 'nšarafa*, al-iqtidAr-u الْأَقْتِدَارُ *al-iqtidār^u*, 'il_A 'l-inti qA.d-i إِلَى الْأَتَقَاضِ *ila 'l-intiqādⁱ*, law istaqbala لَوْ اسْتَقْبَلَ *lawi 'staqbala*.

Other cases: 'awi ismu-hu أَوْ اسْمُهُ *'awi 'smu-hu*, zayduN ibn-u `amriNU زَيْدُ ابْنِ عَمْرٍو *Zayd^{uni} 'bn^u 'Amrⁱⁿ*,²² `umar-u ibn-u 'l_-ha.t.tAb-i عُمَرُ ابْنِ أَحْطَابٍ *Umar^u 'bn^u 'l-Haṭṭābⁱ*,²³ imru'-u 'l-qays-i إِمْرُؤُ الْقَيْسِ *Imru^u 'l-Qaysⁱ*, la-aymun-u 'l-l_ah-i لَا يُؤْمِنُ اللَّهُ *la-'ymun^u 'l-lāhⁱ*.

'alif^u 'l-waṣlⁱ 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-nAs-i فِي النَّاسِ *fi 'n-nāsⁱ*, 'abU 'l-wazIr-i أَبُو الْوَزِيرِ *abu 'l-wazīrⁱ*,
fI 'l-ibtidA'-i فِي الْإِبْتِدَاءِ *fi 'l-ibtidā'ⁱ*, _dU 'l-i`lAl-i ذُو الْأَعْلَالِ *du 'l-i`lālⁱ*,
maqH_A 'l-'amIr-i مَقْهَى الْأَمِيرِ *maqha 'l-'amīrⁱ*.

'alif^u 'l-waṣlⁱ 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:—

²² “Zayd is the son of ‘Amr”: the second noun is not in apposition to the first, but forms part of the predicate. Hence زَيْدُ ابْنِ عَمْرٍو and not زَيْدُ بْنُ عَمْرٍو, “Zayd, son of ‘Amr”.

²³ “Umar is the son of al-Haṭṭāb” (see footnote 22).

fI`aynay 'l-malik-i فِي عَيْنِي الْمَلِكِ 'aynayi 'l-malikⁱ, ix[^]say 'l-qaw
m-a إخْشَى الْقَوْمَ ihšayi 'l-qawm^a, mu.s.tafaw 'l-l_ah-i مُصْطَفَوْا اللَّهَ muṣṭa-
fawu 'l-lāhⁱ.
ramaW 'l-.hijAraT-a رَمَوْا الْحِجَارَةَ ramawu 'l-ḥiğārat^a, fa-lamma ra'aW
'l-najm-a فَلَمَّا رَأَوْا النَّجْمَ fa-lammā ra'awu 'n-nağm^a.

'alif^u 'l-waṣlⁱ 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 *kasrah*; in most of the cases (Wright 1896, i. 22 A–C), it is encoded explicitly, like so:—

'antum 'l-kA_dib-Una أَنْتُمْ الْكَاذِبُونَ 'antum 'l-kādib^{ūna}, ra'aytumu
'l-rajul-a رَأَيْتُمُ الرَّجُلَ ra'aytumu 'r-rağul^a, mani 'l-ka_d_dAb-u مَنِ
الْكَذَّابُ mani 'l-kaddāb^u, qatalati 'l-rUm-u قَتَلَتِ الرُّومُ qatalati 'r-Rūm^u.

However, the Arabic script does not show the *kasrah* or the *ḍammah* which may be 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 مُحَمَّدٌ النَّبِيُّ Muḥammad^{uni} 'n-nabī, salAmuN ud_hulUA
سَلَامٌ أَدْخُلُوا salām^{unu} 'dhulū, qa.sIdata-hu fI qatl-i \uc{'a}bI \uc{m}
uslimiN 'llatI yaqUlu fI-hA قَتَلَ أَبِي مُسْلِمٍ الَّذِي يَقُولُ فِيهَا
hu fī qatlⁱ 'Abī Muslimⁱⁿⁱ 'llatī yaqūlu fī-hā.

4.3 Special orthographies

The name of God The name of God, اللَّهُ, is compounded of the article اَلْ, and إِلَٰه (noted إِلَهِ with the defective 'alif) so that it becomes اِلَٰلَٰه; then the *hamzah* is suppressed, its vowel being transferred to the ل before it, so that there remains اَللَّهِ (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 17), the name of God is therefore encoded al-l_ah-u or 'l-l_ah-u:—

al-l_ah-u اللهُ *al-lāh^u*, yA|²⁴ al-l_ah-u اللهُ يَا *yā al-lāh^u*, 'a-fa|²⁵-al-l_ah-i la-ta.g`alanna أَفَالله لَتَعْلَنَ 'a-fa-al-lāhⁱ la-tag'alanna, bi-'l-l_ah-i بالله *bi-'l-lāhⁱ*, wa-'l-l_ah-i والله *wa-'l-lāhⁱ*, bi-sm-i 'l-l_ah-i بِسْمِ *bi-smⁱ 'l-lāhⁱ*, al-.hamd-u li-l-l_ah-i الْحَمْدُ لِلّهِ *al-ḥamd^u li-l-lāhⁱ*, li-l-l_ah-i 'l-qA'il-u الْقَائِلُ لِلّهِ *li-l-lāhⁱ 'l-qā'il^u*.

The conjunctive الَّذِي 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-malik-i 'lla_dI ya.zlimu 'l-nAs-a أَخَافُ مِنَ الْمَلِكِ 'ahāfu mina 'l-malikⁱ 'lladī yazlimu 'n-nās^a, `udtu 'l-`say_h-a 'lla_dI huwa marI.duN عُدْتُ الشَّيْخَ الَّذِي هُوَ مَرِيضٌ 'udtu 'š-šayh^a 'lladī huwa marīd^{un}, mA 'anA bi-'lla_dI qA'iluN la-ka `say'aN مَا أَنَا بِالَّذِي قَائِلٌ لَكَ شَيْئًا mā 'anā bi-'lladī qā'il^{un} la-ka šay'^{an}.

'ari-nA 'lla_dayni 'a.dallA-nA mina 'l-jinn-i wa-'l-'ins-i أَرِنَا الَّذِينَ الَّذِينَ أَضَلَّانَا مِنَ الْجِنَّ وَالْإِنْسِ 'ari-na 'lladayni 'adallā-nā mina 'l-ḡinnⁱ wa-'l-'insⁱ.

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

fa-'innA na_dkuru 'l-.sawt-ayni 'l-la_dayni rawaynA-humA `an ja.h.zaT-a فَإِنَّا نَذْكُرُ الصَّوْتَيْنِ اللَّذَيْنِ رَوَيْنَاهُمَا عَنْ جَهْظَةِ 'fa-'innā nadkuru 'š-šawt^{ayni} 'l-ladayni rawaynā-humā `an Ġaḥẓat^a.

And also: al-la_dAni اللَّذَانِ *al-ladāni*, al-la_dayni اللَّذَيْنِ *al-ladayni*, al-latAni اللَّتَانِ *al-latāni*, al-latayni اللَّتَيْنِ *al-latayni*, al-lAtI اللَّاتِي *al-lātī*, al-lA' |Ati²⁶ اللَّائِي *al-lā'ī*, al-lA' I اللَّائِي *al-lā'ī*, and so forth.

²⁴Note the “pipe” character ‘|’ here after yA and below after fa before footnote mark 25: it is needed by the dm̄g 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. ‘alā ʔ-ṭarīqⁱ. See section 4.5 on page 24 on the “pipe” and section 8 on page 39 on dm̄g mode.

²⁵See footnote 24.

²⁶Note here the “pipe” character ‘|’: as already stated on page 16, the sequence 'A usually encodes 'alif with hamzah followed by 'alif of prolongation, which is represented in writing 'alif with maddah: ٱ. The “pipe” character prevents this rule from being applied. See section 4.5 on page 24.

4.4 Quoting

It is here referred to “quoting” after the `arabtex` package.²⁷ 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: عَلَّمَ علم الهيئة; is it عَلَّمَ, *he was taught the science of astronomy*, or عَلَّمَ, *he taught the science of astronomy*? In order to disambiguate this clause, it may be sensible to put a *dammah* above the first عَلَّمَ علم الهيئة, 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 *dammah*, *fathah* or *kasrah* appear above the appropriate consonant:—

`ullima `ilm-a 'l-hay'aT-i علم الهيئة ullima 'ilm^a 'l-hay'atⁱ,
ya.gz"UA يغزوا yaǧzū.

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

wa-'innA sawfa tudriku-nA 'l-manAyA muqadd"araT"aN وإنا سوف
تدريكم المنيّة المقدرة wa-'innā sawfa tudriku-na 'l-manāyā muqaddarat^{an}.

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

qAla isma`" اسمع qāla 'sma', jA'at" hinduN جَاءَتْ هِنْدُ gā'at
Hind^{un}, ṣabIhuN bi-man q"u.ti`at" qadamA-hu شَبِيه بِن قُطِعَتْ
qadamā-hu. šabīh^{un} bi-man quṭi'at qadamā-hu.

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

wa-"ust"u`mila وَأَسْتَعْمَل wa-'stu'mila, huwa "inhazama هُوَ أَنْهَزَم huwa
'nhazama, al-"intiqA.d-u الْإِنْتِقَاض al-intiqāḍ^u.

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:—

²⁷See Lagally (2004, p. 22)

q"Ala q"A'iluN قَالَ قَاتِلٌ *qāla qā'il^{un}*, ibn-u 'abI 'u.saybi`aT-
 "a أَصْبَعَةً *Ibn^u 'Abī 'Uṣaybi'at^a*.

(b) *tanwīn*:—

madInaT"aN مَدِينَةٌ *madīnat^{an}*, bAb"aN بَابًا *bāb^{an}*, hud"aN_A هُدًى *hudaⁿ*,
 ^say'"iN سَيِّئٌ *sayⁱⁿ*.

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

(a) *ʿalif^u ʿl-waṣlⁱ*:—

fI "istiq.sA'-iN فِي اسْتِقْصَاءٍ *fi 'stiqṣā'ⁱⁿ*, wa-"istiq.sA'-uN وَأَسْتِقْصَاءٌ
wa-'stiqṣā'^{un}, qAla "uhrub fa-lan tuqtala قَاتِلَ أَهْرُبٍ فَلَنْ تُقْتَلَ
'hrub fa-lan tuqtala.

(b) *sukūn*:—

qAla "uqtul" fa-lan tuqtala قَاتِلَ أَهْرُبٍ فَلَنْ تُقْتَلَ *qāla 'qtul fa-lan tuq-*
tala, mA jA'at" mini imra'aTiN مَا جَاءَتْ مِنْ امْرَأَةٍ *mā ḡā'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” can be used to take away any short vowel (or *tanwīn*, as seen above) or any *sukūn*:—

al-jamr-u 'l-.sayfiyy-u 'lla_dI kAna bi-q"rAn" |nUn-a أَجْرُ الصَّيْفِيِّ
 al-ḡamr^u 'ṣ-sayfiyy^u 'lladī kāna bi-Qrānnūn^a.

4.4.1 Quoting the hamzah

As said above in section 4.2 on page 15, the *hamzah* is always written ⟨ ʾ ⟩, its carrier being determined by contextual analysis. “Quoting” that straight single quote character like so: ⟨ " ʾ ⟩ allows to determine the carrier of the *hamzah* freely, without any consideration for the context. Table 5 gives the equivalents for all the possible carriers the *hamzah* may take.

| Letter | Transliteration ²⁸ | | | ArabT _E X notation |
|--------|-------------------------------|-----|---------|-------------------------------|
| | dmg | loc | arabica | |
| ء | ʾ | ʾ | ء | " ʾ |

Table 5: “Quoted” *hamzah*

²⁸See below section 8 on page 39.

| Letter | Transliteration | | | ArabTeX notation |
|--------|-----------------|-----|---------|------------------|
| | dmg | loc | arabica | |
| آ | 'ā | 'ā | 'ā | A'' |
| أ | ' | ' | ' | a'' |
| إ | ' | ' | ' | u'' |
| ؤ | ' | ' | ' | w'' |
| إ | ' | ' | ' | i'' |
| ئ | ' | ' | ' | y'' |

Table 5: “Quoted” *hamzah*

As one can see from table 5 on the preceding page, the carrier of the *hamzah* is inferred from the letter that precedes the straight double quote (<”). Of course, any “quoted” *hamzah* may take a short vowel, which is to be written *after* the ArabTeX equivalent for the *hamzah* itself, namely (<'). For example, ؤ is encoded <w'' 'a>, while ؤ 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 pages 22–23.

'a`dA'ukum أَعْدَاؤُكُمْ 'a`dā'ukum, 'a`dA|'' 'ukum أَعْدَاءُكُمْ 'a`dā'ukum, 'a`dA'ikum أَعْدَائِكُمْ 'a`dā'ikum, 'a`dA|'' 'ikum أَعْدَاءُكُمْ 'a`dā'ikum.

4.5 The “pipe” character (|)

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

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 24 on page 21 and in footnote 26 on page 21. One more example follows:—

bi-qraAn|nUn-a بَقْرَانُونْ²⁹ *bi-Qrānnūn*^a, “in Crannon” (Thessaly, Greece).

As one can see, the “pipe” character between the two <n> prevents the necessary *tašdīd* rule (page 17) from being applied.

4.6 Putting back on broken contextual analysis rules

In complex documents such as critical editions where footnotes and other kind of annotations can be particularly abundant, the contextual analysis rules that are described above may be broken by L^ATeX commands. To take an example, consider the following:—

²⁹See more context on the preceding page.


```

1 This is wrong:
2 \begin{arab}[fullvoc]
3   fa-lammA ra'aW\LRfootnote{A footnote which interferes with the
4     contextual analysis.} 'l-na^gma...
5 \end{arab}

```

This is wrong:

قَلْبًا رَأَوْا^a النَّجْمَ...

^a A footnote which interferes with the contextual analysis.

According to the rule stated on page 19, the diphthong in *ra'aw* must be resolved into two simple vowels before the *'alif^u* *'l-waṣlⁱ*, as رَأَوْا النَّجْمَ.

`\arbnnull` The `\arbnnull` command is provided so as to put back on contextual analysis rules in such situations. It takes as argument the word that must be brought back for any given rule to be applied as it ought to. Depending on the contexts that have to be restored, `\arbnnull` may be found just after or before Arabic words.

In any case, *no space must be left* after or before the Arabic word that `\arbnnull` is applied to.

The following shows how the Arabic should have been written in the preceding example and gives further illustrations of the same technique:—

```

1 \begin {arab}[fullvoc]
2   fa-lammA ra'aW\arbnnull{'l-na^gma}\LRfootnote{A footnote which
3     interferes with the contextual analysis.} 'l-na^gma...
4
5   qAla\LRfootnote{A footnote which interferes with the contextual
6     analysis.} \arbnnull{qAla}uhrub fa-lan tuqtala.
7
8   \uc{z}ayduN\arbnnull{ibnu}\LRfootnote{A footnote which
9     interferes with the contextual analysis.}
10  \arbnnull{zayduN}ibn-u \uc{'a}mriNU.\LRfootnote{See
11    \vref{fn:zayd-is-son}.}
12 \end{arab}
13
14 \begin{arab}[trans]
15   \uc{z}ayduN\arbnnull{ibnu}\LRfootnote{A footnote which
16     interferes with the contextual analysis.}
17   \arbnnull{zayduN}ibn-u \uc{'a}mriNU.\LRfootnote{See
18     \vref{fn:zayd-is-son}.}
19 \end{arab}

```

فَلَمَّا رَأَوْا^a النَّجْمَ...
 قَالَ^b أَهْرَبْ فَلَنْ تَقْتُلَ.
 زَيْدٌ^c ابْنُ عَمْرٍو^d.

Zayd^{unie} 'bn^u Amrⁱⁿ!^f

^aA footnote which interferes with the contextual analysis.

^bA footnote which interferes with the contextual analysis.

^cA footnote which interferes with the contextual analysis.

^dSee footnote 22 on page 19.

^eA footnote which interferes with the contextual analysis.

^fSee footnote 22 on page 19.

4.7 Stretching characters: the taṭwīl

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

.hunayn-u bn-u 'is.h--_aq-a حُنَيْنُ بْنُ إِسْحَاقَ *Hunayn^u bn^u 'Ishāq^a*

4.8 Digits

4.8.1 Numerical figures

The *Indian numbers*, *ar-raqam^u 'l-hindiyy^u*, 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 sanaT-i 1024 ١٠٢٤ فِي سَنَةِ

4.8.2 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 `\abjad{<number>}` command in any of the `voc`, `fullvoc` and `novoc` modes, where <number> 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, `arabluatex` expresses the *ʿabǧad* numbers in Roman numerals if it finds the `\abjad` command in any of the transliteration modes.

REM. *b.* `\abjad` may also be found outside Arabic environments. In that case, `arabluatex` 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 4.10 on the following page:—

The `\arb[trans]{ʿabjad}` number for 1874 is `\abjad{1874}` The *ʿabǧad* number for 1874 is غصعد.

The `\arb[trans]{ʿabjad}` number for 1874 is `\aemph*{\abjad{1874}}` The *ʿabǧad* number for 1874 is غصعد.

`\abjad` may also be used to convert values of counters into *ʿabǧad* numbers, like so:—

```
1 The \arb[trans]{ʿabǧad} number for the current page (\thepage) is
2   \abjad{\thepage}.
```

The *ʿabǧad* number for the current page (27) is ٢٧.

This technique can be used to produce abjad-numbered lists as will be demonstrated on page 53.

4.9 Additional characters

In the manuscripts, the unpointed letters, *al-ḥurūfu ʿl-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 ³⁰ | | | ArabTeX notation |
|--------|-------------------------------|----------|----------|------------------|
| | dmg | loc | arabica | |
| ب | <i>b</i> | <i>b</i> | <i>b</i> | .b |
| د | <i>d</i> | <i>d</i> | <i>d</i> | ˆd |
| ف | <i>f</i> | <i>f</i> | <i>f</i> | .f |
| ق | <i>q</i> | <i>q</i> | <i>q</i> | .q |
| ك | <i>k</i> | <i>k</i> | <i>k</i> | .k |
| ن | <i>n</i> | <i>n</i> | <i>n</i> | .n |
| و | <i>w</i> | <i>w</i> | <i>w</i> | ((|

Table 6: Additional Arabic codings

³⁰See below section 8 on page 39.

| Letter | Transliteration | | | ArabT _E X notation |
|--------|-----------------|-----|---------|-------------------------------|
| | dmg | loc | arabica | |
| ﴿ |) |) |) |) |

Table 6: Additional Arabic codings

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

4.10 Arabic emphasis

As already seen in section 4.8.2 on page 26, 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 `\aemph{Arabic text}` command to use the same technique to emphasize words, like so:—

`\abjad{45}: kitAbu-hu \aemph{fI 'l-`AdAt-i}` 45: مة: كُأَبُهُ فِي الْعَادَاتِ
kitābu-hu fī 'l-Ādātⁱ.

REM. a. As the above example shows, `arabluatex` places the horizontal stroke *under* the emphasized words in any of the transliteration modes.

REM. b. `\aemph*` is also provided should one wish to always have the horizontal stroke printed over the emphasized words, like so: `\abjad{45}: kitAbu-hu \aemph*{fI 'l-`AdAt-i}`
45: مة: كُأَبُهُ فِي الْعَادَاتِ
kitābu-hu fī 'l-Ādātⁱ.

5 Arabic poetry

`arabluatex` provides a special environment for typesetting Arabic poetry. Every line in this environment must end with `\\`.

`arabverse` The `arabverse` environment may take up to six optional ‘named arguments’ each of which is set using the syntax `<key>=<value>`, like so:—

```

1 \begin{arabverse}[key1=value1, key2=value2, ...]
2 <verses>
3 \end{arabverse}
```

The description of the optional arguments follows:—

`mode` `mode=<mode>`, either `voc`, `fullvoc`, `novoc` or `trans`. The default mode is the one that is set at load time as already seen section 2.2 on page 5.

`width` `width=<length>` Default: 0.3\linewidth

The default width of each hemistich that the verse consists of. It may be expressed in any accepted unit of measurement, such as 4cm or 2in. However, one must keep

in mind that the total length of the two hemistichs added to the one of the gutter that separates them must not exceed the length of the base line, unless one wishes to have the hemistichs distributed on subsequent lines.

gutter `gutter=<width>` Default: 0.15 x (hemistich width)

The gutter consists of the blank space that is between the two hemistichs. By default, it is commensurate with the width of the hemistich, but it may be expressed in any accepted unit of measurement as well.

metre `metre=<name>` Default: none

If the name of the metre is expressed, it is printed after the lines and set flush left in `voc`, `fullvoc` and `novoc` modes or flush right in `trans` mode.

delim `delim=true|false` Default: false

This named argument does not need a value as it defaults to `true` if it is used. If so, a delimiter is printed between each of the hemistichs. By default, it is set to the ‘star’ character ‘*’. The `\SetHemistichDelim{<delimiter>}` command may be used at any point of the document to change this default setting.

`\SetHemistichDelim`

utf `utf=true|false` Default: false

As the preceding one, this named argument does not need a value as it defaults to `true` if it is used. If so, Unicode Arabic input is expected in the `arabverse` environment instead of ASCII ArabTeX or Buckwalter input schemes. See section 10 on page 47 for more details.

\bayt Inside the `arabverse` environment, each line is typeset by the `\bayt` command which takes two mandatory arguments and may accept one optional argument. Additionally, every `\bayt` command *must* be followed with `\\` like so:—

```
\bayt{<ṣadr>}[<tadwīr>] {<ʿağuz>}\\
```

That two subsequent hemistichs should be connected with one another is technically named *tadwīr*. Should that happen, either the *ṣadr* or the *ʿağuz* or both of them, may be connected to one another by letters that are naturally bound to the following or the preceding ones over the *tadwīr*. The optional argument of the `\bayt` command is designed to deal with the various situations that may arise:—

- (a) If the two hemistichs be connected with one another by a prominent horizontal flexible stroke, the *taṭwīl* should be used, like so: `--` (see section 4.7 on page 26). Of course, the ending word of the *ṣadr* and the word at the commencement of the *ʿağuz* must have the *taṭwīl* too so that the proper shapes of the letters be selected. Consider for example the following:—

```
1 \begin{arabverse}[mode=fullvoc, width=.3\linewidth]
2 \bayt{1A 'ar_A man `ahidtu fI-hA fa-'abkI
   'l---}[--]{---yawma
3     dalhaN wa-mA yaruddu 'l-bukA'u}\\
4 \end{arabverse}
```

لَا أَرَى مَنْ عَهِدْتُ فِيهَا فَأُبْكِي الْيَوْمَ دَلْهًا وَمَا يَرُدُّ الْبُكَاءُ

As one can see, *triple hyphens* have been used. In the *ṣadr*, the first hyphen triggers the rules that are related to the definite article and the *ʿalif*³¹ *ʿl-waṣl*, while the following two select the figure of the letter *lām* connected with a following letter. In the *ʿağuz*, the last two hyphens select the letter *yā* connected with a preceding letter, while the first one is simply discarded in this mode, but still may appear as it should, if the *trans* mode be selected:—

```
1 \begin{arabverse}[mode=trans, width=.4\linewidth]
2 \bayt{1A 'ar_A man `ahidtu fI-hA fa-'abkI
   'l---}{--}{---yawma
3 dalhaN wa-mA yaruddu 'l-bukA'u}\\
4 \end{arabverse}
```

lā 'arā man 'ahidtu fī-hā fa-'abki 'l- -yawma dāh^{an} wa-mā yaruddu 'l-bukā'u

- (b) In some other cases, it may seem difficult, if not fairly impossible, to split a given word into two parts. This happens mostly because of the *šaddah*. Consider for example the following:—

```
1 \begin{arabverse}[mode=fullvoc, width=.25\linewidth,
   gutter=1cm]
2 \bayt{.gayra 'annI qad 'asta`Inu `al_A 'l-ha--}{--mmi
   ]{'i_dA
3 _haffa bi-'l-_tawiiyi 'l-na^gA'u}\\
4 \bayt{bi-zaf--UfiN ka-'anna-hA hiq--laTuN}[ 'ummu ]{ri'AlIN
5 dawwiyyaTuN saqfA'u}\\
6 \end{arabverse}
```

غَيْرَ أَنِّي قَدْ أَسْتَعِينُ عَلَى الْهَمِّ إِذَا خَفَّ بِالْثَوِيِّ النَّجَاءُ
بِزَفُوفٍ كَانَتْهَا هِقْلَةٌ أَمْ رِثَالٍ دَوِيَّةٌ سَقْفَاءُ

In the first line, the word *أَلْهَمَ* should be split into *أَلْهَمَ* as the first part of it belongs to the *ṣadr* and the second to the *ʿağuz*. One solution to avoid splitting this word in such a way is to write inside the *tadwīr* the part of it that belongs to either hemistich, without omitting to add a space after it. In the second line,

³¹See section 4.2 on page 18

the word أَمُّ should be split into أُمُّ , so that the only way to avoid splitting it into two parts is to write it all inside the *tadwīr*. In that case, as the word is to be placed in the middle, it has been surrounded by spaces.

Sealing and distortion of characters The `arabverse` environment and the `\bayt` command are designed to typeset the verses in a two-column, fixed width layout. This may result in a somewhat distorted text. Should that happen, one may adapt the layout by modifying the values of the above described `width` and `gutter` named arguments until the visual aspect of the layout be satisfactory. It has to be noted that distortion and warping may be even more perceptible in Roman than in Arabic characters.

Footnotes Footnotes are not set by default inside the `\bayt` command, but there are two easy ways to have them printed.

If they are little in number, each footnote may be split into pairs of `\footnote mark{}` (please mind the braces or “declare” `footnotemark` using `\MkArbBreak` to take it out of the Arabic environment³²) in the argument of the `\bayt` command and `\footnotetext` outside the `\bayt` command.

If the footnotes are abundant in number, it is advised to load the `footnotehyper` package which `arabluatex` will then use to typeset any kind of footnote that is called from the arguments of the `\bayt` command.³³

Line numbering Inside the `arabverse` environment, the `linenumbers` environment of the `lineno` package can be used to have the lines of succeeding verses numbered. Please refer to the documentation of this package for more information or to the example below for a basic implementation of this technique.

5.1 Example

Here follow the first lines of *Imru’u ‘l-Qaysi’s Mu‘allaqah*. In this example, `\SetArbDflt*` has been selected so as to mark the *‘idgām* that is fit to this declamatory poetry:—³⁴

```
1 \begin{arab}[fullvoc]
2   qAla \uc{i}mru'u 'l-\uc{q}aysi fI mu`allaqati-hi:
3 \end{arab}
4
5 \begin{arabverse}[mode=fullvoc, metre={(al-.darbu 'l-_tAnI mina
6   'l-`arU.di 'l-'_U1_A mina 'l-.tawIli)}}]
```

³²See section 11.1 on page 50.

³³The `footnote` package can also be used for the same effect. However, it must be loaded *after* `arabluatex`.

³⁴Please note that for the time being only the assimilation rules that are laid on item (b) on page 17 are applied. See section 2.2.1 on page 6 for more information. None of the editions of the *Mu‘allaqāt* that I know of feature the *‘idgām* in the Arabic text, although it is often strongly marked in declamation.

```

7 \SetArbDflt*
8 \begin{linenumbers*}
9 \bayt{qifA nabki min _dikr_A .habIbiN wa-manzili}{bi-saq.ti
10 'l-liw_A bayna \uc{'l-d}a_hUli fa-\uc{.h}awmali}\\
11 \bayt{fa-\uc{t}U.di.ha fa-'l-\uc{m}iqrATi lam ya`fu
12 rasmu-hA}{limA nasa`gat-hA min `ganUbiN wa-`sam'ali}\\
13 \bayt{tar_A ba`ara 'l-'ar'Ami fI `ara.sAti-hA}{wa-qI`Ani-hA
14 ka-'anna-hu .habbu fulfuli}\\
15 \bayt{ka-'annI .gadATa 'l-bayni yawma ta.hammalUA}{lad_A
16 samurAti 'l-.hayyi nAqifu .han.zali}\\
17 \bayt{wuqUfaN bi-hA .sa.hbI `alayya ma.tiyya-hum}{yaqUlUna lA
18 tahlik 'asaN_A wa-ta`gammali}\\
19 \bayt{wa-'inna `sifa'I `abraTuN muharAqaTuN}{fa-hal `inda
rasmIN
20 dAsiriN min mu`awwali}\\
21 \end{linenumbers*}
22 \end{arabverse}

```

قَالَ أَمْرُؤُ الْقَيْسِ فِي مَعْلَقَتِهِ:

| | | |
|---|---|--|
| 1 | بَسَقَطَ اللَّوَى بَيْنَ الدَّخُولِ فَخَوَّلَ | فَقَنَا نَبِكُ مِنْ ذِكْرِي حَبِيبٌ وَمَنْزِلُ |
| 2 | لَمَّا نَسَجْتَهَا مِنْ جَنُوبٍ وَشَمَالٍ | فَتَوَضَّعَ فَاَلْمَقْرَأَةُ لَمْ يَعْفُ رَسْمُهَا |
| 3 | وَقِيَعَانِهَا كَأَنَّهُ حَبٌّ فَلَفُلٍ | تَرَى بَعَرَ الْأَرَامِ فِي عَرَصَاتِهَا |
| 4 | لَدَى سَمَرَاتٍ أَلْحَى نَافِقُ حَنْظَلٍ | كَأَنِّي غَدَاةَ الْبَيْنِ يَوْمَ تَحَلَّلُوا |
| 5 | يَقُولُونَ لَا تَهْلِكُ أَسَى وَتَجَلَّى | وَوُفُوفًا بِهَا صَحْبِي عَلَيَّ مَطْلَبُهُمْ |
| 6 | فَهَلْ عِنْدَ رَسْمٍ دَاسِرٍ مِنْ مُعَوَّلٍ | وَإِنْ شِفَائِي عَبْرَةُ مَهْرَاقَةٍ |

(الضَرْبُ الثَّانِي مِنَ الْعَرُوضِ الْأُولَى مِنَ الطَّوِيلِ)

qāla Imru'ū 'l-Qaysi fī mu'allaqati-hi:

| | | |
|---|---|--|
| 1 | <i>qifā nabki min dīkrā ḥabīb^{iw} wa-manzili</i> | <i>bi-saqti 'l-liwā bayna 'd-Daḥūli fa-Ḥawmali</i> |
| 2 | <i>fa-Tūḍiḥa fa-'l-Migrāti lam ya`fu rasmu-hā</i> | <i>limā nasaḡat-hā min ḡanūb^{iw} wa-šam'ali</i> |
| 3 | <i>tarā ba`ara 'l-'ar'āmi fī `arašāti-hā</i> | <i>wa-qī'āni-hā ka-'anna-hu ḥabbu fulfuli</i> |
| 4 | <i>ka-'annī gadāta 'l-bayni yawma taḥammalū</i> | <i>ladā samurāti 'l-ḥayyi nāqifu ḥanzali</i> |
| 5 | <i>wuqūf^{un} bi-hā šaḥbī `alayya maṭiyya-hum</i> | <i>yaqūlūna lā tahlik 'asā^w wa-taḡammali</i> |
| 6 | <i>wa-'inna šifā'ī `abrat^{um} muharāqat^{un}</i> | <i>fa-hal 'inda rasmⁱⁿ dāsir^{im} mim mu'awwali</i> |

(aḍ-ḍarbu 't-tānī min 'l-'arūḍi 'l-'ulā min 't-tawīli)

6 Special applications

Linguistics The same horizontal stroke as the *taṭwīl* (see section 4.7 on page 26) may be encoded $\langle B \rangle$; $\langle BB \rangle$ will receive the *tašdīd*. This is useful to make linguistic annotations and comments on vowels:—

Bu Ba Bi BuN BaN BiN $\overset{\text{u}}{\text{---}} \overset{\text{a}}{\text{---}} \overset{\text{i}}{\text{---}}$ $\overset{\text{un}}{\text{---}} \overset{\text{an}}{\text{---}} \overset{\text{in}}{\text{---}}$, BBu BBa BBi $\overset{\text{u}}{\text{---}} \overset{\text{a}}{\text{---}} \overset{\text{i}}{\text{---}}$, B--aN
 $\overset{\text{a}}{\text{---}}$ $\overset{\text{an}}{\text{---}}$, B" $\overset{\text{a}}{\text{---}}$ $\overset{\text{an}}{\text{---}}$.

Brackets The various bracket symbols are useful in technical documents such as critical editions for indicating that some words or some letters must be added or removed. `arabluatex` will automatically fit those symbols to the direction of the text. For the time being, the following symbols are supported:

- parentheses: `()`
- square brackets: `[]`
- angle brackets: `<>`
- braces: `{ }`

`\abraces` Parentheses, square and angle brackets may be input directly at the keyboard; however, words or letters that are to be read between braces must be passed as arguments to the `\abraces` command:—

```
1 \begin{arab}
2   \abraces{wa-qAla} 'inna 'abI kAna mina 'l-muqAtilaTi
3   wa-kAna--<--t> 'ummI min `u.zamA'i buyUti 'l-zamAzimaTi.
4 \end{arab}
```

{وَقَالَ} إِنَّ أَيْ كَانَ مِنَ الْمُقَاتِلَةِ وَكَانَتْ < أُمِّي مِنْ عُظْمَاءِ بُيُوتِ الزَّمَاذِمَةِ.

Additional Arabic marks In addition to common letters, many symbols and ligatures are encoded in Arabic Unicode standard, such as honorifics consisting of complex ligatures, and annotation signs used in the *Qurʾān* or in classical poetry.

`\arbmark` `\arbmark{<shorthand>}` can be used to insert such characters either in Unicode or in romanized Arabic environments. It takes as argument a shorthand defined beforehand in a default list which consists of the following at the time of writing:—

| Codepoint | Shorthand | Glyph | Transliteration |
|-----------|-----------|---------------------------------------|-------------------------------------|
| FDFD | bismillah | بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ | bi-'smi 'Llāhi 'r-rahmāni 'r-rahīmi |
| FDF5 | salam | صَلَام | ṣalla 'Llāhu 'alay-hi wa-sallama |

Table 7: Additional Arabic marks

New feature
v1.4.3

New feature
v1.11

| Codepoint | Shorthand | Glyph | Transliteration |
|-----------|-----------|---------|--|
| FDFA | slm | سَلَامٌ | <i>ṣalla 'Lāhu 'alay-hi wa-sallama</i> |
| FDFB | jalla | جَلَالٌ | <i>ǧalla ǧalāla-hu</i> |

Table 7: Additional Arabic marks

New feature
v1.11

`\newarbmark`

`\newarbmark` is also provided should one wish to define new marks in addition to the marks defined above. This command takes three arguments, like so:—

```
\newarbmark{<shorthand>}{<RTL codepoint>}{<LTR rendition>}
```

As regards the right-to-left codepoint, it may be either typed in Unicode or selected as Unicode codepoint. To that end, the \LaTeX command `\symbol{"XYZT}` or its plain \TeX variant `\char"XYZT\relax` may be used, where XYZT are uppercase hex digits (0 to 9 or A to F).

It is also possible to use the so-called ‘`~~~~` notation’ like so: `~~~~xyzt`, where xyzt are lowercase hex digits (0 to 9 or a to f).

As regards the third argument (left-to-right rendition), it may be either left empty or typed by means of `\arb[trans]{<arabtex code>}` so as to have it printed in romanized Arabic.

It must be noted that `\newarbmark` expects Arab \TeX input scheme inside `\arb[trans]{}` to the exclusion of buckwalter input scheme.

The example below provides an implementation of this technique. It may be observed that `\arbcOLOR` is used so as to have the marks printed in red:—

```
1 \SetArbDflt *
2 \newarbmark{sly}{\arbcOLOR[red]{~~~~06d6}}{}
3 \newarbmark{jim}{\arbcOLOR[red]{~~~~06da}}{}
4 \begin{arab}
5   sUraTu 'l-nisA'i, 19:
6 \end{arab}
7 \begin{center}
8   \begin{arab}
9     \arbmark{bismillah}
10    \end{arab}
11 \end{center}
12 \begin{arab}[fullvoc]
13   y_a'ayyuhA 'lla_dIna 'a'manUA lA ya.hillu la-kum 'an tari_tUA
14   'l-nisA'a karhaN\arbmark{sly} wa-lA ta`.dulU-hunna li-ta_dhabUA
15   bi-ba`.di mA 'a'taytumU-hunna 'illA 'an ya'tIna bi-fA.hi`saTiN
16   mubayyinaTiN\arbmark{jim} wa-`A`sirU-hunna
17   bi-'l-ma`rUfi\arbmark{jim} fa-'in karihtumU-hunna fa-`as_A_a
18   'an takrahUA ^say'aN wa-ya`g`ala 'l-l_ahu fI-hi _hayraN
```

```

19      ka_tIraN ((19))
20      \end{arab}

```

سُورَةُ النَّسَاءِ، ١٩:

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا يَحِلُّ لَكُمْ أَنْ تَرْثُوا النِّسَاءَ كَرِهًا ^ط وَلَا تَعْضُلُوهُنَّ لِتَذْهَبُوا بِبَعْضِ مَا آتَيْتُمُوهُنَّ إِلَّا أَنْ يَأْتِيَنَّ بِفَاحِشَةٍ مُبِينَةٍ ^ج وَعَاشِرُوهُنَّ بِالْمَعْرُوفِ فَإِنْ كَرِهْتُمُوهُنَّ فَعَسَى أَنْ تَكْرَهُوا شَيْئًا وَيَجْعَلَ اللَّهُ فِيهِ خَيْرًا كَثِيرًا ﴿١٩﴾

7 Color

New feature
v1.12

arabluatex is able to render in color either words, parts of words or diacritics. As the techniques implemented in this section may lead to some complexity, the reader should first become well acquainted with the following points:³⁵—

- The “pipe” character (|, section 4.5 on page 24);
- ‘Quoting’ technique (section 4.4 on page 22), and more specifically ‘quoting the *hamzah*’ (on page 23);
- Putting back on broken contextual analysis rules (section 4.6 on page 24);
- Arabic marks (section 6 on page 33).

`\arbcolor` `\arbcolor` takes the text to be colored into `<color>` as an argument:—

```
\arbcolor[<color>]{<Arabic text>}
```

```

1  \begin{arab}
2    \arbcolor[red]{al-bAbu 'l-hAmisu} fI .tabaqAti 'l-'a.tibbA'i
3    'lla_dIna kAnUA mun_du zamAni \uc{^gAlInUsa} wa-qarIbaN
4    min-hu. \arbcolor[red]{\uc{^gAlInUsu}}: wa-li-na.da` 'awwalaN
5    kalAmaN kulliyyaN fI 'a_hbAri \uc{^gAlInUsa} wa-mA kAna
6    `alay-hi...
7  \end{arab}
8  \begin{arab}[trans]
9    \arbcolor[red]{al-bAbu 'l-hAmisu} fI .tabaqAti 'l-'a.tibbA'i
10   'lla_dIna kAnUA mun_du zamAni \uc{^gAlInUsa} wa-qarIbaN
11   min-hu. \arbcolor[red]{\uc{^gAlInUsu}}: wa-li-na.da` 'awwalaN
12   kalAmaN kulliyyaN fI 'a_hbAri \uc{^gAlInUsa} wa-mA kAna
13   `alay-hi...

```

³⁵Regarding the colors themselves and the way new colors can be defined in addition to those that are already available, please refer to the `xcolor` package.

البَابُ الْخَامِسُ فِي طَبَقَاتِ الْأَطِبَّاءِ الَّذِينَ كَانُوا مِنْدُ زَمَانِ جَالِينُوسَ وَقَرِيبًا مِنْهُ. جَالِينُوسُ: وَلِنَضَعُ أَوَّلًا
 كَلَامًا كُلِّيًّا فِي أَخْبَارِ جَالِينُوسَ وَمَا كَانَ عَلَيْهِ...
al-bābu 'l-ḥāmisu fī ṭabaqāti 'l-aṭibbā'i 'lladīna kānū mundu zamāni
Ġālīnūsa wa-qarīb^{an} min-hu. Ġālīnūsu: wa-lī-naḍa^c 'awwal^{an} kalām^{an}
kullīyy^{an} fī aḥbāri Ġālīnūsa wa-mā kāna alay-hi...

As this example shows, `\arbcOLOR` has been used to render headings in red with the same encoding both in vocalized and in romanized Arabic. The same technique also applies to syllables inside words. `arabluatex` takes care of selecting the appropriate shape of the letters while coloring them:—

‘voc’ mode:

`i^stara\arbcOLOR[brown]{y}tu-hu bi-_tama\arbcOLOR[red]{niN}`
`'a^\arbcOLOR[blue]{^ga}ba-ka` اشْتَرَيْتُهُ بَعْنِي أَعْبِكَ *ištaraytu-hu bi-tama-*
nⁱⁿ 'a'ḡaba-ka.

‘fullvoc’ mode:

`i^stara\arbcOLOR[brown]{y}tu-hu bi-_tama\arbcOLOR[red]{niN}`
`'a^\arbcOLOR[blue]{^ga}ba-ka` اشْتَرَيْتُهُ بَعْنِي أَعْبِكَ *ištaraytu-hu bi-tama-*
nⁱⁿ 'a'ḡaba-ka.

7.1 Tricks of the trade

Diacritics Depending on the mode selected, either `voc`, `novoc` or `fullvoc`, coloring the diacritics requires more attention for the insertion of `\arbcOLOR` may prevent contextual analysis from being applied.

Furthermore, depending on the surrounding letters, the standard encoding of short vowels $\langle u, a, i \rangle$ may result either in diacritics or in a connective *ʾalif* with the *waṣlah* or its accompanying vowel. As for the *sukūn*, it is generated by contextual analysis. Thus applying colors to bare diacritics requires them to have specific encodings.

Table 8 gives the ArabTeX equivalents for the diacritics to be printed inside or just after `\arbcOLOR`.

| Diacritic | Transliteration ³⁶ | | | ArabTeX notation |
|-----------|-------------------------------|----------|----------|------------------|
| | dmg | loc | arabica | |
| ˘ | <i>a</i> | <i>a</i> | <i>a</i> | <code>.a</code> |

Table 8: ArabTeX diacritics for `\arbcOLOR`

³⁶See below section 8 on page 39.

| Diacritic | Transliteration | | | ArabTeX notation |
|-----------|-----------------|-----|---------|------------------|
| | dmg | loc | arabica | |
| ˆ | u | u | u | .u |
| ˙ | i | i | i | .i |
| ˚ | | | | o |

Table 8: ArabTeX diacritics for \arbcolor

The following examples show how the letters, or the diacritics above or under them or both the letters and the diacritics can be rendered in different colors:—

‘voc’ mode:

```
i^staraytu-hu bi-_taman\arbcolor[red]{iN} 'a^g\arbcolor[red]
{.a}ba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ i^staraytu-hu bi-_tamanin 'agaba-ka.

i^staraytu-hu bi-_tama\arbcolor[red]{n}iN 'a^\arbcolor[red]
{^g}.aba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ i^staraytu-hu bi-_tamanin 'agaba-ka.

i^staraytu-hu bi-_tama\arbcolor[red]{n}\arbcolor[blue]{iN}
'a^\arbcolor[red]{^g}\arbcolor[blue]{.a}ba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ
i^staraytu-hu bi-_tamanin 'agaba-ka.
```

‘fullvoc’ mode:

```
i^staray"\arbcolor[red]{o}tu-hu bi-_taman"\arbcolor[red]{iN}
'a^g"\arbcolor[red]{.a}ba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ i^staraytu-hu bi-_tamanin
'agaba-ka.

i^stara\arbcolor[red]{y"}otu-hu bi-_tama\arbcolor[red]{n"}iN
'a^\arbcolor[red]{^g"}.aba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ i^staraytu-hu bi-_tamanin
'agaba-ka.

i^stara\arbcolor[red]{y"}\arbcolor[blue]{o}tu-hu bi-_tama\arb
color[red]{n"}\arbcolor[blue]{iN} 'a^\arbcolor[red]{^g"}\arb
color[blue]{.a}ba-ka أَجَبَكَ بَتْنِ اشْتَرَيْتَهُ i^staraytu-hu bi-_tamanin 'agaba-
ka.
```

As can be seen, fullvoc required the letters y, n and ^g before \arbcolor to be ‘quoted’. Otherwise, unwanted *sukūns* would have been generated because of the absence of a vowel after those consonants.

tanwīn \arbnul must be used with *fathatān* (َ) so as to put back on contextual analysis rules:—

```
mu`allim\arbcOLOR[red]{\arbnul{m}aN} مَعْلَمٌ mu'alliman,
istisqA'\arbcOLOR[red]{\arbnul{A'}aN} اِسْتِسْقَاءٌ istisqā'an,
`say'\arbcOLOR[red]{\arbnul{ay'}aN} شَيْئًا šay'an,
`gAmi`aT|\arbcOLOR[red]{\arbnul{T}aN} جَامِعَةٌ ḡāmi'atan.
```

REM. Note that in the last example (*ḡāmi'at^{an}*), the ‘pipe’ character has been inserted before `\arbcOLOR`. Otherwise, the `dmg` mode of the transliteration rules would have interpreted the *tā'* *marbūṭah* as *final* (e.g. *h* instead of the expected *t*).³⁷

The *tanwīn* preceding a `ى` conveys even more intricate business to the rendering with the utmost accuracy in both romanized and non-romanized modes. First, a new Arabic mark needs to be defined. It should print `ى` in Arabic script and not a thing in transliteration. It is to be appended after `\arbcOLOR`, like so:—

```
1 \newarbmArk {Y}{~~~~0649}{~}
2 \arB{hud\arbcOLOR[red]{aN\arbnul{~}_A}}\arbmArk{Y}}
3 \arB{trans}{hud\arbcOLOR[red]{aN\arbnul{~}_A}}\arbmArk{Y}}
```

هُدًى *hudā'*ⁿ

waṣlah and maddah Both can be generated with the help of `\arbnul`:—

```
wa-\arbcOLOR[red]{\arbnul{wa}i}stisqA'uN وَأَسْتِسْقَاءٌ wa-'stisqā'un38.
fI "al".i-\arbcOLOR[red]{\arbnul{'l-}i}btidA'i فِي الْإِبْتِدَاءِ
fi 'li-'btidā'i.
\arbcOLOR[red]{'a'\arbnul{k}}kulu أَكُلُ 'akulu,
\arbcOLOR[red]{'A'\arbnul{k}}kiluN أَكِلُ 'akilun.
```

The Unicode codepoint of the *maddah* is 0653, while bare *ʿalif* is 0627. So:—

```
1 \newarbmArk {alifmaddahred}{~~~~0627\arbcOLOR[red]{~~~~0653}}%
2 {\arB{trans}{\arbcOLOR[red]{'a'\arbnul{k}}}}
3 \arB{\arbmArk{alifmaddahred}kulu}
4 \arB{trans}{\arbmArk{alifmaddahred}kulu}.
```

أَكُلُ 'akulu.

REM. In the preceding example, any consonant could have been passed as argument to the `\arbnul` command.

³⁷See also on page 43 “Discarding the *ʿrāb*” for more information.

³⁸To the knowledge of the writer, the *waṣlah* alone is not part of the Arabic Unicode block.

šaddah In the following example, it is assumed that the *šaddah* above the letter ل in المَعْلُونِ, *al-mu‘allimūna*, is to be rendered in red. Thus the Arabic mark must generate the *šaddah* alone—of which the Unicode codepoint is 0651—in Arabic script and the letter ‘l’ in transliteration:—

```
1 \newarbmak {lamshaddah}{^^^^0651}{l}
2 \arb[fullvoc]{al-mu`al"\arbcOLOR[red]{\arbmark{lamshaddah}}.imUna}
3 \arb[trans]{al-mu`al"\arbcOLOR[red]{\arbmark{lamshaddah}}.imUna}.
```

المَعْلُونِ *al-mu‘allimūna*.

The definite article and the euphonic tašdīd The intricate business of rendering in color the initial *‘alif al-waṣl* of the definite article followed by a solar consonant must be unraveled.

From the examples provided above, in فِ النَّاسِ *fi ‘n-nāsi*, the initial *‘alif* “*l-waṣl*” can be rendered in red like so: `\arbcOLOR[red]{\arbnUL{al-}a}`. Then, the following two letters, namely l-n, must print the string *lām + nūn + šaddah* in Arabic, and exactly *n-n* in transliteration. Thus an Arabic mark is needed:—

```
1 \newarbmak {lnn}{^^^^0644^^^^0646^^^^0651}{n-n}
2 \arb[fullvoc]{fI\arbnUL{al-}
3 \arbcOLOR[red]{\arbnUL{al-}a}\arbmark{lnn}Asi}
4 \arb[trans]{fI\arbnUL{al-}
5 \arbcOLOR[red]{\arbnUL{al-}a}\arbmark{lnn}Asi}.
```

فِ النَّاسِ *fi ‘n-nāsi*.

hamzah The ‘quoting’ technique provides an easy way to determine the carrier of the *hamzah*, as shown in table 5 on page 23—:

yatasA\arbnUL{‘a}\arbcOLOR[red]{|''}.alUna يَتَسَاءَلُونَ *yatasā‘al-lūna*, ^say\arbcOLOR[red]{|''}\arbnUL{‘}aN شَيْءٌ *šay‘an*, ^say\arbcOLOR[red]{|''}iN شَيْءٍ *šay‘in*, \arbcOLOR[red]{a''}.as\arbcOLOR[red]{y''}.ilaTuN أَسْئَلَةٌ *as‘ilatun*.

8 Transliteration

It may be more appropriate to speak of “romanization” than “transliteration” of Arabic. As seen above in section 2.2 on pages 5–9, the “transliteration mode” may be selected globally or locally.

New feature
v1.8

This mode transliterates the ArabTeX input into one of the accepted standards. As said above on page 6, three standards are supported at present:

dmg *Deutsche Morgenländische Gesellschaft*, which was adopted by the International Convention of Orientalist Scholars in Rome in 1935.³⁹ **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;⁴⁰

arabica *Journal of Arabic and Islamic Studies/Revue d'études arabes et islamiques*: this standard is most widely used by scholars in the field of Arabic studies.⁴¹

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 `\SetTranslitConvention{<mode>}` command, where `<mode>` may be either **dmg**, **loc** or **arabica**. This command is also accepted in the preamble should one wish to set the transliteration mode globally, e.g.:—

```
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 `\SetTranslitStyle{<style>}` command, where `<style>` may be any font shape selection command, e.g. `\upshape`, `\itshape`, `\slshape`, and so forth.

New feature
v1.4

\SetTranslitFont **Font** `\SetTranslitFont{}` allows any specific font to be selected for rendering transliterated text with the font-selecting commands of the `fontspec` or `luaotfload` package. Of course, this font must have been defined properly. To take one example, here is how the *Gentium Plus* font can be used for rendering transliterated text:—

```
1 \newfontfamily\translitfont{Gentium Plus}[Ligatures=TeX]
2 \SetTranslitFont{\translitfont}
```

\uc **Proper names** Proper names or book titles that must have their first letters up-

³⁹See Brockelmann et al. (1935).

⁴⁰See <http://www.loc.gov/catdir/cpsd/roman.html> for the source document concerning Arabic language.

⁴¹See http://www.brill.nl/files/brill.nl/specific/authors_instructions/ARAB.pdf.

percased may be passed as arguments to the `\uc{⟨word⟩}` command. `\uc` 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:—

```
\uc{.hunayn-u} bn-u \uc{'is.h_aq-a} حُنَيْنُ بْنُ إِسْحَقَ Hunaynu bnu
Ishāqa, \uc{`u_tm_an-u} عُثْمَنُ Utmānu, .daraba \uc{zayd-u} bn-u
\uc{`h_alidiN} \uc{sa`d-a} bn-a \uc{`awf-i} bn-i \uc{`abd-i}
\uc{'l-l_ah-i} ضَرْبَ زَيْدُ بْنُ خَلْدٍ سَعْدُ بْنُ عَوْفٍ بْنِ عَبْدِ اللَّهِ daraba Zaydu bnu
Hālidin Sa`da bna Awfi bni Abdi 'Llāhi.
```

However, `\uc` 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 20, the transliteration of مُحَمَّدٌ النَّبِيُّ must be *Muḥammad^{uni} 'n-nabī*, as nouns having the *tanwīn* take a *kasrah* in pronunciation before *'alifu 'l-waṣli*. In that case, encoding مُحَمَّدٌ like so: `\uc{mu.hammaduN}` is wrong, because the closing brace would prevent `arabluatex` from detecting the sequence `⟨-uN⟩` immediately followed by `⟨'l-⟩`. Fortunately, this can be circumvented in a straightforward way by inserting only part of the noun in the argument of `\uc` viz. up to the first letter that is to be uppercased, like so: `\uc{m}u.hammaduN`.

Hyphenation In case transliterated Arabic words break the `TeX` hyphenation algorithm, one may use the `\-` command to insert discretionary hyphens. This command will be discarded in all of the Arabic modes of `arabluatex`, but will be processed by any of the transliteration modes:—

```
\uc{'abU} \uc{bakriN} \uc{mu\-.ham\_-madu} bnu \uc{za\_-ka \_-
riy\_-yA'a} \uc{'l-rAziyyu} أَبُو بَكْرٍ مُحَمَّدُ بْنُ زَكْرِيَّا الرَّازِيُّ Abū Bakrin Mu-
hammadu bnu Zakariyyāa 'r-Rāziyyu.
```

New feature
v1.10

‘Long’ proper names `\uc` is also able to process proper names consisting of several subsequent words:—

```
\arb[trans]{\uc{'abU zaydiN .hunaynu bnu 'is.h_aqa 'l-`ibAdiyyu}}
'Abū Zaydin Hunaynu bnu Ishāqa 'l-Ibādiyyu.
```

New feature
v1.10

`\prname` **Proper names outside Arabic environments** Transliterated proper names inserted in paragraphs of English text should be printed in the same typeface as the surrounding text. `\prname{⟨Arabic proper name⟩}` is provided to that effect:⁴²—

- 1 From `\textcite[i. 23 C]{Wright}`:--- If the name following
- 2 `\arb[fullvoc]{ibnuN}` be that of the mother or the grandfather, the

⁴²Just as `\uc`, `\prname` is also able to process proper names consisting of several subsequent words.

```

3 \arb[fullvoc]{a} is retained; as \arb[fullvoc]{`Is_A ibnu
   maryama},
4 \enquote{Jesus the son of Mary}; \arb[fullvoc]{`ammAru ibnu
5   man.sUrIN}, \enquote{\prname{`ammAr} the (grand)son of
6   \prname{man.sUr}}.

```

From Wright (1896, i. 23 C):— If the name following ^ٱٱبن be that of the mother or the grandfather, the ٱ is retained; as عيسى ٱبن مريم, “Jesus the son of Mary”; عمار ٱبن منصور, “Ammār the (grand)son of Maṣṣūr”.

The following example shows how `\prname` can be used in conjunction with the `nameauth` package to have Arabic proper names printed first in full then in partial forms:⁴³—

```

1 \begin{nameauth}
2   \< Hunayn & \prname{'abU zayd} & \prname{.hunayn}, \prname{{i}bn
3     'is.h_aq al-`ibAdiyy} & > %
4   \< Razi & \prname{'abU bakr mu.hammad ibn zakariyyA'} &
5     \prname{al-rAziyy} & > %
6 \end{nameauth}
7
8 On first occurrence, proper names are printed as \Hunayn, \Razi.
9 Then as \Hunayn, \Razi.

```

On first occurrence, proper names are printed as ʾAbū Zayd Ḥunayn ibn ʾIshāq al-ʾIbādī, ʾAbū Bakr Muḥammad ibn Zakariyyāʾ ar-Rāzī. Then as Ḥunayn, ar-Rāzī.

8.1 Additional note on dmḡ convention

According to Brockelmann et al. (1935, p. 6), Arabic *ʾrāb* may be rendered into dmḡ in three different ways:

- (a) In full: *Amrun*;
- (b) As superscript text: *Amr^{un}*;
- (c) Discarded: *Amr*.

`\arbup` By default, `arabuatex` applies rule (b). Once delimited by a set of Lua functions, *ʾrāb* is passed as an argument on to a `\arbup` command which is set to `\textsuperscript`.

`\NoArbUp` `\NoArbUp` may be used either in the preamble or at any point of the document in case one wishes to apply rule (a). The default rule (b) can be set back with `\ArbUpDflt` at any point of the document.

`\SetArbUp` Finally, `\SetArbUp{<formatting directives>}` can be used to customize the way

⁴³See the documentation of `nameauth` for more details: <https://ctan.org/pkg/nameauth>

ʾirāb is displayed. To take one example, here is how Arabic *ʾirāb* may be rendered as subscript text:—

```
1 \SetArbUp {\textsubscript{#1}}
2 Arabic |dmg| transliteration for \arb{ra'aytu ḡami`aN
3 muhaddamaTaN mi`_danatu-hu}: \arb[trans]{ra'aytu
4 ḡami`aN muhaddamaTaN mi`_danatu-hu.}
```

Arabic dmg transliteration for رَأَيْتُ جَامِعًا مَهْدَمَةً مِثْلَتَهُ: *ra'aytu ḡāmi`an muhaddamat_{an} mi`danatu-hu*.

As shown in the above example, #1 is the token that is replaced with the actual *tanwīn* in the formatting directives of the `\SetArbUp` command.

ʾirāb boundaries Every declinable noun (*muʿrab*) may be declined either with or without *tanwīn*, viz. *munṣarif^{un}* or *ḡayr^u munṣarifⁱⁿ*. The former is automatically parsed by *arabluatex*, whereas the latter has to be delimited with an hyphen, like so:—

munṣarif: mu`allimuN ^{مُعَلِّمٌ} *mu'allim^{un}*, kA'inuN ^{كَائِنٌ} *kā'in^{un}*, kA'inAtuN ^{كَائِنَاتٌ} *kā'in^{ātun}*, \uc{`amraNU} ^{عَمَرُوا} *Amr^{an}*, fataN_A ^{فَتَى} *fataⁿ*, qA.diNI ^{قَاضٍ} *qāḍiⁿ*.

ḡayr munṣarif: al-mu`allim-u ^{المُعَلِّمُ} *al-mu'allim^u*, kitAb-Ani ^{كِتَابَانِ} *kitāb^{āni}*, ra`sa'-Ani ^{رَاسَانِ} *raša'āni*, sAriq-Una ^{سَارِقُونَ} *sāriq^{āna}*, qA.d-Una ^{قَاضُونَ} *qāḍ^{āna}*, al-.zulm-Atu ^{الظُّلُمَاتُ} *aḡ-ḡulm^{ātu}*.

REM. a. As the *tanwīn* is passed over in pronunciation when it is followed by the letters ج, ل, ي, و, م (see item (b) on page 17), it may be desirable to further distinguish it by putting it above the line, but not to do the same for *ḡayr munṣarif* terminations. This can be achieved by simply omitting the hyphen before any *ḡayr munṣarif* termination:—

kAna .ganiyyaN l_akinna-hu labisa ḡubbaTaN mumazzaqaN 'aydu-hA ^{كَانَ غَنِيًّا لَكِنَّهُ لَيْسَ جَبَّةً مَزْمَرًا} *kāna ḡaniyy^{an} lākinna-hu labisa ḡubbat^{an} mumazzaq^{an} 'aydu-hā*.

REM. b. Although the hyphen before the *tanwīn* is optional as *arabluatex* always parses nouns with such termination, it may also be used to mark better the inflectional endings:—

mana`a 'l-nAs-a kAffaT-aN min mu_hA.tabati-hi 'a.had-uN bi-sayyidi-nA ^{مَنَّعَ النَّاسَ كُلَّهُ مِنْ أَحَدٍ بِسَيِّدٍ} *mana'a 'n-nās^a kāffat^{an} min muḡāṭabati-hi 'aḡad^{un} bi-sayyidi-nā*.

Discarding the ʾirāb As said above (item (c) on the previous page), the *ʾirāb* may be discarded in some cases, as in transliterated proper names or book titles. *arabluatex* is able to render words ending with *tā' marbūṭah* in different ways, depending on their function:—

- (a) Nouns followed by an adjective in apposition: *madInaT kabIraT madīnah kabīrah*, *al-madInaT al-kabIraT al-madīnah al-kabīrah*.

- (b) Nouns followed by another noun in the genitive (construct state): `.hikmaT al-l_ah hikmat Allāh, fi.d.daT al-darAhim fiḍḍat ad-darāhim`.

REM. It may so happen, as in the absence of the article before the annexed word, that `arablu-atex` be unable to determine which of the above two cases the word ending with *tā' marbūṭah* falls into. The 'pipe' character (see section 4.5 on page 24) may be appended to that word to indicate that what follows is in the construct state: `\uc{r}isAlaT fI tartIb qirA'aT| kutub \uc{g}AlInUs Riṣālah fī tartīb qirā'at kutub Ġālīnūs`.

Uncertain short vowels In some printed books, it may happen that more than one short vowel be placed on a consonant in cases where the vocalization is uncertain or ambiguous, like so: *فَعِلَ*. In transliteration, the uncertain vowels go between slashes and are separated by commas: *fa`uaila فَعِلَ fa'/u,a,i/la*.

8.2 Examples

Here follows in transliteration the story of Ġuḥā and his donkey (جُحَا وَحَمَارُهُ). See the code on page 8:—

‘dmg’ standard: *atā ṣadīq^{un} ilā Ġuḥā yaṭlubu min-hu ḥimāra-hu li-yarkaba-hu fī safratⁱⁿ qaṣīratⁱⁿ fa-qāla la-hu: “sawfa uṭdu-hu ilay-ka fī l-masā’i wa-adfa’u la-ka uḡrat^{an}.”* *fa-qāla Ġuḥā: “anā āsīf^{un} ḡidd^{an} annī lā astatī’u an uḥaqqīqa la-ka raḡbata-ka fa-l-ḥimār^u laysa huna l-yawm^a.”* *wa-qabla an yutimma Ġuḥā kalāma-hu bada’a l-ḥimār^u yanhaqu fī iṣṭabli-hi.* *fa-qāla la-hu ṣadīqu-hu: “innī asma’u ḥimāra-ka yā Ġuḥā yanhaqu.”* *fa-qāla la-hu Ġuḥā: “ḡarīb^{un} amru-ka yā ṣadīqī a-tuṣaddiqu l-ḥimār^a wa-tukaddība-nī?”*

‘loc’ standard: *atā ṣadīqun ilā Juḥā yaṭlubu min-hu ḥimāra-hu li-yarkaba-hu fī safratīn qaṣīratīn fa-qāla la-hu: “sawfa uṭdu-hu ilay-ka fī al-masā’i wa-adfa’u la-ka uḡratan.”* *fa-qāla Juḥā: “anā āsīfun jiddan annī lā astatī’u an uḥaqqīqa la-ka raḡbata-ka fa-al-ḥimāru laysa hunā al-yawm^a.”* *wa-qabla an yutimma Juḥā kalāma-hu bada’a al-ḥimāru yanhaqu fī iṣṭabli-hi.* *fa-qāla la-hu ṣadīqu-hu: “innī asma’u ḥimāra-ka yā Juḥā yanhaqu.”* *fa-qāla la-hu Juḥā: “ḡharībun amru-ka yā ṣadīqī a-tuṣaddiqu al-ḥimāra wa-tukadhdhiba-nī?”*

‘arabica’ standard: *atā ṣadīqun ilā Ġuḥā yaṭlubu min-hu ḥimāra-hu li-yarkaba-hu fī safratīn qaṣīratīn fa-qāla la-hu: “sawfa uṭdu-hu ilay-ka fī l-masā’i wa-adfa’u la-ka uḡratan.”* *fa-qāla Ġuḥā: “anā āsīfun ḡiddan annī lā astatī’u an uḥaqqīqa la-ka raḡbata-ka fa-l-ḥimāru laysa hunā l-yawm^a.”* *wa-qabla an yutimma Ġuḥā kalāma-hu bada’a l-ḥimāru yanhaqu fī iṣṭabli-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-tukaddība-nī?”*

9 Buckwalter input scheme

Even though `arabluatex` is primarily designed to process the `ArabTeX` notation, it

can also process the Buckwalter input scheme to a large extent.⁴⁴ The Buckwalter scheme is actually processed in two steps, as it is first converted into ArabTeX. Then, once this is accomplished, the ArabTeX scheme is processed through the above described functions. In this way, the Buckwalter input scheme can make the most of the `arabluatex` special features that are presented in section 2.2 on page 5.

`\SetInputScheme`

The input scheme, which is set to `arabtex` by default, may be changed at any point of the document by the `\SetInputScheme{<scheme>}` command, where `<scheme>` may be either `arabtex` or `buckwalter`. This command is also accepted in the preamble should one wish to set the input scheme globally, like so:—

```
1 \usepackage{arabluatex}
2 \SetInputScheme{buckwalter}
```

‘base’, ‘xml’ and ‘safe’ schemes `arabluatex` can use any of the so-called Buckwalter ‘base’, ‘xml’ or ‘safe’ schemes as they are described in Habash (2010, pp. 25–26).⁴⁵ However, the following limitation apply to the ‘base’ and ‘xml’ schemes: the braces { and }, which are used to encode ا and ؤ, must be replaced with square brackets viz. [and] respectively.

It is therefore recommended to use the Buckwalter ‘safe’ scheme.

Table 9 gives the Buckwalter equivalents that are currently used by `arabluatex`. The additional characters that are defined in table 6 on page 27 are also available.

| Letter | Transliteration ⁴⁶ | | | Buckwalter notation | |
|--------|-------------------------------|-----------|-----------|---------------------|------|
| | dmg | loc | arabica | base/xml | safe |
| ا | <i>a</i> | <i>a</i> | <i>a</i> | A | A |
| ب | <i>b</i> | <i>b</i> | <i>b</i> | b | b |
| ت | <i>t</i> | <i>t</i> | <i>t</i> | t | t |
| ث | <i>ṭ</i> | <i>th</i> | <i>ṭ</i> | v | v |
| ج | <i>ǧ</i> | <i>j</i> | <i>ǧ</i> | j | j |
| ح | <i>ḥ</i> | <i>ḥ</i> | <i>ḥ</i> | H | H |
| خ | <i>ḫ</i> | <i>kh</i> | <i>ḫ</i> | x | x |
| د | <i>d</i> | <i>d</i> | <i>d</i> | d | d |
| ذ | <i>ḏ</i> | <i>dh</i> | <i>ḏ</i> | * | V |
| ر | <i>r</i> | <i>r</i> | <i>r</i> | r | r |
| ز | <i>z</i> | <i>z</i> | <i>z</i> | Z | Z |
| س | <i>s</i> | <i>s</i> | <i>s</i> | s | s |
| ش | <i>š</i> | <i>sh</i> | <i>š</i> | \$ | c |
| ص | <i>ṣ</i> | <i>ṣ</i> | <i>ṣ</i> | S | S |

Table 9: Buckwalter scheme

⁴⁴See <http://www.qamus.org/transliteration.htm>

⁴⁵I am grateful to Graeme Andrews who suggested that the ‘safe’ scheme be included in `arabluatex`.

⁴⁶See section 8 on page 39.

| Letter | Transliteration | | | Buckwalter notation | |
|------------|-----------------|-----------|-----------|---------------------|------|
| | dmg | loc | arabica | base/xml | safe |
| ض | <i>d</i> | <i>d</i> | <i>d</i> | D | D |
| ط | <i>t</i> | <i>t</i> | <i>t</i> | T | T |
| ظ | <i>ẓ</i> | <i>ẓ</i> | <i>ẓ</i> | Z | Z |
| ع | | | | E | E |
| غ | <i>g</i> | <i>gh</i> | <i>g</i> | g | g |
| ف | <i>f</i> | <i>f</i> | <i>f</i> | f | f |
| ق | <i>q</i> | <i>q</i> | <i>q</i> | q | q |
| ك | <i>k</i> | <i>k</i> | <i>k</i> | k | k |
| ل | <i>l</i> | <i>l</i> | <i>l</i> | l | l |
| م | <i>m</i> | <i>m</i> | <i>m</i> | m | m |
| ن | <i>n</i> | <i>n</i> | <i>n</i> | n | n |
| ه | <i>h</i> | <i>h</i> | <i>h</i> | h | h |
| و | <i>w</i> | <i>w</i> | <i>w</i> | w | w |
| ي | <i>y</i> | <i>y</i> | <i>y</i> | y | y |
| ى | <i>ā</i> | <i>á</i> | <i>ā</i> | Y | Y |
| ة | <i>ah</i> | <i>ah</i> | <i>a</i> | p | p |
| ء | ’ | ’ | ’ | ’ | C |
| آ | <i>’ā</i> | <i>’ā</i> | <i>’ā</i> | | M |
| أ | ’ | ’ | ’ | > | O |
| ؤ | ’ | ’ | ’ | & | W |
| إ | ’ | ’ | ’ | < | I |
| ئ | ’ | ’ | ’ |] | Q |
| ـ | — | — | — | ~ | ~ |
| ـ | , | , | — | [| L |
| ـ | <i>a</i> | <i>a</i> | <i>a</i> | a | a |
| ـ | <i>u</i> | <i>u</i> | <i>u</i> | u | u |
| ـ | <i>i</i> | <i>i</i> | <i>i</i> | i | i |
| ـ | <i>an</i> | <i>an</i> | <i>an</i> | F | F |
| ـ | <i>un</i> | <i>un</i> | <i>un</i> | N | N |
| ـ | <i>in</i> | <i>in</i> | <i>in</i> | K | K |
| ـ | — | — | — | o | o |
| ـ | <i>ā</i> | <i>ā</i> | <i>ā</i> | ` | e |
| ـ (taṭwīl) | — | — | — | - | - |

Table 9: Buckwalter scheme

Transliteration The Buckwalter notation can also be transliterated into any accepted romanization standard of Arabic. See above section 8 on page 39 for more information. However, it should be pointed out again that only accurate coding produces accurate transliteration. It is therefore at the very least highly advisable to use the hyphen for tying the definite article and the inseparable particles (viz. prepositions, adverbs and conjunctions) to words, like so:—

Al-EaAlamu الْعَالَمُ *al-‘ālam^u*, Al-camsu الشَّمْسُ *aš-šams^u*, bi-SinaAEapi
 Al-T~ib~i, بِصِنَاعَةِ الطِّبِّ *bi-ṣinā‘atⁱ ‘t-ṭibbⁱ*.
 wa-Al-l~ehi وَاللَّهِ *wa-‘l-lāhⁱ*, Al-Hamdu li-l~ehi لِلَّهِ الْحَمْدُ *al-ḥamd^u li-llāhⁱ*.

Similarly, it is not advisable to use | and [(‘base’ and ‘xml’ schemes) or M and L (‘safe’ scheme) to encode the ‘*alif^u*’ *‘l-mamdūdātⁱ* and the ‘*alif^u*’ *‘l-waṣlⁱ* for such signs are supposed to be generated by *arabluatex* internal functions. Besides, as they do not *per se* convey any morphological information on what they are derived from, they cannot be transliterated accurately. To take one example, <ilY Al-LntiqāADi gives إِلَى الْإِنْتِقَاضِ as expected, but only <ilY Al-intiqADi can be transliterated as *‘ilā ‘l-intiqāḍi* with the correct vowel ⟨i⟩ in place of the ‘*alif^u*’ *‘l-waṣlⁱ*.

10 Unicode Arabic input

As said above in section 9 on page 44 about the Buckwalter input scheme, even though *arabluatex* is primarily designed to process the ArabTeX notation, it also accepts Unicode Arabic input. It should be noted that *arabluatex* does in no way interfere with Unicode Arabic input: none of the *voc*, *fullvoc*, *novoc* or *trans* options will have any effect on plain Unicode Arabic for the time being.

That said, there are two ways of inserting Unicode Arabic:

- `\txarb` (a) The `\txarb{Unicode Arabic}` command for inserting Unicode Arabic text in paragraphs;
- `txarab` (b) The `txarab` environment for inserting running paragraphs of Arabic text, like so:—

```
1 \begin{txarab}
2 <Unicode Arabic text>
3 \end{txarab}
```

11 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 any single-argument command with up to *two optional arguments*—that is: `\command[⟨opt1⟩][⟨opt2⟩]{⟨arg⟩}`—such as `\emph{⟨text⟩}`, `\textbf{⟨text⟩}` and the like, is assumed to have Arabic text in its mandatory argument:—

```
\abjad{45} kitAbu-hu \emph{fI 'l-\uc{`AdAt-i}} مَهْ كِتَابُهُ فِي الْعَادَاتِ
kitābu-hu fi 'l-Ādāt47
\arb{\abjad{45} \framebox[1in][s]{kitAbu-hu fI 'l-`AdAti}}
مَهْ كِتَابُهُ فِي الْعَادَاتِ
```

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-muqAtilaT-i\footnote{al-muqAtilaT-i:
5 al-muqAtil-Ina.}, wa-kAnat 'ummI min `u.zamA'-i buyUt-i
6 'l-zamAzimaT-i\footnote{al-zamAzimaT-u: .tA'ifaT-u mina
7 'l-furs-i.}.
8 \end{arab}
```

إِنَّ أَيْ كَانَ مِنَ الْمُقَاتِلَةِ^a، وَكَانَتْ أُمِّي مِنْ عُظَمَاءِ بُيُوتِ الزَّمَازِمَةِ^b.

^aالمُقَاتِلَةُ: الْمُقَاتِلِينَ.
^bالزَّمَازِمَةُ: طَائِفَةٌ مِنَ الْقُرْسِ.

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

\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 L^AT_EX 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{⟨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.

\LRfootnote `\LRfootnote{⟨text⟩}` and **\RLfootnote** `\RLfootnote{⟨text⟩}` typeset left-to-right and right-

⁴⁷This is odd in Arabic script, but using such features as `\emph` or `\textbf` is a matter of personal taste.

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` can be used to insert English footnotes in running Arabic text:—

```

1 \begin{arab}[fullvoc]
2   \uc{z}ayd-uN\arbnnull{ibnu}\LRfootnote{%
3     \enquote{\arb[trans]{\uc{z}ayd} is the son of
4       \arb[trans]{\uc{`a}mr}}: the second noun is not in
5       apposition to the first, but forms part of the
6       predicate\ldots} \arbnnull{zayduN}ibn-u \uc{`a}mr-iNU
7 \end{arab}

```

زَيْدٌ^a ابْنُ عَمْرٍو

^a “*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 `\FixArbFtnmk` command. However, for more control over the layout of footnotes marks, it is advisable to use the `scrextend` package.⁴⁸

`\LRmarginpar`

The `\LRmarginpar[<left>]{<right>}` command 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 the previous page applies.

`\setRL`

`\setLR`

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

```

1 \setRL
2 \section*{\arb{barzawayhi li-buzurjumihra bn-i 'l-buxtikAni}}
3 \setLR
4 \begin{arab}
5   qAla barzawayhi bn-u 'azhar-a, ra's-u 'a.tibbA'-i fAris-a...
6 \end{arab}

```

⁴⁸See <http://ctan.org/pkg/koma-script>; read the documentation of KOMA-script for details about the `\deffootnotemark` and `\deffootnote` commands.

بِرَزْوِيهِ لِبُزْجُمِهِرِ بْنِ الْبُخْتِكَانِ
قَالَ بِرَزْوِيهِ بْنُ أَزْهَرَ، رَأْسُ أَطِبَّاءِ فَارِسَ...

11.1 New commands

New feature
v1.9

In some particular cases, it may be useful to define new commands to be inserted in Arabic environments. From the general principle laid on page 48, it follows that any command that is found inside an Arabic environment is assumed to have Arabic text in its argument which `arabluatex` will process as such before passing it on to the command itself for any further processing. As a result of this feature, such a command as:

```
\newcommand{\fvarabic}[1]{\arb[fullvoc]{#1}}
```

will work as expected, but will always output non-vocalized Arabic if it is inserted in a `novoc` Arabic environment because its argument will have been processed by the `novoc` rules before the command `\fvarabic` itself can see it.

`\MkArbBreak`

The `\MkArbBreak{<csv list of commands>}` command can be used in the preamble to give any command—either new or already existing—the precedence over `arabluatex` inside Arabic environments. It takes as argument a comma-separated list of commands each of which must be stripped of its leading character `\`, like so:—

```
\MkArbBreak{onecmd, anothercmd, yetanothercmd, ...}
```

For example, here follows a way to define a new command `\fvred` to distinguish words with a different color and always print them in fully vocalized Arabic:—

```
1 \MkArbBreak {fvred}
2 \newcommand{\fvred}[1]{\arbcolor[red]{\arb[fullvoc]{#1}}}
3 \begin{arab}[voc]
4 _tumma "intalaqa_dU 'l-qarn-ayni 'il_A 'ummaT-iN 'u_hr_A fI
5 \fvred{((ma.tli`-i 'l-^sams-i))} wa-lA binA'-a la-hum
6 yu'amminu-hum mina 'l-^sams-i.
7 \end{arab}
```

ثُمَّ اتَّخَذَ ذُو الْقَرْنَيْنِ إِلَى أُمَّةٍ أُخْرَى فِي **مَطْلَعِ الشَّمْسِ** وَلَا بِنَاءَ لَهُمْ يُؤْمِنُهُم مِّنَ الشَّمْسِ.

It must be noted that the arguments, either optional or mandatory, of commands declared with `\MkArbBreak` are not to be processed by `arabluatex`. Therefore, as in the previous example, any of their argument to be rendered in Arabic must be inserted again in `\arb`. These commands themselves may have up to two optional and/or mandatory arguments followed by one optional argument, like so:—

- (a) `\command` (no argument, lowermost combination)
- (b) `\command[<optI>]` (one optional argument)

New feature
v1.12

New feature
v1.12

- (c) `\command{⟨arg1⟩}` (one mandatory argument)
- (d) `\command[⟨opt1⟩]{⟨arg1⟩}` (one optional and one mandatory argument)
- (e) [...]
 - (f) `\command[⟨opt1⟩][⟨opt2⟩]{⟨arg1⟩}{⟨arg2⟩}`
 - (g) `\command[⟨opt1⟩][⟨opt2⟩]{⟨arg1⟩}{⟨arg2⟩}[⟨opt3⟩]` (uppermost combination)

`\MkArbBreak*`

As said above, `\MkArbBreak` prevents `arabluatex` from processing the arguments of ‘declared’ commands as Arabic text. This technique proves sufficient in most cases. However, a ‘starred’ version of this command—`\MkArbBreak*{⟨csv list of commands⟩}`—is also provided. It goes a step further, as it directs `arabluatex` to *close* the current Arabic environment before any of the ‘declared’ commands, then *resume* it just after.

It must be noted that `\MkArbBreak*` must be used with the utmost care and *should never be used* if `\MkArbBreak` gives satisfaction. At any rate, the latter must always be tested before the former.

11.2 Environments

New feature
v1.5

Environments such as `\begin{quote}... \end{quote}` may be nested inside the `arab` environment. Up to one optional argument may be passed to each nested environment, like so:—

```
1 \begin{arab}
2   \begin{environment}[<options>]
3     <Arabic text>
4   \end{environment}
5 \end{arab}
```

In the following example, the quoting package is used:—

```
1 \setquotestyle {arabic}
2 \begin{arab}[fullvoc]
3   kAna \uc{'abU} \uc{'l-hu_dayli} 'ahd_A 'il_A \uc{muwaysiN}
4   dajAjaTaN. wa-kAnat dajAjatu-hu 'llatI 'ahdA-hA dUna mA kAna
5   yuttaxa_du li-\uc{muwaysiN}. wa-l_akinna-hu bi-karami-hi
6   wa-bi-.husni xuluqi-hi 'a.zhara 'l-ta`ajjuba min simani-hA
7   wa-.tIbi la.hmi-hA. wa-kAna <\uc{'abU} \uc{'l-hu_dayli}>
8   yu`rafu bi-'l-'imsAki 'l-^sadIdi. fa-qAla: \enquote{wa-kayfa
9     ra'ayta yA \uc{'abA} \uc{'imrAna} tilka 'l-dajAjaTa?} qAla:
10  \enquote{kAnat `ajabaN mina 'l-`ajabi!} fa-yaqUlu:
11  \begin{quoting}[begintext=\textquotedblright,
12    endtext=\textquotedblleft]
13    wa-tadrI mA jinsu-hA? wa-tadrI mA sinnu-hA? fa-'inna
14    'l-dajAjaTa 'inna-mA ta.tIbu bi-'l-jinsi wa-'l-sinni.
```

```

15      wa-tadrI bi-'ayyi ^say'iN kunna nusamminu-hA? wa-fI 'ayyi
16      makAniN kunna na`lifu-hA?
17      \end{quoting}
18      fa-lA yazAlu fI h_a_dA wa-'l-'A_haru ya.d.haku .da.hkaN
19      na`rifu-hu na.hnu wa-lA ya`rifu-hu \uc{'abU} \uc{'l-hu_dayli}.
20      \end{arab}

```

كَانَ أَبُو الْهَذِيلِ أَهْدَى إِلَى مُوسَى دَجَاجَةً. وَكَانَتْ دَجَاجَتُهُ الَّتِي أَهْدَاهَا دُونَ مَا كَانَ يَتَّخِذُ لِمُوسَى. وَلَكِنَّهُ بِكَرَمِهِ وَبِحُسْنِ خُلُقِهِ أَظْهَرَ التَّعَجُّبَ مِنْ سَمَنِهَا وَطِيبِ لَحْمِهَا. وَكَانَ <أَبُو الْهَذِيلِ> يَعْرِفُ بِالْإِمْسَاكِ الشَّدِيدِ. فَقَالَ: "وَكَيْفَ رَأَيْتَ يَا أَبَا عِمْرَانَ تِلْكَ الدَّجَاجَةَ؟" قَالَ: "كَانَتْ عَجَبًا مِنَ الْعَجَبِ!" فَيَقُولُ:

"وَتَدْرِي مَا جَنَسُهَا؟ وَتَدْرِي مَا سَنُهَا؟ فَإِنَّ الدَّجَاجَةَ إِذَا تَطَيَّبَ بِالْجَنَسِ وَالسِّنِّ. وَتَدْرِي بِأَيِّ شَيْءٍ كُنَّا نُسَمِّيُهَا؟ وَفِي أَيِّ مَكَانٍ كُنَّا نَعْلِفُهَا؟"

فَلَا يَزَالُ فِي هَذَا وَالْآخِرِ يَضْحَكُ ضَحْكَاً نَعْرِفُهُ نَحْنُ وَلَا يَعْرِفُهُ أَبُو الْهَذِيلِ.

11.2.1 Lists

Lists environments are also accepted inside the `arab` environment. One may either use any of the three standard list environments, viz. `itemize`, `enumerate` and `description` or use packages that provide additional refinements such as `paralist` or `enumitem`.

To take a first example, should one wish to typeset a list of manuscripts, the `description` environment can be used like so:—

```

1  \setRL \paragraph{\arb[novoc]{rumUzi 'l-kitAbi}}\setLR
2  \begin{arab}[novoc]
3  \begin{description}
4  \item[b] max.tU.tu 'l-maktabaTi 'l-'ahliyyaTi bi-\uc{bArIs} 2860
5  `arabiyyuN.
6  \item[s] max.tU.tu 'l-maktabaTi 'l-'ahliyyaTi bi-\uc{bArIs} 2859
7  `arabiyyuN.
8  \item[m] max.tU.tu majlisi \arb[novoc]{^sUrAY maly} .tahrAna
9  521.
9  \end{description}
10 \end{arab}

```

رموز الكتاب

ب مخطوط المكتبة الأهلية بباريس ٢٨٦٠ عربي.
 س مخطوط المكتبة الأهلية بباريس ٢٨٥٩ عربي.
 م مخطوط مجلس شوراى ملي طهران ٥٢١.

As a second example, the contents of a treatise may be typeset with the standard list environments, like so:—

```

1 \setRL \centerline{\arb{\textbf{al-qAnUnu fI 'l-.tibbi}}}\setLR
2 \begin{arab}
3 \begin{itemize}
4 \item \textbf{al-fannu 'l-'awwalu} fI .haddi 'l-.tibbi
5 wa-maw.dU`Ati-hi mina 'l-'umUri 'l-.tabI`iyyaTi wa-ya`stamilu
6 `al_A sittaTi ta`AlImiN
7 \begin{itemize}
8 \item \textbf{al-ta`lImu 'l-'awwalu} [wa-huwa fa.slAni]
9 \begin{itemize}
10 \item \textbf{al-fa.slu 'l-'awwalu}
11 \end{itemize}
12 \end{itemize}
13 \end{itemize}
14 \end{arab}

```

الْقَانُونُ فِي الطِّبِّ
 - الْفَنُّ الْأَوَّلُ فِي حَدِّ الطِّبِّ وَمَوْضُوعَاتِهِ مِنَ الْأُمُورِ الطَّبِيعِيَّةِ وَيَشْتَمِلُ عَلَى سِتَّةِ تَعَالِيمٍ
 - التَّعْلِيمُ الْأَوَّلُ [وَهُوَ فَصْلَانِ]
 - الْفَصْلُ الْأَوَّلُ

As a third example, abjad-numbered lists can be typeset in conjunction with the `enumitem` package,⁴⁹ like so:—

```

1 % preamble:---
2 \usepackage{enumitem}
3 \newlist{enumabjad}{enumerate}{10}
4 \setlist[enumabjad]{nosep, label={\abjad{\arabic*}}}
5 \usepackage{multicol}

```

```

1 From \textcite[i. 29 B--C]{Wright}:--- The derived forms of the
2 trilateral verb are usually reckoned fifteen in number, but the
3 learner may pass over the last four, because (with the exception
4 of the twelfth) they are of very rare occurrence.
5 \RLmulticolcolumns
6 \begin{multicols}{3}
7 \begin{arab}[fullvoc]
8 \begin{enumabjad}
9 \item fa`ala
10 \item fa``ala

```

⁴⁹See the documentation of `enumitem` for more details: <https://ctan.org/pkg/enumitem>

```

11      \item fA`ala
12      \item 'af`ala
13      \item tafa``ala
14      \item tafA`ala
15      \item infA`ala
16      \item ifta`ala
17      \item if`alla
18      \item istaf`ala
19      \item if`Alla
20      \item if`aw`ala
21      \item if`awwala
22      \item if`anlala
23      \item if`anl_A
24      \end{enumabjad}
25  \end{arab}
26 \end{multicols}

```

From Wright (1896, i. 29 B–C):— The derived forms of the trilateral verb are usually reckoned fifteen in number, but the learner may pass over the last four, because (with the exception of the twelfth) they are of very rare occurrence.

| | | |
|-------------|-----------------|------------------|
| أَفْعَلَّ | وَتَفَاعَلَّ | يَأْإِفْعَالَّ |
| بَفْعَلَّ | زَانْفَعَلَّ | يَبْإِفْعَوَلَّ |
| جَفَاعَلَّ | حَإِفْعَلَّ | يَجْإِفْعَوَلَّ |
| دَفْعَلَّ | طَإِفْعَلَّ | يَدْإِفْعَلَّلَّ |
| هَتَفْعَلَّ | يَإِسْتَفْعَلَّ | يَهْإِفْعَنَلَّ |

Caveat The various French definition files of the babel package viz. `acadian`, `canadien`, `français`, `frenchb` or `french` all redefine the list environments, which breaks the standard definition file that is used by `arabluatex`. Therefore, `babel-french` must be loaded with the `StandardLists=true` option, like so:—

```

1 \usepackage[french]{babel}
2 \frenchsetup{StandardLists=true}

```

This option will prevent `babel-french` from interfering with the layout of the document. Then the `paralist` or `enumitem` packages can be used to make the lists ‘compact’ as `babel-french` do.

11.3 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 {\textquotedblright}{\textquotedblleft}
4 {\textquoteright}{\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{.garIb-uN 'amru-ka yA .sadIqI
4   'a-tu.saddiqu 'l-.himAr-a wa-tuka_d_diba-nI?}
5 \end{arab}
6 \setquotestyle{english}
```

فَقَالَ لَهُ جُحَا: "غَرِيبٌ أَمْرُكَ يَا صَدِيقِي أَتَصَدِّقُ الْحِمَارَ وَتَكْفُرُ بِي؟"

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.

11.4 Two-argument special commands

textcolor The two-argument command `\textcolor{<color>}{<Arabic text>}` is supported inside `\begin{arab} ... \end{arab}`. One simple example follows:⁵⁰—

```
1 \begin{arab}
2   \textcolor{red}{\uc{m}uha_d_dabu \uc{'l-d}Ini \uc{'a}bdu
3   \uc{'l-r}a.hImi bnu \uc{'a}liyyiN} huwa ^say_hu-nA 'l-'imAmu
4   'l-.sadru 'l-kabIru 'l-`Alimu 'l-fA.dilu \uc{m}uha_d_dabu
5   \uc{'l-d}Ini \uc{'a}bU \uc{m}u.hammadIN \uc{'a}bdu
6   \uc{'l-r}a.hImi bnu \uc{'a}liyyi bni \uc{.h}AmidiN wa-yu`rafu
7   bi-\uc{'l-d}a_hwari.
8 \end{arab}
9 \begin{arab}[trans]
10  \textcolor{red}{\uc{m}uha_d_dabu \uc{'l-d}Ini \uc{'a}bdu
11  \uc{'l-r}a.hImi bnu \uc{'a}liyyiN} huwa ^say_hu-nA 'l-'imAmu
12  'l-.sadru 'l-kabIru 'l-`Alimu 'l-fA.dilu \uc{m}uha_d_dabu
13  \uc{'l-d}Ini \uc{'a}bU \uc{m}u.hammadIN \uc{'a}bdu
14  \uc{'l-r}a.hImi bnu \uc{'a}liyyi bni \uc{.h}AmidiN wa-yu`rafu
```

⁵⁰`arabluatex` provides its own `\arbcOLOR` command which is able to render syllables or diacritics in colors. See section 7 on page 35.

```

15 bi-\uc{'l-d}a_hwari.
16 \end{arab}

```

مُهَدَّبُ الدِّينِ عَبْدُ الرَّحِيمِ بْنِ عَلِيٍّ هُوَ شَيْخُنَا الْإِمَامُ الصَّدْرُ الْكَبِيرُ الْعَالِمُ الْفَاضِلُ مُهَدَّبُ الدِّينِ أَبُو مُحَمَّدٍ عَبْدُ
الرَّحِيمِ بْنِ عَلِيٍّ بْنِ حَامِدٍ وَيُعرفُ بِالْأَخْوَ.

*Muhaddabu 'd-Dīni 'Abdu 'r-Raḥīmi bnu 'Aliyyiⁱⁿ huwa šayḥu-na 'l-
imāmu 'ṣ-ṣadru 'l-kabīru 'l-ālimu 'l-fādilu Muhaddabu 'd-Dīni 'Abū
Muḥammadⁱⁿ 'Abdu 'r-Raḥīmi bnu 'Aliyyi bni Ḥāmidⁱⁿ wa-yu'rafu bi-'d-
Daḥwari.*

reledmac The two-argument command `\edtext{<lemma>}{<commands>}` is supported inside `\begin{arab} ... \end{arab}`.⁵¹ 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-jild-i
5 \edtext{\arb{.sadId-uN}}{\Afootnote{M: \arb{.sadId-aN} E1}}
6 \end{arab}
7 \pend
8 \endnumbering

```

11.5 quran

`arabluatex` is compatible with the `quran` package so that both can be used in conjunction with one another for typesetting the *Qur'ān*. As `quran` draws the text of the *Qur'ān* from a Unicode encoded database, its commands have to be passed as arguments to the `\txarb` command for short insertions in left-to-right paragraphs, or inserted inside the `txarab` environment for typesetting running paragraphs of *Qur'ānic* text (see above section 10 on page 47 for more details). Please note that `arabluatex` takes care of formatting the Arabic: therefore, it is recommended to load the `quran` package with the `nopar` option, after `arabluatex` itself has been loaded, like so:—

```

1 \usepackage{arabluatex}
2 \usepackage[nopar]{quran}

```

As an example, the following code will typeset the *sūrat al-Fātiḥah*:—

⁵¹`\pstart` and `\pend` are also supported inside the `arab` environment.


```

1 \begin{txarab}
2   \quransurah[1]
3 \end{txarab}

```

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ ﴿١﴾ الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ ﴿٢﴾ الرَّحْمَنِ الرَّحِيمِ ﴿٣﴾ مَالِكِ يَوْمِ الدِّينِ ﴿٤﴾
إِيَّاكَ نَعْبُدُ وَإِيَّاكَ نَسْتَعِينُ ﴿٥﴾ اهْدِنَا الصِّرَاطَ الْمُسْتَقِيمَ ﴿٦﴾ صِرَاطَ الَّذِينَ أَنْعَمْتَ عَلَيْهِمْ غَيْرِ الْمَغْضُوبِ
عَلَيْهِمْ وَلَا الضَّالِّينَ ﴿٧﴾

12 Future work

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

- (a) Short-term:
 - i. 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 4.
- (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 4.9 on page 27.

13 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 [2018/06/24 v1.12 An ArabTeX-like interface for LuaLaTeX]
4 \RequirePackage{ifluatex}

```

`arabluatex` requires `LuaLaTeX` 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 \RequirePackage{xkeyval}
13 \DeclareOptionX{voc}{\def\al@mode{voc}}
14 \DeclareOptionX{fullvoc}{\def\al@mode{fullvoc}}
15 \DeclareOptionX{novoc}{\def\al@mode{novoc}}
16 \DeclareOptionX{trans}{\def\al@mode{trans}}
17 \ExecuteOptionsX{voc}
18 \ProcessOptionsX\relax
19 \def\al@mode@voc{voc}
20 \def\al@mode@fullvoc{fullvoc}
21 \def\al@mode@novoc{novoc}
22 \def\al@mode@trans{trans}

```

Packages that are required by arabluatex:

```

23 \RequirePackage{xcolor}
24 \RequirePackage{luacolor}
25 \RequirePackage{etoolbox}
26 \RequirePackage{arabluatex-patch}
27 \RequirePackage{fontspec}
28 \RequirePackage{amsmath}
29 \RequirePackage{luacode}
30 \RequirePackage{xparse}
31 \RequirePackage{environ}
32 \RequirePackage{adjustbox}

```

The following boolean will be set to true in RL mode:

```

33 \newbool{al@rlmode}

```

Here begins the real work: load arabluatex.lua:

```

34 \luadirect{dofile(kpse.find_file("arabluatex.lua"))}

```

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

```

35 \AtBeginDocument{\ifdefined\arabicfont\relax\else
36 \PackageWarning{arabluatex}{\string\arabicfont\ is not defined.^^JI
37 will try to load Amiri}%
38 \newfontfamily\arabicfont[Script=Arabic]{Amiri}\fi}%

```

`\setRL` This neutralizes what may be defined by other packages:

```

39 \AtBeginDocument{\def\setRL{\booltrue{al@rlmode}\pardir TRT\textdir TRT}}

```

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

```

40 \AtBeginDocument{\def\setLR{\boolfalse{al@rlmode}\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:

```

41 \AtBeginDocument{\ifdef{\LR}%
42 {\RenewDocumentCommand{\LR}{m}{\bgroup\textdir TLT\rmfamily#1\egroup}}
43 {\NewDocumentCommand{\LR}{m}{\bgroup\textdir TLT\rmfamily#1\egroup}}}

```

| | |
|------------------------------|---|
| <code>\RL</code> | This one typesets its argument from right to left. Same remark as above regarding the need of redefinition. |
| | <pre> 44 \AtBeginDocument{\ifdef{\RL}% 45 {\RenewDocumentCommand{\RL}{m}{\bgroup\textdir TRT\rmfamily#1\egroup}} 46 {\NewDocumentCommand{\RL}{m}{\bgroup\textdir TRT#1\rmfamily\egroup}}}</pre> |
| <code>\MkArbBreak</code> | The <code>\MkArbBreak{<csv list of commands>}</code> command can be used to give any command—either new or already existing—the precedence over <code>arabluatex</code> inside Arabic environments. It is actually coded in Lua. |
| <code>\MkArbBreak*</code> | <code>\MkArbBreak*</code> goes a step further as it directs <code>arabluatex</code> to close the current Arabic environment before processing any ‘declared’ command then resume it just after. |
| | <pre> 47 \NewDocumentCommand{\MkArbBreak}{s m}{% 48 \IfBooleanTF{#1}{% 49 {\luadirect{mkarbbreak(\luastringN{#2}, "out")}} 50 {\luadirect{mkarbbreak(\luastringN{#2}, "dflt")}} 51 } }</pre> |
| <code>\aemph</code> | Arabic emphasis. Needs to be redefined as well. The function is actually coded in Lua. |
| <code>\aemph*</code> | The ‘starred’ version of this command always puts the stroke over its argument. |
| | <pre> 52 \AtBeginDocument{\ifdef{\aemph}% 53 {\RenewDocumentCommand{\aemph}{s m}{% 54 \IfBooleanTF{#1}{% 55 \luadirect{tex.sprint(aemph(\luastringN{#2}, "over"))}} 56 {\luadirect{tex.sprint(aemph(\luastringN{#2}, "dflt"))}}} 57 {\NewDocumentCommand{\aemph}{s m}{% 58 \IfBooleanTF{#1}{% 59 \luadirect{tex.sprint(aemph(\luastringN{#2}, "over"))}} 60 {\luadirect{tex.sprint(aemph(\luastringN{#2}, "dflt"))}}}}}</pre> |
| <code>\arbcolor</code> | <code>\arbcolor[<color>]{<Arabic text>}</code> takes the Arabic text to be colored as argument. |
| | <pre> 61 \NewDocumentCommand{\arbcolor}{o m}{% 62 \IfNoValueTF{#1}{#2}{\textcolor{#1}{#2}}}</pre> |
| <code>\SetInputScheme</code> | <code>arabluatex</code> is designed for processing Arab _T _E _X input notation. <code>\SetInputScheme</code> may be used in the preamble or at any point of the document should the user wish to use a different notation such as the ‘Buckwalter scheme’. |
| | <pre> 63 \def\al@input@scheme{arabtex} 64 \NewDocumentCommand{\SetInputScheme}{m}{\def\al@input@scheme{#1}}</pre> |
| <code>\SetArbEasy</code> | By default, <code>arabluatex</code> applies complex rules to generate euphonic <i>tašdīd</i> , <i>ʿalif</i> |
| <code>\SetArbEasy*</code> | <i>mamdūdah</i> and <i>sukūn</i> depending on the modes which are selected, either <i>voc</i> , |
| <code>\SetArbDflt</code> | <i>fullvoc</i> or <i>trans</i> . Such refinements can be discarded with <code>\SetArbEasy</code> , either globally in the preamble or at any point of the document. Note that <code>\SetArbEasy</code> keeps the <i>sukūn</i> that is generated, while the starred version <code>\SetArbEasy*</code> takes it away. Default complex rules can be set back at any point of the document with <code>\SetArbDflt</code> . |

`\SetArbDflt*` As of v1.6, arabuatex does not apply any more the assimilation rules that are laid on item (b) on page 17; a new starred version `\SetArbDflt*` is now available to the user should he wish to apply them.

```
65 \def\al@arb@rules{dflt}
66 \NewDocumentCommand{\SetArbEasy}{s}{%
67   \IfBooleanTF{#1}
68   {\def\al@arb@rules{easynosukun}}
69   {\def\al@arb@rules{easy}}}
70 \NewDocumentCommand{\SetArbDflt}{s}{%
71   \IfBooleanTF{#1}
72   {\def\al@arb@rules{idgham}}
73   {\def\al@arb@rules{dflt}}}
```

`\SetTranslitFont` By default, the font that is used for transliterated text is the main font of the document. Any other font may also be selected with the font-selecting commands of the fontspec package.

```
74 \def\al@trans@font{\rmfamily}%
75 \NewDocumentCommand{\SetTranslitFont}{m}{\def\al@trans@font{#1}}
```

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

```
76 \def\al@trans@style{\itshape}%
77 \NewDocumentCommand{\SetTranslitStyle}{m}{\def\al@trans@style{#1}}
```

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

```
78 \def\al@trans@convention{dmg}
79 \NewDocumentCommand{\SetTranslitConvention}{m}{\def\al@trans@convention{#1}}
```

`\arbup` By default, `\arbup` is set to `\textsuperscript`. This is how the *tanwīn* that takes place at the end of a word should be displayed in `dmg` mode. `\NoArbUp` may be used either in the preamble or at any point of the document in case one wishes to have the *tanwīn* on the line. The default rule can be set back with `\ArbUpDflt` at any point of the document. Finally `\SetArbUp` can be used to customize the way *tanwīn* is displayed: this command takes the formatting directives as argument, like so: `\SetArbUp{<code>}`.

```
80 \NewDocumentCommand{\al@arbup@dflt}{m}{\textsuperscript{\thinspace#1}}%
81 \NewDocumentCommand{\al@arbup}{m}{\al@arbup@dflt{#1}}
82 \NewDocumentCommand{\arbup}{m}{\al@arbup{#1}}
83 \NewDocumentCommand{\ArbUpDflt}{}{\let\al@arbup=\al@arbup@dflt}
84 \NewDocumentCommand{\NoArbUp}{}{\RenewDocumentCommand{\al@arbup}{m}{##1}}
85 \NewDocumentCommand{\SetArbUp}{m}{\RenewDocumentCommand{\al@arbup}{m}{#1}}
```

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

```
86 \NewDocumentCommand{\uc}{m}%
87   {\luairect{tex.sprint(uc(\luastringN{#1}))}}
```

`\Uc` `\uc` can be used safely in all of the modes that are provided by `arabluatex` as any of the `voc`, `fullvoc` and `novoc` modes discard it on top of any other functions to be run. `\Uc` does the same as `\uc` except that *it is never discarded*. For that reason, `\Uc` *should never be used outside the `trans` mode*. `arabluatex` uses `\Uc` internally so as to prevent `\uc` from being discarded in case words that are to be transliterated are inserted into Arabic commands or environments where transliteration is not required. Therefore, it is not documented.

```
88 \let\Uc\uc
```

`\prname` `\prname` is to be used outside Arabic environments for proper names. It takes as argument one or more Arabic words, each of which will be rendered in upright roman style with its first letter uppercased.

```
89 \NewDocumentCommand{\prname}{m}{%
90   \bgroup\SetTranslitStyle{\relax}\arb[trans]{\uc{#1}}\egroup}
```

`\txarb` `\txarb` sets the direction to right-to-left and selects the Arabic font. It is used internally by several Lua functions, 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. Therefore, it is not documented.

```
91 \NewDocumentCommand{\txarb}{+m}{\bgroup\textdir
92   TRT\arabicfont#1\egroup}
93 \NewDocumentCommand{\txtrans}{+m}{\bgroup\textdir
94   TLT\al@trans@font\al@trans@style#1\egroup}
```

`txarab` The `txarab` environment does for paragraphs the same as `\txarb` does for short insertions of utf8 Arabic text.

```
95 \NewDocumentEnvironment{txarab}{}{%
96   \par%
97   \booltrue{al@rlmode}%
98   \pardir TRT\textdir TRT\arabicfont}\par}
```

`\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.

```
99 \NewDocumentCommand{\arb}{0{\al@mode} +m}%
100 {\edef\@tempa{#1}%
101   \ifx\@tempa\al@mode@voc%
102     \booltrue{al@rlmode}%
103     \bgroup\textdir TRT\arabicfont%
104     \luadirect{tex.sprint(processvoc(\luastringN{#2},
105       \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}))}\egroup%
106     \boolfalse{al@rlmode}%
107     \else%
108     \ifx\@tempa\al@mode@fullvoc%
109       \booltrue{al@rlmode}%
110       \bgroup\textdir TRT\arabicfont%
111       \luadirect{tex.sprint(processfullvoc(\luastringN{#2},
```

```

112 \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}}))\egroup%
113 \boolfalse{al@rlmode}%
114 \else%
115 \ifx\@tempa\al@mode@novoc%
116 \booltrue{al@rlmode}%
117 \bgroup\textdir TRT\arabicfont%
118 \luadirect{tex.sprint(processnovoc(\luastringN{#2},
119 \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}}))\egroup%
120 \boolfalse{al@rlmode}%
121 \else%
122 \ifx\@tempa\al@mode@trans%
123 \bgroup\textdir TLT\al@trans@style%
124 \luadirect{tex.sprint(processstrans(\luastringN{#2},
125 \luastring0{\al@trans@convention},
126 \luastring0{\al@arb@rules},
127 \luastring0{\al@input@scheme}}))\egroup%
128 \else%
129 \fi\fi\fi\fi}

```

\arbmark **\arbmark** takes one argument from a list of defined elements. This command is coded in Lua.

```

130 \NewDocumentCommand{\arbmark}{m}{%
131 \bgroup%
132 \SetInputScheme{arabtex}%
133 \luadirect{tex.sprint(processarbmarks(\luastringN{#1}))}%
134 \egroup}

```

\newarbmark **\newarbmark** lets the user define additional Arabic marks. As **\arbmark**, this command is coded in Lua. It takes three arguments: the abbreviated form to be used as argument of **\arbmark**, the rendition in Arabic script and the rendition in romanized Arabic.

```

135 \NewDocumentCommand{\newarbmark}{m m m}{%
136 \luadirect{newarbmark(\luastringN{#1}, \luastringN{#2}, \luastringN{#3})}}

```

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

```

137 \NewEnviron{arab}[1][\al@mode]%
138 {\par\edef\@tempa{#1}%
139 \ifx\@tempa\al@mode@voc%
140 \booltrue{al@rlmode}%
141 \bgroup\pardir TRT\textdir TRT\arabicfont%
142 \luadirect{tex.sprint(processvoc(\luastring0{\BODY},
143 \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}}))\egroup%
144 \else%
145 \ifx\@tempa\al@mode@fullvoc%
146 \booltrue{al@rlmode}%
147 \bgroup\pardir TRT\textdir TRT\arabicfont%
148 \luadirect{tex.sprint(processfullvoc(\luastring0{\BODY},
149 \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}}))\egroup%

```

```

150 \else%
151 \ifx\@tempa\al@mode@novoc%
152 \booltrue{al@rlmode}%
153 \bgroup\pdir TRT\textdir TRT\arabicfont%
154 \luadirect{tex.sprint(processnovoc(\luastring0{\BODY},
155 \luastring0{\al@arb@rules}, \luastring0{\al@input@scheme}))}\egroup%
156 \else%
157 \ifx\@tempa\al@mode@trans%
158 \bgroup\pdir TLT\textdir TLT\al@trans@style%
159 \luadirect{tex.sprint(processtrans(\luastring0{\BODY},
160 \luastring0{\al@trans@convention},
161 \luastring0{\al@arb@rules},
162 \luastring0{\al@input@scheme}))}\egroup%
163 \else \fi\fi\fi\fi}\par

```

arabverse The arabverse environment may receive different options: mode, width, gutter, metre, utf and delim; all of them are defined here just before the arabverse environment:

```

164 \newlength{\al@bayt@width}
165 \newlength{\al@gutter@width}
166 \setlength{\al@bayt@width}{.3\textwidth}
167 \setlength{\al@gutter@width}{.15\al@bayt@width}
168 \define@key{al}{verse}{width}{\setlength{\al@bayt@width}{#1}}
169 \define@key{al}{verse}{gutter}{\setlength{\al@gutter@width}{#1}}
170 \define@key{al}{verse}{metre}{\arb{#1}}
171 \define@boolkey{al}{verse}{utf}[true]{}
172 \define@boolkey{al}{verse}{delim}[true]{}
173 \define@choicekey{al}{verse}{mode}{fullvoc, voc, novoc,
174 trans}{\def\al@mode{#1}}
175 \presetkeys{al}{verse}{metre={}, utf=false,
176 delim=false}{}

```

Then follows the environment itself:

```

177 \NewDocumentEnvironment{arabverse}{0{}}%
178 {\par\centering\noindent\bgroup\setkeys{al}{verse}[metre]{#1}%
179 \ifx\al@mode\al@mode@trans%
180 \ifal@verse@utf\setRL\else\setLR\fi%
181 \else\setRL\fi}%
182 {\hfill\setkeys{al}{verse}[width,gutter,utf,mode]{#1}\egroup}

```

\bayt Each verse consists of two hemistichs; therefore the `\bayt` command takes two arguments, the first receives the *ṣadr* and the second the *ʿağuz*. That two subsequent hemistichs should be connected with one another is technically named *tadwīr*. In some of these cases, the hemistichs may be connected by a prominent horizontal flexible stroke which is drawn by the `\al@verse@stroke` command.

\SetHemistichDelim A hemistich delimiter also may be defined. By default, it is set to the ‘star’ character: `*`. The `\SetHemistichDelim{<delimiter>}` command can be used at any point of the document to change this default setting.

```

183 \NewDocumentCommand{\arb@utf}{m}{%
184   \ifal@verse@utf\txarb{#1}\else\arb{#1}\fi}
185 \def\al@hemistich@delim{*}
186 \NewDocumentCommand{\SetHemistichDelim}{m}{\def\al@hemistich@delim{#1}}
187 \def\al@verse@stroke{\leavevmode\xleaders\hbox{\arb{--}}\hfill\kern0pt}
188 \NewDocumentCommand{\bayt}{m o m}{%
189   \ifdefined\savenotes\savenotes\else\fi%
190   \edef\al@tatweel{--}%
191   \adjustbox{width=\al@bayt@width, height=\Height}{\arb@utf{#1}}%
192   \IfNoValueTF{#2}{%
193     \ifal@verse@delim\makebox[\al@gutter@width][c]{\al@hemistich@delim}%
194     \else%
195       \hspace{\al@gutter@width}%
196     \fi
197   }{%
198     \edef\@tempa{#2}%
199     \ifx\@tempa\al@tatweel%
200       \ifx\al@mode\al@mode@trans%
201         \hspace{\al@gutter@width}%
202       \else%
203         \makebox[\al@gutter@width][s]{\al@verse@stroke}%
204       \fi%
205     \else%
206       \ifx\al@mode\al@mode@trans%
207         \adjustbox{width=\al@gutter@width, height=\Height}{\arb@utf{#2}}%
208       \else%
209         \makebox[\al@gutter@width][s]{\arb@utf{#2}}%
210       \fi\fi}%
211   \adjustbox{width=\al@bayt@width, height=\Height}{\arb@utf{#3}}%
212   \ifdefined\spewnotes\spewnotes\else\fi%
213 }

```

`\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.

```

214 \AtBeginDocument{%
215   \ifdefined\abjad%
216     \RenewDocumentCommand{\abjad}{m}%
217     {\ifbool{al@rlmode}%
218       {\aemph{\luadirect{tex.sprint(abjadify(\luastring{#1}))}}}
219       {\luadirect{tex.sprint(abjadify(\luastring{#1}))}}}
220     \else%
221     \NewDocumentCommand{\abjad}{m}%
222     {\ifbool{al@rlmode}%
223       {\aemph{\luadirect{tex.sprint(abjadify(\luastring{#1}))}}}
224       {\luadirect{tex.sprint(abjadify(\luastring{#1}))}}}
225     \fi}

```


`\arbnul1` The `\arbnul1` command does nothing by itself. It is processed only if it is found in Arabic context so as to put back on contextual analysis in case it has been broken by other commands.

```
226 \NewDocumentCommand{\arbnul1}{m}{\relax}
```

`\abraces` `\abraces{⟨Arabic text⟩}` puts its argument between braces. This macro is written in Lua and is dependent on the current value of `tex.textdir`.

```
227 \NewDocumentCommand{\abraces}{+m}{%
228   \luadirect{tex.sprint(abraces(\luastringN{#1}))}}
```

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

```
229 \DeclareDocumentCommand{\LRmarginpar}{o m}{%
230   \IfNoValueTF{#1}
231     {\marginpar{\textdir TLT #2}}
232     {\marginpar[\textdir TLT #1]{\textdir TLT #2}}}
```

`\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.

```
233 \DeclareDocumentCommand{\LRfootnote}{m}{\bgroup\pardir
234   TLT\textdir TLT\footnote{#1}\egroup}
235 \DeclareDocumentCommand{\RLfootnote}{m}{\bgroup\pardir
236   TRT\textdir TRT\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.

```
237 \NewDocumentCommand{\FixArbFtnmk}{}{%
238   \@ifpackageloaded{scrextend}%
239   {\AtBeginDocument{\deffootnote{2em}{1.6em}{\LR{\thefootnotemark}.\enskip}}}%
240   {\RequirePackage{scrextend}
241    \AtBeginDocument{\deffootnote{2em}{1.6em}{\LR{\thefootnotemark}.\enskip}}}
```

Errors and Warnings

```
242 \newcommand{\al@warning}[1]{\PackageWarning{arabluatex}{#1}}
243 \newcommand{\al@error}[2]{\PackageError{arabluatex}{#1}{#2}}
244 \newcommand{\al@wrong@nesting}{\al@error{%
245   (RL/LR)\string\footnote\space is not allowed\MessageBreak inside
246   \string\RL{} and \string\RL{} commands}{%
247   Get rid of the surrounding \string\RL{} or \string\LR{} command.}}
248 \newcommand{\al@wrong@mark}{\al@warning{%
249   Unknown Arabic mark in \string\arbmark{}. Replaced
250   with\MessageBreak <??>. Please check your code}}
```

That is it. Say goodbye before leaving.

Patches

```
251 \NeedsTeXFormat{LaTeX2e}
252 \ProvidesPackage{arabluatex-patch}%
253 [2016/11/14 v1.0 patches for arabluatex]
```

I have put in a separate .sty file external lines of code that I had to patch for a good reason. I hate doing this, and hopefully, most of these lines will disappear as soon as they are not required anymore.

The following is taken from latex.ltx. I had to make this patch for I could not find a way to process the list environments in right-to-left mode. The LuaTeX primitives \bodydir and \pagedir will eventually allow us to get rid of this:

```
254 \def\list#1#2{%
255   \ifnum \@listdepth >5\relax
256     \@toodeep
257   \else
258     \global\advance\@listdepth\@ne
259   \fi
260   \rightmargin\z@
261   \listparindent\z@
262   \itemindent\z@
263   \csname @list\romannumeral\the\@listdepth\endcsname
264   \def\@itemlabel{#1}%
265   \let\makelabel\@mklab
266   \@nmbrrlistfalse
267   #2\relax
268   \@trivlist
269   \parskip\parsep
270   \parindent\listparindent
271   \advance\linewidth -\rightmargin
272   \advance\linewidth -\leftmargin

patch begins:
273   \ifbool{al@rlmode}{\advance\@totalleftmargin \rightmargin}%
274   {\advance\@totalleftmargin \leftmargin}

patch ends.
275   \parshape \@ne \@totalleftmargin \linewidth
276   \ignorespaces}
277 \def\@item[#1]{%
278   \if@nparitem
279     \@donoparitem
280   \else
281     \if@inlabel
282       \indent \par
283     \fi
284     \ifhmode
285       \unskip\unskip \par
286     \fi
287     \if@newlist
288       \if@nobreak
```

```

289     \@nbitem
290     \else
291     \addpenalty\@beginparpenalty
292     \addvspace\@topsep
293     \addvspace{-\parskip}%
294     \fi
295     \else
296     \addpenalty\@itempenalty
297     \addvspace\itemsep
298     \fi
299     \global\@inlabeltrue
300 \fi
301 \everypar{%
302     \@minipagefalse
303     \global\@newlistfalse
304     \if@inlabel
305     \global\@inlabelfalse
306     {\setbox\z@\lastbox
307     \ifvoid\z@
308     \kern-\itemindent
309     \fi}%
310     \box\@labels
311     \penalty\z@
312 \fi
313 \if@nobreak
314     \@nobreakfalse
315     \clubpenalty \@M
316 \else
317     \clubpenalty \@clubpenalty
318     \everypar{}%
319 \fi}%
320 \if@noitemarg
321     \@noitemargfalse
322     \if@nmbolist
323     \refstepcounter\@listctr
324     \fi
325 \fi

patch begins:
326 \ifbool{al@rlmode}{\sRLbox\@tempboxa{\makelabel{#1}}}{%
327 \sbox\@tempboxa{\makelabel{#1}}}%
328 \ifbool{al@rlmode}{\global\setbox\@labels\hbox dir TRT}%
329 {\global\setbox\@labels\hbox}{%

patch ends.
330 \unhbox\@labels
331 \hskip \itemindent
332 \hskip -\labelwidth
333 \hskip -\labelsep
334 \ifdim \wd\@tempboxa >\labelwidth
335     \box\@tempboxa

```

```

336 \else
337 \hbox to\labelwidth {\unhbox\@tempboxa}%
338 \fi
339 \hskip \labelsep}%
340 \ignorespaces}

```

This is adapted from Vafa Khalighi's bidi package. Thanks to him.

```

341 \long\def\sRLbox#1#2{\setbox#1\hbox dir TRT{%
342 \color@setgroup#2\color@endgroup}}

```

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 vorgelegt von der Transkriptionkommission der Deutschen Morgenländischen Gesellschaft*. In collab. with Ph. S. van Ronkel and Otto Spies. Leipzig: Deutsche Morgenländische Gesellschaft, in Kommission bei F. A. Brockhaus. URL: http://www.naher-osten.uni-muenchen.de/studium_lehre/werkzeugkasten/dmgtransliteration.pdf.
- Habash, Nizar Y. (2010). *Introduction to Arabic Natural Language Processing*. Synthesis Lectures on Human Language Technologies 10. Toronto: Morgan & Claypool Publishers.
- Hosny, Khaled (Dec. 13, 2017). *Amiri*. URL: <http://www.amirifont.org/>.
- DIN 31 635 (July 2011). *Information and Documentation - Romanization of the Arabic Alphabet for Arabic, Ottoman-Turkish, Persian, Kurdish, Urdu and Pushto*. URL: <http://www.din.de>.
- Lagally, Klaus (Nov. 3, 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. | v1.1. |
| General: Initial release | 1 |
| v1.0.1. | |
| General: Minor update of the documentation | 1 |
| | \abjad: New and more flexible \abjad command. 64 |

| | | | | |
|---------|--|----|---|----|
| v1.2. | \SetArbEasy: New \SetArbEasy/\SetArbDflt for ‘modern’ or ‘classic’ Arabic styles. | 59 | \SetHemistichDelim: New \SetHemistichDelim command for changing the default delimiter between hemistichs . . | 63 |
| v1.3. | \arbup: ‘ <i>rāb</i> ’ is now written as superscript text in dmg mode by default. | 60 | v1.7. \arbnnull: New \arbnnull command for putting back on any contextual analysis rule broken by other commands. | 64 |
| v1.4. | \SetInputScheme: \SetInputScheme can be used to process other input schemes such as ‘Buckwalter’ | 59 | v1.8. General: arabica transliteration standard is now supported . . . | 40 |
| | \SetTranslitFont: For selecting a specific font for transliterated texts | 60 | v1.8.5. General: Six additional Persian characters are now available . . | 11 |
| v1.4.3. | \abraces: New \abraces command which expresses its argument between braces. | 65 | v1.9. \MkArbBreak: New \MkArbBreak command for inserting user-defined macros in Arabic environments | 59 |
| v1.4.4. | \SetArbEasy*: this starred version discards the <i>sukūn</i> in addition to what is already discarded by \SetArbEasy. | 59 | v1.9.2. \ae \emph *: Starred version which always puts the stroke over its argument | 59 |
| v1.5. | General: Compatibility with the quran package | 56 | v1.10. General: \uc supersedes \cap | 41 |
| | Environments may be nested inside the arab environment . . | 51 | \prname: New command for typesetting Arabic proper names in transliteration | 61 |
| | txarab : New txarab environment for typesetting running paragraphs in Unicode Arabic . | 61 | v1.11. \arbmark: New command for inserting additional marks in Arabic environments | 62 |
| v1.6. | arabverse : New environment arabverse for typesetting Arabic poetry | 63 | \newarbmark: Allows defining additional sets of Arabic marks | 62 |
| | \bayt: New macro \bayt for typesetting each verse inside the arabverse environment | 63 | v1.12. General: \abjad can now process L ^A T _E X counters | 26 |
| | \SetArbDflt*: This starred version applies the assimilation rules in addition to what \SetArbDflt already does. | 60 | \arbcolor: Standard color command for Arabic environments | 59 |
| | | | \MkArbBreak*: ‘starred’ version which closes Arabic environments before processing declared commands. | 59 |

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 | | |
|-----------------------------|----------------|--|
| \@M | 315 | \adjustbox 191, 207, 211 |
| \@beginparpenalty . | 291 | \advance 258, |
| \@clubpenalty | 317 | 271, 272, 273, 274 |
| \@donoparitem | 279 | \aemph . 28, <u>52</u> , 218, 223 |
| \@ifpackageloaded . | 238 | \aemph* 28, <u>52</u> |
| \@inlabelfalse | 305 | \al@arb@rules 65, 68, |
| \@inlabeltrue | 299 | 69, 72, 73, 105, |
| \@item | 277 | 112, 119, 126, |
| \@itemlabel | 264 | 143, 149, 155, 161 |
| \@itempenalty | 296 | \al@arbuf |
| \@labels | | . 81, 82, 83, 84, 85 |
| 310, 328, 329, 330 | | \al@arbuf@dflt |
| \@listctr | 323 | 80, 81, 83 |
| \@listdepth 255, 258, 263 | | \al@bayt@width |
| \@minipagefalse | 302 | 164, 166, |
| \@mklab | 265 | 167, 168, 191, 211 |
| \@nbitem | 289 | \al@error 243, 244 |
| \@ne | 258, 275 | \al@gutter@width |
| \@newlistfalse | 303 | 165, 167, |
| \@nmbrlistfalse | 266 | 169, 193, 195, |
| \@nbreakfalse | 314 | 201, 203, 207, 209 |
| \@noitemargfalse | 321 | \al@hemistich@delim |
| \@tempa 100, 101, | | 185, 186, 193 |
| 108, 115, 122, | | \al@input@scheme |
| 138, 139, 145, | | 63, 64, 105, |
| 151, 157, 198, 199 | | 112, 119, 127, |
| \@tempboxa 326, | | 143, 149, 155, 162 |
| 327, 334, 335, 337 | | \al@mode 13, 14, |
| \@toodeep | 256 | 15, 16, 99, 137, |
| \@topsep | 292 | 174, 179, 200, 206 |
| \@totalleftmargin | | \al@mode@fullvoc |
| 273, 274, 275 | | 20, 108, 145 |
| \@trivlist | 268 | \al@mode@novoc |
| | | 21, 115, 151 |
| | | \al@mode@trans |
| | | 22, 122, |
| | | 157, 179, 200, 206 |
| _ | 36 | \al@mode@voc 19, 101, 139 |
| | | \al@tatweel 190, 199 |
| A | | \al@trans@convention |
| \abjad | 26, <u>214</u> | 78, 79, 125, 160 |
| \abraces | 33, <u>227</u> | \al@trans@font |
| \addpenalty | 291, 296 | 74, 75, 94 |
| \addvspace 292, 293, 297 | | |
| | | \al@trans@style . 76, |
| | | 77, 94, 123, 158 |
| | | \al@verse@stroke |
| | | 187, 203 |
| | | \al@warning 242, 248 |
| | | \al@wrong@mark 248 |
| | | \al@wrong@nesting . 244 |
| | | <i>Amiri</i> 5 |
| | | amiri (package) 5 |
| | | arab (environment) 8, <u>137</u> |
| | | <i>Arabic-English Lexicon</i> , |
| | | <i>An</i> 20 |
| | | \arabicfont |
| | | 35, 36, 38, 92, |
| | | 98, 103, 110, |
| | | 117, 141, 147, 153 |
| | | arabluatex (package) |
| | | 2–7, 9, |
| | | 11–14, 16–21, |
| | | 26–28, 31, 33, |
| | | 35, 36, 40–45, |
| | | 47–51, 54–61 |
| | | <i>ArabTeX</i> 3, 21 |
| | | arabtex (package) |
| | | 3, 4, 17, 21, 57 |
| | | arabulatex (package) . . 18 |
| | | arabverse (environ- |
| | | ment) 28, <u>164</u> |
| | | arabxetex (package) . 4, 17 |
| | | \arb 7, 90, |
| | | 99, 170, 184, 187 |
| | | \arb@utf 183, |
| | | 191, 207, 209, 211 |
| | | \arbcOLOR 35, <u>61</u> |
| | | \arbmark 33, <u>130</u> , <u>249</u> |
| | | \arbnul 24, <u>226</u> |
| | | \arbuf 42, <u>80</u> |
| | | \ArbUpDflt 42, <u>80</u> |
| | | |
| | | B |
| | | babel (package) 54 |
| | | babel-french (package) . 54 |
| | | \bayt 29, <u>183</u> |

| | | | |
|-----------------------------|--------------------|--------------------------------|-------------------------------|
| bidi (package) | 68 | <i>Grammar of the Arabic</i> | <i>Introduction to Arabic</i> |
| \BODY | 142, 148, 154, 159 | <i>Language, A</i> | <i>Natural Lan-</i> |
| \boolfalse | | 5, 6, 8, 9, 14–20, | <i>guage Process-</i> |
| . 40, 106, 113, 120 | | 26, 27, 42, 54 | <i>ing</i> |
| \booltrue | 39, | gutter (option) | 29 |
| 97, 102, 109, | | | |
| 116, 140, 146, 152 | | | |
| \box | 310, 335 | | |
| Brockelmann, Carl | 11, 40, 42 | | |
| | | H | |
| | | Habash, Nizar Y. | 45 |
| | | \hbox | 187, |
| | | 328, 329, 337, 341 | |
| | | \Height | 191, 207, 211 |
| | | \hfill | 182, 187 |
| | | Hosny, Khaled | 5 |
| | | \hskip | 331, 332, 333, 339 |
| | | \hspace | 195, 201 |
| | | | |
| | | I | |
| | | \if@inlabel | 281, 304 |
| | | \if@newlist | 287 |
| | | \if@nmbrrlist | 322 |
| | | \if@nobreak | 288, 313 |
| | | \if@noitemarg | 320 |
| | | \if@noparitem | 278 |
| | | \ifal@verse@delim | 193 |
| | | \ifal@verse@utf | 180, 184 |
| | | \ifbool | 217, |
| | | 222, 273, 326, 328 | |
| | | \IfBooleanTF | |
| | | . 48, 54, 58, 67, 71 | |
| | | \ifdim | 334 |
| | | \ifhmode | 284 |
| | | \ifluatex | 5 |
| | | \IfNoValueTF | 62, 192, 230 |
| | | \ifnum | 255 |
| | | \ifvoid | 307 |
| | | \ignorespaces | 276, 340 |
| | | \indent | 282 |
| | | <i>Information and Docu-</i> | |
| | | <i>mentation - Ro-</i> | |
| | | <i>manization of</i> | |
| | | <i>the Arabic Al-</i> | |
| | | <i>phabet for Ara-</i> | |
| | | <i>bic, Ottoman-</i> | |
| | | <i>Turkish, Per-</i> | |
| | | <i>sian, Kur-</i> | |
| | | <i>dish, Urdu and</i> | |
| | | <i>Pushto</i> | 11 |
| | | | |
| | | K | |
| | | \kern | 187, 308 |
| | | KOMA-script (package) | 49 |
| | | | |
| | | L | |
| | | \labelsep | 333, 339 |
| | | \labelwidth | 332, 334, 337 |
| | | Lagally, Klaus | 3, 21 |
| | | Lane, Edward William | 20 |
| | | \lastbox | 306 |
| | | \leavevmode | 187 |
| | | \leftmargin | 272, 274 |
| | | lineno (package) | 31 |
| | | \linewidth | 271, 272, 275 |
| | | \list | 254 |
| | | \listparindent | 261, 270 |
| | | \long | 341 |
| | | \LR | 41, 48, 239, 241, 247 |
| | | \LRfootnote | 48, 233 |
| | | \LRmarginpar | 49, 229 |
| | | \luadirect | 34, 49, 50, |
| | | 55, 56, 59, 60, | |
| | | 87, 104, 111, | |
| | | 118, 124, 133, | |
| | | 136, 142, 148, | |
| | | 154, 159, 218, | |
| | | 219, 223, 224, 228 | |
| | | luaotfload (package) | 40 |
| | | \luastring | |
| | | 218, 219, 223, 224 | |
| | | \luastringN | |
| | | | 49, 50, 55, |
| | | 56, 59, 60, 87, | |
| | | 104, 111, 118, | |
| | | 124, 133, 136, 228 | |
| | | \luastringO | |
| | | 105, 112, 119, | |
| | | 125, 126, 127, | |
| | | 142, 143, 148, | |
| | | 149, 154, 155, | |
| | | 159, 160, 161, 162 | |

| | | | |
|---------------------------------------|---------------------|---------------------------------------|--------------------|
| M | | T | |
| <code>\makebox</code> . | 193, 203, 209 | <code>\textcolor</code> | 62 |
| <code>\makelabel</code> | 265, 326, 327 | <code>\textdir</code> | |
| <code>\marginpar</code> . . . | 231, 232 | | 39, 40, 42, 43, |
| <code>metre</code> (option) | 29 | | 45, 46, 91, 93, |
| <code>\MkArbBreak</code> | 47, 50 | | 98, 103, 110, |
| <code>\MkArbBreak*</code> | 47, 51 | | 117, 123, 141, |
| <code>mode</code> (option) | 28 | | 147, 153, 158, |
| N | | | 231, 232, 234, 236 |
| <code>nameauth</code> (package) . . . | 42 | <code>\textsuperscript</code> . . . | 80 |
| <code>\newarbmark</code> . . . | 34, 135 | <code>\textwidth</code> | 166 |
| <code>\newbool</code> | 33 | <code>\the</code> | 263 |
| <code>\NewDocumentEnvironment</code> | | <code>\thefootnotemark</code> . . | |
| | 95, 177 | | 239, 241 |
| <code>\NewEnviron</code> | 137 | <code>\thinspace</code> | 80 |
| <code>\newfontfamily</code> | 38 | <code>trans</code> (option) | 6, 9 |
| <code>\newlength</code> . . . | 164, 165 | <i>Transliteration der ara-</i> | |
| <code>\NoArbUp</code> | 42, 80 | <i>bischen Schrift,</i> | |
| <code>\noindent</code> | 178 | <i>Die</i> . . . | 11, 40, 42 |
| <code>novoc</code> (option) | 6, 9 | <code>txarab</code> (environment) | |
| O | | | 47, 95 |
| options: | | <code>txarab</code> (environment) . . | 56 |
| <code>delim</code> | 29 | <code>\txarb</code> | 47, 91, 184 |
| <code>fullvoc</code> | 6, 9 | <code>\txtrans</code> | 91 |
| <code>gutter</code> | 29 | U | |
| <code>metre</code> | 29 | <code>\Uc</code> | 88 |
| <code>mode</code> | 28 | <code>\uc</code> | 40, 86, 88, 90 |
| <code>novoc</code> | 6, 9 | <code>\unhbox</code> | 330, 337 |
| <code>trans</code> | 6, 9 | <code>\unskip</code> | 285 |
| <code>utf</code> | 29 | <code>utf</code> (option) | 29 |
| <code>voc</code> | 5, 9 | V | |
| <code>width</code> | 28 | <code>voc</code> (option) | 5, 9 |
| P | | W | |
| <code>\par</code> | 96, 98, 138, | <code>\wd</code> | 334 |
| | 163, 178, 282, 285 | <code>width</code> (option) | 28 |
| <code>paralist</code> (package) . | 52, 54 | <code>Wright, W. LL.D</code> . . | 5, |
| <code>\pardir</code> | 39, 40, | 6, 8, 9, 14–20, | |
| | 98, 141, 147, | 26, 27, 42, 54 | |
| | 153, 158, 233, 235 | X | |
| <code>\parindent</code> | 270 | <code>xcolor</code> (package) | 35 |
| <code>\parsep</code> | 269 | <code>\xleaders</code> | 187 |
| <code>\parshape</code> | 275 | Z | |
| <code>\parskip</code> | 269, 293 | <code>\z@</code> | 260, 261, |
| <code>\penalty</code> | 311 | | 262, 306, 307, 311 |
| <code>polyglossia</code> (package) | 5, 64 | | |
| <code>\presetkeys</code> | 175 | | |
| <code>\prname</code> | 41, 89 | | |
| Q | | | |
| <code>\ProcessOptionsX</code> . . . | 18 | | |
| <code>\ProvidesPackage</code> . | 2, 252 | | |
| R | | | |
| <code>\refstepcounter</code> . . . | 323 | | |
| <code>reledmac</code> (package) . . | 56 | | |
| <code>\RequirePackage</code> . . . | | | |
| . . . | 4, 12, 23, 24, | | |
| | 25, 26, 27, 28, | | |
| | 29, 30, 31, 32, 240 | | |
| <code>\rightmargin</code> | | | |
| | 260, 271, 273 | | |
| <code>\RL</code> . . . | 44, 48, 246, 247 | | |
| <code>\RLfootnote</code> . . . | 48, 233 | | |
| <code>\rmfamily</code> | | | |
| . . . | 42, 43, 45, 46, 74 | | |
| <code>\romannumeral</code> | 263 | | |
| S | | | |
| <code>\savenotes</code> | 189 | | |
| <code>\sbox</code> | 327 | | |
| <code>scrextend</code> (package) | 49, 65 | | |
| <code>\SetArbDflt</code> | 7, 65 | | |
| <code>\SetArbDflt*</code> | 7, 65 | | |
| <code>\SetArbEasy</code> | 6, 65 | | |
| <code>\SetArbEasy*</code> | 6, 65 | | |
| <code>\SetArbUp</code> | 42, 80 | | |
| <code>\setbox</code> | | | |
| | 306, 328, 329, 341 | | |
| <code>\SetHemistichDelim</code> | | | |
| | 29, 183, 186 | | |
| <code>\SetInputScheme</code> . . . | | | |
| | 45, 63, 132 | | |
| <code>\setkeys</code> | 178, 182 | | |
| <code>\setlength</code> | | | |
| | 166, 167, 168, 169 | | |
| <code>\setLR</code> | 40, 49, 180 | | |
| <code>\setRL</code> . | 39, 49, 180, 181 | | |
| <code>\SetTranslitConvention</code> | | | |
| | 40, 78 | | |
| <code>\SetTranslitFont</code> | 40, 74 | | |
| <code>\SetTranslitStyle</code> . | | | |
| | 40, 76, 90 | | |
| <code>\space</code> | 245 | | |
| <code>\spewnotes</code> | 212 | | |
| <code>\sRLbox</code> | 326, 341 | | |