Objective
To provide tools that help analysts create more effective search queries in Iraqi and other Arabic dialects.

Relevance
The Problem—In Arabic the written form of a word does not always reflect the spoken form; this is especially true of dialects. Furthermore, the spoken or written form of a word may not match the dictionary citation form. Exhaustive searches of authentic materials also require knowledge of morphologically related variants not listed in the dictionary. These issues all pose difficulties for an analyst who wants to create meaningful queries.

Our Solution—To help analysts frame their queries in Arabic dialects, CASL researchers are pursuing three strands of research: (1) dictionary enhancements, (2) query analysis that works like Google’s *Did You Mean...?* tool, correcting potential spelling errors, and (3) analyzing and generating morphologically related variants. With a dialect-based Arabic query analysis tool and improved dictionary support, analysts will be able to find relevant material more quickly.

Products
To assist analysts, CASL researchers are developing a series of tools to assess and repair queries, including these previously developed tools:

- An improved Arabic dictionary—The dictionary provides expanded and corrected entries in the forms needed for search of both text and speech, including multiple phonetic representations, native orthography, Buckwalter transliteration, and Standard Arabic Technical Transliteration System (SATTS).
- A spell correction tool—When users are unsure of a term’s proper spelling, they can find the term by entering an approximate spelling into CASL’s *Did You Mean?* spell correction tool.

We are currently developing this tool:
- A morphological query expansion tool—An extension of the spell correction tool, this tool will identify the citation form of the queried word and generate a list of morphological variants.

Report
- Supporting Iraqi Arabic speech search: An updated and improved lexicon helps analysts effectively search for Arabic terms (October 2008)

For More Information
For more information about this project or for a copy of the report, contact us:

U.S. Government Point of Contact
David Cox
Government Technical Director at CASL
University of Maryland Center for Advanced Study of Language
(301) 226-8970 | dcox@casl.umd.edu | www.casl.umd.edu

CASL Principal Investigator
C. Anton Rytting, PhD
Assistant Research Scientist
University of Maryland Center for Advanced Study of Language
(301) 226-8883 | crytting@casl.umd.edu | www.casl.umd.edu