

User's Guide to ArabSpell Version 1.00

Otakar Smrž

April 30, 2000

This User's Guide may be modified and improved independently of the current ArabSpell release. That will happen if we reveal some serious inaccuracies or if more detailed commentary on any of the features is required.

All components of ArabSpell are copyrighted, but free use for noncommercial purposes is granted. There is no warranty of any kind and the entire risk rests with the user.

In case you have any query, suggestion of improvement or if you find a bug, do not be shy to get in touch with the author:

Otakar Smrž
Topolská 749
537 05 Chrudim
Czech Republic
arabic@seznam.cz

Corrections of my English and translations of the error messages into other languages are welcome as well and will carry the name of their writer. Thank you very much for your interest.

Copyright © 1999, 2000 Otakar Smrž

Contents

1	Introduction to ArabSpell	3
2	Obtaining and Running ArabSpell	5
2.1	Interactive execution	5
2.2	Parameter execution	6
2.3	Editing parameter & language files	7
2.4	Output error files	7
2.5	Letter statistics	7
3	Document Preparation	9
3.1	Preferred coding	9
3.2	ArabSpell commands	12
3.3	Limitations	12
4	Error Messages	13
4.1	On the edge	13
4.2	On thin ice	16

<i>CONTENTS</i>	2
Acknowledgements	17
Bibliography	18
A On <i>Hamza</i> Writing	19

Chapter 1

Introduction to ArabSpell

During our work on some processing software for the *Arabic-Czech and Czech-Arabic Dictionary* being compiled by Opus arabicum, an idea of creating a simple interface between a common ArabTeX source file and our generalized spelling algorithms descended from heavens.

ArabSpell is a newborn tool for checking the correctness of ASCII Transliteration encoding normalized by ArabTeX [1]. The three principal aims of the program are following:

- to facilitate *beginner's first contact* with ArabTeX transcription rules for Arabic by giving advice and explanation
- to help anyone including experts reveal *a certain type of typing errors* in text documents which use the specified encoding of the Arabic language
- to perform basic *letter statistics* on larger Arabic data recorded in this encoding

The spelling algorithms were originally devised for checking an arising dictionary, and therefore ArabSpell operates with almost no vocabulary. Its power

rests in verification of some formal rules of Arabic morphology. The motto runs as follows [4]:

Every syllable in literary Arabic starts with one and only one consonant.

For the reasons stated, do not rely on ArabSpell to leave your document perfect. It does not distinguish between `ma_taluN` مَثَلٌ *maṭalun* and `ma_tluN` مَثَلٌ *maṭlun*, `mi_tluN` مِثْلٌ *mitlun*, `ma_tIlLuN` مَثِيلٌ *maṭīlun*, `mi_tAluN` مِثَالٌ *miṭālun*, `ma_tAluN` مَثَالٌ *maṭālun*, `maAluN` مَالٌ *malun* or whatsoever. Nevertheless it does protest if you write `ma_talN` مَثَلٌ *mataln*, `m_taluN` مَثَلٌ *mtalun*, `ma_talUN` مَثَلٌ *matalūn*, `ma_talun` مَثَلٌ *matalun* etc.

It is now obvious that the input text must be *properly vocalized*, otherwise ArabSpell cannot cut the words into syllables and the whole thing becomes senseless. Read on (or skip to 3.2) to see the way of excluding non-vocalized parts of your data from processing.

When using ArabSpell, you should also be aware of the fact that the program is not a grammar-checker at all. It is still you who must keep your Arabic Arabic.

Chapter 2

Obtaining and Running ArabSpell

You may download the current version of ArabSpell packed as `aspell.zip` from <http://come.to/arabspell>.

Create an extra directory on your disk and extract all files into it. The total occupied disk space hardly ever exceeds 1 MB!

To run ArabSpell, you only need an operation system which supports DOS applications. There are two main ways of executing the program — either with some parameters on the command line, or without any. Try the latter alternative first.

2.1 Interactive execution

Having chosen your preferred language and specified the name of the file (or the keyboard input) to check, you answer some questions in order for ArabSpell to report errors and suggest corrections accordingly.

Handling the program afterwards should be very primitive, and that is why we shall say no more about it here.

2.2 Parameter execution

The executable file swallows up to two parameters, as you can see in the verbatim output:

```
C:\ArabSpell>aspell.exe /?
```

```
syntax: aspell [interface_opt] file_name[.file_ext]
```

```
interface_opt: file with interface options
```

The default extension for the file to check is `.tex`, the interface parameter file takes `.ipf` implicitly. To specify a file which has no extension, type a dot `.` after its name.

If you do not indicate any interface parameter file, default values are assigned to the variables (see the original `aspell.ipf`, too): `Language = 1`, `Screen = T`, `ForTeX = F`, `ArabContext = F`, `SpellNovocalize = T`.

In contrast with interactive execution, here you spare the necessity to set the options in each run. We believe that you will find out ways of utilizing this feature. A simple batch file can check several files at once, or another may take the appropriate steps and serve ArabSpell the current file being edited e.g. in TextPad, after you have pressed just one shortcut key¹!

¹This invention comes from Karel Mokřý, my close colleague in the Dictionary team, and I owe him thanks. He proposed `Ctrl+E` for ArabSpell.

2.3 Editing parameter & language files

All *interface parameter files* require a fixed line format, otherwise the value of the variable would not be reset. ArabSp ϵ ll is case and blank sensitive. The variable links up with its value exclusively by the sequence `□=□`.

If you want to *translate the error messages* into your native language, copy the files `aspell.eng` and `aspells.eng` to new ones and rename their extensions. Do not mix, add or delete the lines! The total length of the sentences on one line may not exceed 249 characters.

Then, edit `aspell.11f`, the essential *language list file*. When creating your own item, pay a bit of your attention to the `□|□` separators.

2.4 Output error files

Independently of your wishes, ArabSp ϵ ll produces an *error file* differing from the input file name in extension `.err`. This file contains non-tagged error messages and source strings which they refer to. You may keep it as the result of the check and put off the corrections till the next day.

Optionally, you can get a *TEX error file* with extension `.ter` and its contents much neater. With the help of `roman.def` or `arabic.def`, it is ready to be ArabT ϵ Xed.

2.5 Letter statistics

For many of Arabic enthusiasts, *letter statistics file* denoted by `.stt` might be the most interesting. You may look up the absolute and relative frequency of occurrence within Arabic context to each ASCII ordinal number and read its incidental Arabic letter interpretation.

There is a limitation as to the size of the input file. The Arabic context must not count more than $256^2 - 1$ letters (not characters) to be sure that there was no range overflow. That means roughly 820 lines of plain Arabic text, 80 letters per line.

Statistical data from several input files may be summed up e.g. in Excel and processed in various ways. Remember that ArabSpell removes redundant and marginal blanks from the string before performing other operations.

Chapter 3

Document Preparation

3.1 Preferred coding

Please, compare the following instructions with [1]. Note that diacritical marks of the Arabic transcription, i.e. $_$, \cdot and $\hat{}$, must always be indicated *before* their carrier.

- The regular consonants are listed in the table. ArabSpell regards any illegal character as a consonant, too.

ﺀ	’	ء	<i>d</i>	<i>d</i>	د	<i>ḍ</i>	<i>.d</i>	ض	<i>k</i>	<i>k</i>	ك
<i>b</i>	<i>b</i>	ب	<i>ḏ</i>	<i>_d</i>	ذ	<i>ṭ</i>	<i>.t</i>	ط	<i>l</i>	<i>l</i>	ل
<i>t</i>	<i>t</i>	ت	<i>r</i>	<i>r</i>	ر	<i>ẓ</i>	<i>.z</i>	ظ	<i>m</i>	<i>m</i>	م
<i>ṭ</i>	<i>_t</i>	ث	<i>z</i>	<i>z</i>	ز	‘	‘	ع	<i>n</i>	<i>n</i>	ن
<i>ǧ</i>	<i>ˆg</i>	ج	<i>s</i>	<i>s</i>	س	<i>ǧ</i>	<i>.g</i>	غ	<i>h</i>	<i>h</i>	ه
<i>ḥ</i>	<i>.h</i>	ح	<i>š</i>	<i>ˆs</i>	ش	<i>f</i>	<i>f</i>	ف	<i>w</i>	<i>w</i>	و
<i>ḫ</i>	<i>_h</i>	خ	<i>ṣ</i>	<i>.s</i>	ص	<i>q</i>	<i>q</i>	ق	<i>y</i>	<i>y</i>	ي

- Short vowels are coded *u*, *i* and *a*, the long ones *U*, *I* and *A*. Occasional *uw* for *U* and *iy* for *I* are tolerated, *aa* is not.

- Defective notations `_a` and `_U` can be understood well: `'il_ahuN` إِلَهٌ *ilāhun*, `l_akinna` لَكِنَّ *lākinna*, `.t_ah_a` طَه *tāhā*, `h_aka_dA` هَكَذَا *hākadā*, `h_a_dA` هَذَا *hādā*, `_d_alika` ذَلِكَ *dālīka*, `h_a'ulA'i` هُوَ لِأَيِّ *hā'ulā'i*, `'_Ul_a'ika` أُولَئِكَ *ulā'ika*
- Historical writing of long vowels, i.e. `_i`, `_u`, `^A`, `^I`, `^U` and newly `_aU`, `_aI`, `_aY`, is enabled. Even so, ArabSpell will leave you a message.
- The indefinite article must be distinguished by capitalization: `baytuN` بَيْتٌ *baytun*, `madInaTiN` مَدِينَةٌ *madīnatin*, `^sa_h.saN` شَخْصًا *šaḥṣan*, `mA.diN` مَادِيْنٍ *māḍīn*, `'a.saN` عَصًا *aṣan*. In formal masculines, `-aNA` and `-aNU` pass with a report of their occurrence.
- The definite article is separated by `-` and assimilated: `al-.hamdu` الْحَمْدُ *al-ḥamdu*, `a.z-.zillu` الظِّلُّ *az-zillu*, `ad-dirAsaTu` الدِّرَاسَةُ *ad-dirāsatu*, `al-'asadu` الْأَسَدُ *al-'asadu*, `al-"i^gtimA'u` الْأَجْتِمَاعُ *al-iġtimā'u* (the double quote " is necessary, otherwise we get الْأَجْتِمَاعُ *al-iġtimā'u*)
- You must have noticed `tā'` *marbūṭa* T, the ending of the formal feminine singular: `mu'minaTuN` مُؤْمِنَةٌ *mu'minatun*, `fatATaN` فَتَاةٌ *fatātan*, `'as'ilaTuN` أَسْئَلَةٌ *as'ilatun* (however `'as'ilatunA` أَسْئَلَتُنَا *as'ilatunā*)
- Being a long vowel in fact, `Y` generates *'alif maqṣūra*: `'atY` أَتَى *atā*, `'aLY` عَلَى *alā*, `yalqY` يَلْقَى *yalqā*, `kubrY` كُبْرَى *kubrā*
- In words where final `-ayun`, `-ayin` or `-ayan` group transformed to `-an`, the indefinite article is coded `-aNY`: `quwaNY` قُوَى *quwan*, `ma'naNY` مَعْنَى *ma'nan*, `mu.ga.t.taNY` مُغَطَّى *mugaṭṭan*

- In perfect, subjunctive, apocopate and imperative forms of 3rd person masculine plural, *silent 'alif* is denoted
 - either by -UA: .sArUA صَارُوا *ṣārū*, .ha^g^gUA حَجُّوا *ḥaġġū*, yam^sUA يَمَشُّوا *yamšū*, yu'.tUA يُعْطُوا *yuṭū*, iytUA اَيْتُوا *iytū*
 - or by -aWA: da'aWA دَعَا *da'aw*, yalqaWA يَلْقَا *yalqaw*, tatamannaWA تَتَمَنَّوْا *tatamannaw*, raWA رَا *raw*
- The prepositions *bi-* and *ka-* are tied by a hyphen: bi-bu.t'iN بِبُطْءٍ *bi-butʔin*, bi-intibAhiN بِإِنْتِبَاهٍ *bi-'ntibāhin*, bi-al-qalami بِالْقَلَمِ *bi-'l-qalami*, bi-a^s-^sawkaTi بِالشُّوكَةِ *bi-'š-šawkati*, bi-yadayhi بِيَدَيْهِ *bi-yadayhi*, bi-dUni بِدُونِ *bi-dūni*, ka-al-waladi كَالْوَلَدِ *ka-'l-waladi*, ka-'annanI كَأَنَّي *ka-ʔannanī*, ka-_d_alika كَذَلِكَ *ka-dālika*
- The preposition *li-* must be treated carefully enough if followed by a definite article and
 - the consonant *lām*: li-l_ahi لِلْهِ *li-llāhi*, li-l_aylaTi لِلَّيْلَةِ *li-llaylati*
 - the other ones: lil-maliki لِلْمَلِكِ *lil-maliki*, lil-'asafi لِلْأَسْفِ *lil-ʔasafi*, lin-nawmi لِلنَّوْمِ *lin-nawmi*, li.s-.sadIqaTi لِلصَّدِيقَةِ *liš-ṣadīqati* (writing *lil-l_ahi* produces nonused *lil-lāhi*)
- The particles *sa-*, *li-* and *fa-* are hyphenated as well, another short words should be written separately: sa-taktubIna سَتَكْتُبِينَ *sa-taktubīna*, li-'as'alaka لِأَسْأَلُكَ *li-ʔasʔalaka*, fa-in.sarafa فَانصَرَفَ *fa-'nṣarafa*, wa lahu وَ لَهُ *wa lahu*, fI baytI فِي بَيْتِي *fī baytī*

3.2 ArabSpell commands

Open `aspell.def` to see the definitions of new commands. `\ArabSpell` along with `\ArabTeyX` enable you to typeset the equation

$$\text{ArabSpell} + \text{ArabTeX} = \text{ArabTeX}$$

The next two ones, which may enclose *non-vocalized* parts of the source file, mathematical insertions or interfering verbatim sequences, are much more profitable. `\ArabSpellOff` tells ArabSpell to ignore the input until `\ArabSpellOn` crops up. Do not confuse this feature with the effect of the `SpellNovocalize` variable which considers the declared vocalization degree of the output.

3.3 Limitations

ArabSpell cannot handle lines longer than 255 characters. Note that ends of lines must conform to DOS needs if you take your data from another platforms.

For the sake of letter statistics, keep the size of your files moderate if possible (see 2.5). Editing such files is more convenient anyway.

ArabSpell does not trace your own macros, which implies that you only may indicate the extent of the Arabic context by the standard ArabTeX means and by the `ArabContext` variable.

As regards grouping, braces may be nested up to the tenth level.

Chapter 4

Error Messages

ArabSpell reads the input file line by line. As soon as a compact sequence of Arabic context is recognized, spelling algorithms are applied to it. Thus a line may be split into shorter substrings.

In case a suspect word is found, its position within the highlighted expression and a relevant error message are reported. A word may be accused of a series of crimes at the same time. Make your corrections gradually, for some of the discrepancies just happen to be induced.

4.1 On the edge

There is a sort of error messages in ArabSpell adequacy of which must be judged by a human. They were implemented in order to prevent the most silly, yet very easy-to-make typing errors.

Check the justification of the ending -AT-! Such an ending is alright in words like fatATuN فَتَاةٌ *fatātun*, .hayATuN حَيَاةٌ *ḥayātun*, .saLATuN صَلَاةٌ *ṣalātun* having *wāw* or *yā'* as the final consonant of their root. These

semi-vowels usually reappear in plural, e.g. fatayAtuN فَتَيَاتُ *fatayātun*,
.hayawAtuN حَيَوَاتُ *ḥayawātun*, .salawAtuN صَلَوَاتُ *ṣalawātun*.

If we are not dealing with a form of a verb, then you may have made a typing error in the second last character of the word and wrote **t** instead of **T**! Confined by its working methods, ArabSpell is unable to say that sakata سَكَّتَ *sakata* versus sanaTa سَنَةً *sanata*, yatalaffatu يَتَلَفَّفْتُ *yatalaffatu* versus ma`gallaTu مَجَلَّةٌ *mağallatu*, fa-darasati al-fatATu فَدَرَسَتِ الْفَتَاةُ *fa-darasati al-fatātu* versus li-dirAsaTi al-lu.gAti لِدِرَاسَةِ اللَّغَاتِ *li-dirāsati al-luġāti* are all correct expressions.

If the word is fully vocalized and if we are not dealing with an apocopate or an imperative, then there should come an indefinite article **N** instead of the final character **n**! Longer, good-looking words¹ with the exception of l_akin لَكِنْ *lākin* and 'i_dan إِذَنْ *ʾidān* (or 'i_daN إِذًا *ʾidān*) generate this *tanwīn* warning automatically. It may bother you in the indicated cases.

Let us have a look at some examples. From sakana سَكَنَ *sakana*, the apocopate yaskun يَسْكُنُ *yaskun* and the imperative uskun اُسْكُنْ *uskun* can be derived. ArabSpell objects if you replace mA 'atqana مَا أَتَقَنَّ *mā ʾatqana* by lam yutqin لَمْ يُتَقَنَّ *lam yutqin*. 'atqin أَتَقَنَّ *ʾatqin* is no better. Similarly, yakun يَكُنْ *yakun* from kAna كَانَ *kāna* is the stumbling block, however kun كُنْ *kun* is too short for ArabSpell to expose to your scrutiny.

¹Neither do 'an أَنْ *ʾan*, min مِنْ *min*, 'an عَنْ *ʿan* etc. nor lub"nAn لُبْنَان *lubnān* rank in this category.

The word might possibly end with a silent 'alif -UA! The detailed analysis is given in the documentation. Here it comes! In general, occurrence of -U at the very end of the word is scarce. There is a set of words called *al-'asmā' al-ḥamsa* taking this shape: 'abU أَبُو *abū*, 'a_hU أَخُو *aḥū*, .hamU حَمُو *ḥamū*, fU فُو *fū*, _dU ذُو *dū*.

In addition, final -U can show up as a reduced form of the masculine plural ending -Una in genitive phrases, e.g. mu'allimU al-madrasaTi مُعَلِّمُو الْمَدْرَسَةِ *mu'allimū al-madrasati*, _dawU ذُو شَانٍ *dawū ṣa'niN* *ša'nin*, _dawU ال-إِقْطَاعِ ذُو *dawū al-ʾiqṭāʿi*, 'UlU ال-'amri أُولُو الْأَمْرِ *ulū al-ʾamri*.

As for verbs, imperfect indicatives with *wāw* as the final root consonant have no reason to take *silent 'alif* either: 'ar^gU أَرْجُو *arġū*, tad'U تَدْعُو *tad'ū*, ya^skU يَشْكُو *yaškū*, nam.hU نَمَحُو *namḥū*.

The word might possibly end with a silent 'alif -aWA or -aW! The detailed analysis is given in the documentation. There is always a lot of fun about Arabic irregular roots. For instance, the nominative plural of *mustawliN* مُسْتَوِلٍ *mustawlin* reads *mustawlUna* مُسْتَوِلُونَ *mustawlūna* which is quite a digestible word, whereas in the passive participle *mustawlaNY* مُسْتَوَلِيٌ *mustawlan* becomes *mustawlawna* مُسْتَوَلُونَ *mustawlawna*, or *mustawlaw* مُسْتَوَلُوٌ *mustawlaw* in a genitive phrase.

Because of examining *raWA* رَوَا *raw* and all its look-alikes, 'aw أَوْ *aw* or *law* لَوْ *law* fall into the trap, too.

4.2 On thin ice

There are some omissions and drawbacks in ArabSpell the most serious of which follow:

- If you use hyphens in other than suggested functions, the words might not be spelled properly. We will try to fix this in the forthcoming versions.
- On account of morphological ambiguity, *hamza* correction system cannot be complete. In questionable cases, see the Appendix for reference.

Acknowledgements

I would like to thank my brother Jan Smrř for having programmed several Turbo Pascal procedures used in ArabSpell code to make multilingual appearance possible. Jan Smrř is credited with a great deal of the Dictionary software, too.

I owe the theoretical grounding in Arabic to my kind teachers Mgr. Jiří Fleissig and PhDr. Charif Bahbouh, CSc.

My thanks also belong to Karel Mokřý, a T_EXpert I always ask advice of, and to Jitka Charvátová who provided for some English linguistic terms.

Bibliography

- [1] Klaus Lagally: *ArabTeX — a System for Typesetting Arabic, User Manual Version 3.09*.
Institut für Informatik, Universität Stuttgart, 1999
- [2] Jiří Fleissig, Charif Bahbouh: *Základy moderní spisovné arabštiny, I. díl (Essentials of Modern Literary Arabic, Book I)*.
Dar Ibn Rushd, Prague, 1992 (1st edition) and 1997 (2nd edition).
- [3] Jiří Fleissig, Charif Bahbouh: *Základy moderní spisovné arabštiny, II. díl (Essentials of Modern Literary Arabic, Book II)*.
Dar Ibn Rushd, Prague, 1995.
- [4] Karel Petráček: an unknown script on Arabic morphology and semantics.
Faculty of Arts, Charles University, Prague.
- [5] Luboš Kropáček: *Arabsko-český a česko-arabský slovník (Arabic-Czech and Czech-Arabic Dictionary)*.
Státní pedagogické nakladatelství, Prague, 1987 (4th edition).
- [6] al-Majani: *Arabic-English and English-Arabic Pocket Dictionary*.
Dar al-Majani, Beirut, 1996.

Appendix A

On *Hamza* Writing

As we know, the carrier of *hamza* is determined by *her* position and adjacency. Such decision-making is up to ArabTeX. But it fully rests with us whether to indicate *hamza* or not!

Let us formulate the criteria of *hamza* writing as observed from [2, 3, 5] and present some remarkable examples. This study was inspired by [1] and its earlier versions, although and since they do not keep to it systematically.

Being one of the *moon consonants*, *hamza* may occur

- in the root of a word: 'a_taruN أَتْرُ \triangleright atarun, ra'IsIyuN رَيْسِي \triangleright ra'sīyūn,
'AlAtuN آلَات \triangleright ālātun, qara'a قَرَأَ \triangleright qara'a, 'Uly أُولَى \triangleright ulā, ra''asa رَأْسُ
ra''asa, mutasa' 'iluN مُتَسَالٍ \triangleright mutasa'ilun, 'imAraTuN إِمَارَةٌ \triangleright imāratun,
'ay.daN أَيضًا \triangleright ayḍan, 'A_ha_da أَخَذَ \triangleright āḥada, 'i.hdY إِحْدَى \triangleright iḥdā
- as a morphological consonant
 - in noun-like words: 'an.zAruN أَنْظَارٌ \triangleright anẓārūn, 'alwAnuN أَلْوَانٌ \triangleright alwānūn,
'albisaTuN أَلْبِسَةٌ \triangleright albīsatun, 'A_tAruN أَتَارٌ \triangleright ātārūn, 'AdAbuN

أَدَابٌ \triangleright *ādābun*, 'a'mY أَعْمَى \triangleright *a'mā*, .hamrA'u حَمْرَاءُ \triangleright *hamrā'u*, 'a.s.garu
أَصْغَرُ \triangleright *ašgaru*, 'arba'aTu أَرْبَعَةٌ \triangleright *arba'atu*

- in verbs and *mašdars* of 4th stem and in every 1st person singular imperfect form of a verb: 'a^gIasA أَجَلَسَ \triangleright *ağlasa*, 'AmanA آمَنَ \triangleright *āmāna*, 'arAda أَرَادَ \triangleright *arāda*, 'arsil أَرْسَلَ \triangleright *arsil*, 'a'in أَعْيَنَ \triangleright *a'in*, 'ari أَرَى \triangleright *ari*, 'intA^guN إِنْتَاجٌ \triangleright *intāğun*, 'iqAmaTuN إِقَامَةٌ \triangleright *iqāmatun*, 'I^gAruN إِيجَارٌ \triangleright *iğārun*, 'aktubu أَكْتُبُ \triangleright *aktubu*, 'u'allimu أُعَلِّمُ \triangleright *u'allimu*, 'aqtari.hu أَقْتَرِحُ \triangleright *aqtarihu*, 'asta.tI'u أَسْتَطِيعُ \triangleright *astatī'u*

- in other words: 'ilY إِلَى \triangleright *ilā*, 'an أَنْ \triangleright *an*, 'illA إِلَّا \triangleright *illā*, 'ayna أَيْنَ \triangleright *ayna*, 'anti أَنْتِ \triangleright *anti*, 'amma أَمَّا \triangleright *ammā*, 'ayyaTuN أَيَّةٌ \triangleright *ayyatun*, 'i_daN إِذَا \triangleright *idan*, 'aw أَوْ \triangleright *aw*, 'ayyuhA أَيُّهَا \triangleright *ayyuhā*, '_Ul_a'ika أَوْلِيكَ \triangleright *ulā'ika*

On the contrary, the *hamza* mark does not belong before so called auxiliary vowels (otherwise *wašla* could not be introduced)

- in definite articles: al-qAhiraTu الْقَاهِرَةُ \triangleright *al-qāhiratu*, mina al-qAhiraTi مِنَ الْقَاهِرَةِ \triangleright *mina al-qāhirati*, ad-dawA'u الدَّوَاءُ \triangleright *ad-dawā'u*, 'ani ad-dawA'i عَنِ الدَّوَاءِ \triangleright *'ani ad-dawā'i*
- in prefixes of 7th to 10th stem: in.tilAQuN إِنْطِلَاقٌ \triangleright *inṭilāqun*, inta.zara إِنْتَظَرَ \triangleright *intazara*, u'tubira أُعْتَبِرَ \triangleright *ustubira*, i_htAra إِخْتَارَ \triangleright *ihtāra*, izdAda إِزْدَادَ \triangleright *izdāda*, i^starY إِشْتَرَى \triangleright *ištarā*, idda'Y إِدْعَى \triangleright *idda'ā*, itta.hada اِتَّحَدَ \triangleright *ittahada*, i'tamara اِتَّمَرَ \triangleright *itamara*, i.hmarra اِحْمَرَ \triangleright *iħmarra*, istifAdaTuN اِسْتِفَادَةٌ \triangleright *istifādatun*, istI.sAluN اِسْتِصَالٌ \triangleright *istiṣālun*

- in imperatives of 1st stem: uktub اُكْتُبْ *uktub*, u.s.hu اُضْحِ *uṣḥu*, irmi اِزْمِ *irmi*, is'alI اِسْأَلِي *isʿalī*, ifrirna اِفْرِزْنَ *ifrirna*, iyqa.z اِيَقْظْ *iyqaz* (qAla iyqa.z قَالَ اَيَقْظُ *qāla iyqaz*), uwmul اُوْمَلْ *uwmul* (qAla uʿmul قَالَ اُوْمَلْ *qāla uʿmul*), iysaf اِيَسْفْ *iysaf* (qAla iʿsaf قَالَ اَيَسْفُ *qāla iʿsaf*)
- in relative pronouns: alla_dI اَلَّذِي *alladī*, al-la_dAni اَللَّذَانِ *al-ladāni*, al-la_dayni اَللَّذِينَ *al-ladayni*, alla_dIna اَلَّذِينَ *alladīna*, allatI اَلَّتِي *allatī*, al-latAni اَللَّتَانِ *al-latāni*, al-latayni اَللَّتَيْنِ *al-latayni*, al-lAtI اَللَّاتِي *al-lātī*, al-lawAtI اَللَّوَاتِي *al-lawātī*, al-lAʿI اَللَّائِي *al-lāʿī*
- in some more words: i_tnAni اِثْنَانِ *itnāni*, ismuN اِسْمٌ *ismun*, ibnuN اِبْنٌ *ibnun* (but beware of 'a_tnA'a اِثْنَاءُ *ʾatnāʿa*, 'asma'uN اَسْمَاءُ *ʾasmāʿun*, 'abnA'uN اَبْنَاءُ *ʾabnāʿun*)