CHALLENGE

Accurate language resources, such as grammars and dictionaries, speed the work of government language analysts, especially junior and recently cross-trained analysts. Unfortunately, reliable resources may not always exist, and the usefulness of those that do may be limited by complex word formation rules, dialectal differences, hard-to-hear sounds, and variations in spelling, pronunciation, and transcription.

SOLUTION

CASL facilitates analysis in strategic languages by correcting and enhancing existing electronic dictionaries and building replicable software tools to help users find words in dictionaries, gazetteers, and name lists. Some of the tools that can help analysts process language data more quickly and accurately include:

- Enhanced dictionaries and spell correctors for Pashto and Urdu: Analysts can access these electronic dictionaries in CASL's Did You Mean...? fuzzy lookup tool.
- Morphological parsers for Bengali, Urdu, and Pashto: Parsers provide the dictionary forms of complex words and are designed to prevent software obsolescence.
- Transcription system to enable Pashto-Urdu multi-language search: Users will be able to simultaneously search related languages with a single set of letters.
- Language familiarization resources for Dhivehi: CASL created an introduction to the unique Dhivehi writing system and a short, dictionary-type vocabulary list.
- Basic grammars for Dhivehi and Punjabi: Concise, searchable, and available for immediate distribution, these serve as a base for more comprehensive grammars.
- Morphological grammars for Bengali, Urdu, standard Pashto, and Pashto dialects: CASL works with native speakers, written sources, and audio databases to produce new, modern, comprehensive, and searchable descriptions.

UPCOMING PUBLICATIONS

CASL is currently preparing an eight-volume series of descriptive grammars for publication through De Gruyter Mouton, covering underdocumented world languages such as Bengali, Dhivehi, Punjabi, and Swahili. Written by CASL researchers, the volumes are based on fieldwork and include dialectal information, original example sentences, and transcription of native scripts. The first volume, covering standard and dialectal Pashto, is slated for release in 2014.

More at www.casl.umd.edu/grammars