

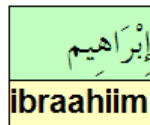
Highlight Arabic Editor

Compose, Edit, and Analyze Complex Arabic Documents

The Highlight Arabic Editor provides the language professional a powerful environment for composing, editing, and analyzing complex Arabic documents. It brings together a rich suite of analytical tools in a single framework, including automatic diacritization, automatic transliteration, online dictionaries, and syntactic analysis. It also provides the technology professional with a flexible tool for developing software applications which process Arabic text.

A unique feature of Arabic Editor is its system for entering and editing fully diacritized Arabic text from a standard PC keyboard, also known as the "QWERTY" layout. The input system is based on a transcription scheme used to approximate Arabic sounds in English. It is easily learned in less than one hour, yet provides productivity and accuracy gains of 2x to 4x that of a conventional Arabic keyboard layout.

For example, to type the Arabic name إبراهيم, the user enters "ibraahiim". The input box at right appears as the user types.



As the name is typed, the Latin spelling appears in the yellow box from left-to-right, while the Arabic spelling appears in the green box from right-to-left. The input box disappears upon completion of a word, but can be re-entered when desired to edit existing text. Text may be entered with all, some, or no diacritical marks.

REVERSIBLE TRANSLITERATION

Arabic Editor's input system is based upon Basis Technology's proprietary, fully-reversible transliteration system for Modern Written Arabic. This system provides a guaranteed "round trip" for any Arabic text (in Unicode) into and out of the Latin alphabet. It is an intuitive and meaningful representation for Arabic speakers while also being easy to learn for non-speakers.

The following sentence is presented both in Latin script form and Arabic script form:

khaTaba al-shaykh ams qaa'ila-n inna al-qaahirā madiinā `aZiimā.

حَطَبَ الشَّيْخِ أَمْسَ قَائِلًا إِنَّ الْقَاهِرَةَ مَدِينَةٌ عَظِيمَةٌ.



AUTOMATIC ORTHOGRAPHY

Arabic Editor handles all the difficult tasks associated with Hamza (ء) "seating", by automatically choosing the correct "chair". To enter a Hamza, the user need only enter an apostrophe (') and the correct orthography is automatically presented:

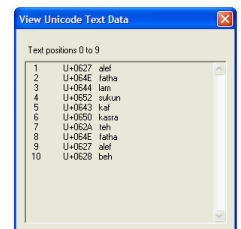
Entry	Result	Entry	Result	Entry	Result
jaa'a	جَاءَ	qur'aan	قُرْآن	ra'iisu-n	رَئِيسٌ
juz'a-n	جُزْءًا	ijraa'aat	إِجْرَاءَات	ar-ru'uusu	الرُّؤُوسُ
qara'a	قَرَأَ	sha'nu	شَأْنٌ	ra'asa	رَأَسَ

Unlike conventional Arabic keyboards, no special keystroke is required to type the Lam-Alif ligature. This is handled automatically:

laa ⇒ لا al-islam ⇒ الإسلام

MULTIPLE TEXT ENCODING SYSTEMS

Arabic Editor supports all of the major text encoding systems used in the Microsoft Windows environment, including Code Page 1256, ISO 8859-6, Unicode UTF-8, and Unicode UTF-16.



A built-in Unicode text inspector is also provided. Invoking this inspector on the word الكِتَاب yields the display shown at right.

FUZZY SEARCH

Arabic Editor's "Fuzzy Search" capability searches Arabic text using approximate Latin strings as input criteria. For example, search input such as "Gaddafi", "Qadhafiy", "Khadaafy", or any of the countless variants of the former Libyan ruler's surname will find the one correct Arabic spelling — القذافي — within a text.

AUTOMATIC TRANSLITERATION

Arabic Editor contains built-in support for the following transliteration systems:

Scheme	Developer	Attributes
Basis	Basis Technology	reversible, phonetic
BGN	Board on Geographic Names	partially reversible, phonetic
Buckwalter	Tim Buckwalter / QAMUS	reversible, non-phonetic
FBIS	Foreign Broadcast Information Service (DNI Open Source Center)	non-reversible, phonetic
IC	U.S. Intelligence Community	non-reversible, phonetic
SATTS	Standard Arabic Technical Transliteration System	reversible (consonants only), non-phonetic

Invoking the transliterator on the following text:

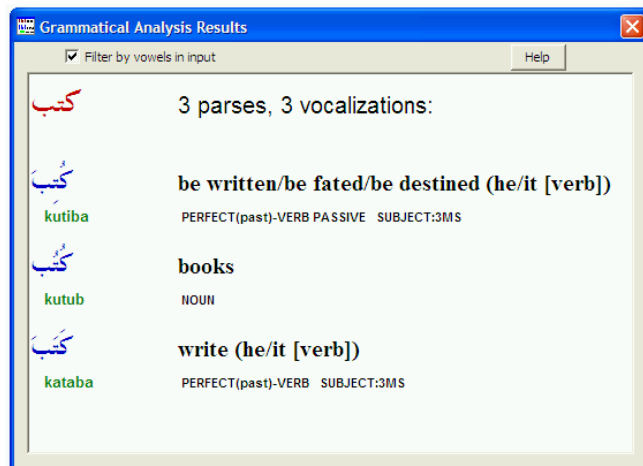
الشَّيْخُ عَبْدُ الْمَجِيدِ الْخُوَيْي

yields the following display:



GRAMMATICAL ANALYSIS

Arabic Editor's grammatical analyzer allows the user to parse Modern Written Arabic on a word-by-word basis. For example, grammatical analysis of the word كتب produces the following display:



The left column of the grammatical analysis results window shows possible vocalizations of كتب by adding the proper combinations of short vowels to the Arabic script. Below each vocalization, the Basis transliteration of the Arabic word is presented in green text. To the right, the corresponding translation and part-of-speech tag is displayed for each parsing.

EXPLORE FURTHER

For more information or to request an evaluation copy, please call us at 617-386-2090 or 800-697-2062, or write to info2011@basistech.com. We will be happy to assist you in evaluating the performance of our product on your data.

VISIT www.basistech.com WRITE info2011@basistech.com CALL 617-386-2090

One Alewife Center
Cambridge, MA 02140

171 Second Street
San Francisco, CA 94105

2553 Dulles View Drive
Herndon, VA 20171

9-6 Nibancho, Chiyoda-ku
Tokyo 102-0084

