The University of Maryland Center for Advanced Study of Language (CASL) has developed a number of tools that will be integral to the LanguageNation platform. CASL researchers have created measures of language aptitude that can provide diagnostics to tailor the learning experience (i.e., support adaptivity). CASL’s research on working memory training has demonstrated that focused training of the working memory system can enhance a learner’s attentional control abilities – a skill that is important for language learning. And CASL researchers have developed a number of human language technology (HLT) tools that will support LanguageNation’s web fetching capabilities and enhance language learning activities.

In this report, we first begin by identifying the tools that are currently functional and ready for integration into the LanguageNation platform. Next, we identify a number of tools that will require additional work prior to integration and describe, to the level of detail possible, what steps are required to complete this work and assess how much of this work will be possible within the scope of the current contract. Finally, we conclude by identifying functionality that we envision being developed in future work to further enhance the capabilities of the LanguageNation platform. Wherever possible, we discuss conceptual and/or technical roadblocks that must be overcome.

TOOLS READY FOR INTEGRATION

Two tools should require minimal effort to prepare them for integration with the LanguageNation platform: CASL’s *Did You Mean...?* and CASL Research Online. It is worth noting that, even after successful integration has been achieved, these tools will be developed continually and improved throughout the period of performance, but they will be usable at each stage of development.

**Did You Mean...?**

CASL’s *Did You Mean...?* (DYM) allows users to search an electronic dictionary with results returned as a ranked list based on how closely the string entered by the user matches dictionary entries. Matches and relative rankings are determined computationally using a matrix of non-native English sounds and their likely English confusions. The currently available DYM modules most relevant to LN will be ready to begin integration with the LanguageNation platform as of 30Sept13, and will be delivered in beta form by 31Jan14. An updated version of the tools will be delivered by 30June14. CASL continues to work on extending the DYM tool to
additional languages and more sophisticated cross- and multi-language functionality. The DYM language materials can be used in conjunction with a dictionary display. The legal use of dictionary content is possible if the dictionary is out of copyright; otherwise, the displayer must acquire the requisite rights.

The list of DYM modules anticipated to begin integration by 30Sept13 includes:

- DYM modules for Persian, including:
  - Ranked approximate matches for single-word lexicon queries in Persian, with rankings based on sound-alike spellings and a model of auditory confusions for native English listeners.
  - Morphological analysis of single Persian words (out of context), allowing for morphologically complex queries to be matched to their corresponding citation form in a Persian dictionary.
  - Display of selected fields from the Steingass (1892) *Comprehensive Persian-English Dictionary*.

- DYM modules for Somali, including:
  - Ranked approximate matches for single-word lexicon queries in Somali, with rankings based on a model of auditory confusions for native English listeners.
  - Morphological analysis of single Somali verbs (out of context), allowing for morphologically complex queries in Somali to be matched to their corresponding citation form.
  - CASL is currently using a dictionary that has been provided to CASL by the USG for specific purposes.

**CASL Research Online**

CASL Research Online (CRO) is a web-based platform for administering cognitive tests at CASL. As such, CRO will allow the administration of the aptitude tests and working memory training applications developed by CASL researchers in support of language learning. CASL staff are working under separate funding to enhance CRO’s functionality by the end of FY13 (30Sept13), at which point CRO should be ready for basic integration with the LanguageNation platform. What is meant by basic integration is the following. The CRO platform is hosted on CASL servers and will allow LanguageNation users to complete assessments of language aptitude or working memory training tasks, as directed by the LanguageNation platform. In the first six months of this project, CASL researchers and IT staff will develop an Application Programming Interface (API) for the CRO system and its components, allowing external systems (LanguageNation platform) to interface with CRO in the following ways:

- Activating/deactivating cognitive tests
- Creating participant accounts, with a unique-to-CRO identifier
- Adding/removing participants
- Scheduling cognitive tests
- Requesting participant data

The initial utility of the aptitude testing component of CRO will be limited to providing the test results to users and information for the instructors/guides. Future work (see below) will allow for aptitude scores to feed into tailored training, although this will require both conceptual and technical work by CASL and our LanguageNation partners.

CRO is implemented in Java using the Spring MVC framework and the Hibernate object-relational mapping framework. It runs on the Apache Tomcat webserver and uses a MySQL database to store data. The server delivers HTML/CSS/JavaScript-based tests to study participants and stores their responses as JSON strings in the MySQL database backend. The CRO framework provides a RESTful API that accepts requests over HTTP and returns data in JSON or CSV format. It also provides basic authentication and authorization using the Spring Security library. The CRO platform consists of two applications: the CRO test delivery server (user-facing, for data collection) and the CRO administration tool (researcher tools).

The CRO test delivery server allows researchers to:

- Administer cognitive tests, collecting user activity data and response times
The CRO administration tool allows researchers to:
- set up data collection
- schedule cognitive tests
- monitor participation
- export participant data in CSV or JSON format

Current development work on CRO
- Porting of Hi-LAB tests to CRO (to be completed by 30Sep13)
- Conversion of existing working memory tests into game or game-like modules (ongoing)

REQUIRED ADDITIONAL WORK: CRO PROCESSING TOOL

Additional work is needed for the CRO Processing Tool (CROPT; working title, system still in design) to prepare for integration with the LanguageNation platform. This back-end work is necessary for successful integration of the aptitude measurement capabilities offered by CRO. CROPT will read raw data and process raw data from CRO and other existing and future CASL systems, either via:

- Push (CRO dumps raw data to CROPT)
- Pull (CROPT requests raw data from CRO)

CROPT will provide new, unique user CRO-IDs on request to associate with results coming from CRO or other systems. Funding for the development of CROPT in FY14 and beyond has not yet been made available.