CHALLENGE

Working with Arabic and Persian languages is challenging due to regional variations in spelling, pronunciation, and vocabulary, as well as the presence of embedded cultural content. Navigating the available language reference material can be difficult and time consuming for language analysts who need the information at their fingertips. CASL researchers are working to provide analysts with a better understanding of Arabic and Persian by creating innovative tools to help them quickly and effectively find the relevant reference material they need.

SOLUTION

CASL is supporting Arabic and Persian language analysts by providing new lexical resources—and new ways to search the available information. By focusing on common errors made by native English speakers who are learning these languages, CASL is working to provide analysts with the following tools to help them better understand Arabic and Persian communications:

- **Persian Did you mean…? (DYM) spell correction tool.** CASL's DYM tool for Persian allows users to find dictionary words with misspelled queries. In addition, a computational morphology component is coupled with the DYM tool to allow users to find inflected word forms (even if misspelled).

- **Improved Arabic DYM search results.** CASL's study of auditory perception has confirmed that certain phonological features found in Arabic but not in English cause the most confusion for Arabic learners. The study also indicates that the position of a sound within a word affects a learner's potential confusion. CASL is investigating the potential for a new DYM module that takes positions of errors within words into consideration.

- **Yemeni dictionary supplement.** CASL is further developing its new dictionary supplement to expand lexical coverage of Yemeni dialects of Arabic. This resource relies on harvesting a corpus of current Yemeni usage from the Web. CASL is eliciting native speaker feedback to select and translate example sentences and new lexical items from this corpus.

EXPANDING YEMENI VOCABULARY

CASL is expanding the vocabulary found in its new Yemeni dictionary supplement. The following terms are some of the newest additions:

- طنفَس / نَقْطَنَس (Car lift or jack)
- حُرَوَاء / حَرَاوَة (Wedding)
- حُرَاوَا (Groom)
- رُوُاس (Headache)
- سنوُس (Chains, fetters)
- صُنُفَ (Type of thick bread)

More at [www.casl.umd.edu/search](http://www.casl.umd.edu/search)