Technical Details

DO0043 Obj. 5.3 ● CDRL A017 ● DID DI-MISC 80508B ● Contract No. H98230-07-D-0175

Individualizing pedagogy in LanguageNation
Recommendations for tailoring training to learner profiles

Medha Tare, PhD, Carrie Bonilla, PhD, Martyn Clark, PhD, Ewa Golonka, PhD, Jared Linck, PhD

The LanguageNation platform aims to make online language learning more effective and efficient by leveraging the resources of Human Language Technology and Second Language Acquisition research. Part of this goal involves tailoring training to learner profiles at the individual level, which may be accomplished more easily in an online environment than in traditional classroom instruction. Further, when training autonomously, learners may not choose the appropriate lessons or modules, relying instead on what they think they might like or need; tailoring can provide the necessary guidance. Tare et al. (2013) outline how to gather information about and build a learner profile which speaks to learner needs, proficiency, and aptitude. This companion report proposes how to use that information to tailor the training more specifically, using examples and illustrations from the LanguageNation platform pilot to discuss current and future possibilities.

RECOMMENDATIONS BASED ON LEARNER APTITUDES

Part of CASL’s ongoing research on aptitude-by-treatment interaction is to establish how foreign language pedagogy can be tailored to match individual differences in cognitive and perceptual abilities as measured by the High Level Language Aptitude Battery (Hi-LAB). Each Hi-LAB sub-score provides unique information about a student’s cognitive aptitude for language learning. Thus, tailoring training to the different scores encompassed by a learner’s aptitude profile should be beneficial to assist the learner in achieving higher learning outcomes. Tailoring training aims to improve cognitive outcomes, such as retention and ability to use the target language, as well as have non-cognitive effects such as lowering levels of anxiety, increasing motivation, and fostering more positive attitudes towards language learning.

Below, we present pedagogical recommendations related to the aptitude constructs which were tested in the Spring 2014 pilot of LanguageNation (Working Memory, Implicit Learning and Explicit Learning/Rote Memory) as well as two others which can be used in the future given more testing time (Inhibitory Control and Perceptual Acuity). Screenshots and examples from the LanguageNation platform as it currently stands are incorporated to explain relevant points at which to insert tailoring. In order for the suggested tailoring to be implemented, training activities would need to be programmed to have access to a learner’s aptitude profile that would specify which adjustments should be presented to the learner. This type of adaptivity would represent a very sophisticated improvement over current forms of automated adaptive training.
Working Memory

The tasks or activities that learners perform in the target language may be adapted according to their aptitude profiles, in particular, those abilities related to Working Memory, or the ability to keep information in memory while updating that information without making mistakes. For those with high Working Memory, learners are better able to handle online tasks in real time (e.g., interpretation, listening comprehension) and more cognitively complex tasks (e.g., tasks about ‘there and then’ rather than ‘here and now,’ tasks on unfamiliar topics, tasks with no planning time, or two-way interactive tasks). Learners with low Working Memory, on the other hand, may need more scaffolding or modifications to such tasks; for example, including additional planning time before a complex task, providing opportunities to engage in pre-task activities to familiarize themselves with the content, or allowing them to engage in shorter, less complex tasks offline before engaging in a more complex, online task. For those with low Working Memory, providing multiple repetitions of language tasks with slight variations will encourage frequent memory store updating for better long-term retention. Task repetition will also be beneficial for those with high Working Memory, although the benefit for them may be additional exposure to the language, which will allow for mastering of fluency or different, and potentially deeper, processing of the input.

Examples of activities that can be tailored based on Working Memory ability include tasks that involve listening to authentic materials for comprehension, gisting, transcription, or dictogloss. Learners with lower Working Memory may benefit from audio input that is slowed down at first and builds up to normal speeds over the course of several lessons. They may also benefit from listening to audio clips multiple times or starting with input with fewer speakers (1-2) compared to many speakers whose speech may overlap. Learners with higher Working Memory can handle native speaker speeds for input more easily. In the screenshot below (Figure 1) from LanguageNation, the learner is asked to fill in the closed caption transcript of the video; learners’ ability to complete this task may be affected by task demands that tax working memory. Having this aptitude information in the learner profile can inform adjustments to the task and input (e.g., whether or not the original clip needs to be slowed down or how many times it should be played).

Figure 1. Fill in Closed Caption

---

1 Dictogloss is an instructional technique in which learners collaborate to recreate a text that was read or played to them.
Implicit Learning

Learners may also differ on their Implicit Learning ability, or ability to pick up on patterns and associations that are not explicitly identified for them. A strong Implicit Learning ability means a learner is able to process and internalize language patterns holistically and in context rather than analytically and out of context. Learners who have high Implicit Learning ability can benefit from exposure to input (extensive listening and reading) in meaning-based tasks. Learners with low Implicit Learning ability can also benefit from these kinds of tasks, but may need more scaffolding (e.g., introduction of a topic/vocabulary/structure before meaning-based tasks) in order to maximize task benefits. For example, in the following activity (Figure 2) from LanguageNation, learners are asked to review flashcards of new vocabulary. Instead of completing the flashcard activity, which may be a waste of time for them, learners who are higher on Implicit Learning ability than Explicit Learning ability will benefit from seeing the new vocabulary in context (i.e., sentences) rather than seeing isolated words. LanguageNation has the capability to show example sentences of the words in context using a concordancer. This option should be automatically implemented for implicit learners. Ideally, the activity would allow the learners to see the target word in multiple sentences where they were focused on meaning and had to infer the meaning of the word from context; this process may suit such learners better than being given the translation pair from the outset.

If a learner is strong in Implicit Learning, it may be best to avoid explicitly discussing grammar rules, instead focusing on using many examples of the target language structure in context; this stands in contrast to an explicit approach for a learner with a high Explicit Learning aptitude. In the screenshot below (Figure 3), learners are being exposed to grammatical rules explicitly; however, a more implicit approach would be to have a learner do a meaning-based task (e.g., matching sentences and pictures) which forces attention to the task at hand, but the learner can also focus on form and notice the relevant gender and singular/plural word endings, thus allowing the learner to pick up on the patterns.
Explicit Learning/Rote Memory

Exposure to input in an original context, such as a newspaper article, is essential for language acquisition, but some learner profiles may indicate potential benefit from occasional decontextualized (e.g., isolated words), or semi-contextualized (e.g., phrases) language input. With regard to the contextualization of input, such as new vocabulary words, learners with high Rote Memory will not only be able to handle decontextualized input, such as target words presented with L1 translations, but may even learn more efficiently when new words are presented in this manner. This method matches the flashcard activity currently used in the LanguageNation system to introduce new vocabulary. In contrast, profiles of learners with low Rote Memory scores should be examined for other strengths, such as in Implicit Learning, which might suggest learning words from context as a more beneficial activity.

As noted above, learners with high Rote Memory should be able to handle lengthy vocabulary lists without the words in context, or use practice techniques such as electronic flashcards in order to remember words. Practicing with flashcards involves repeated retrieval of the meaning of new vocabulary, which has been found to be effective in memorizing new information. However, it should be noted that even though individuals with high Rote Memory ability may find it easy to memorize words out of context (e.g., with flashcards), this may or may not be effective for building the kind of lexical knowledge that will be useful for the ultimate goals of the learner. For those with low Rote Memory, it would be helpful to first be trained in vocabulary learning strategies that would help them to memorize a new word (e.g., connect the word to its synonyms and antonyms, imagine a word’s meaning, build associations for new vocabulary via tasks in the target language, use visually enhanced input, or study a word with a pictorial representation of its meaning). After being introduced to some mnemonic strategies, learners with low Rote Memory can benefit more from using flashcards to memorize new vocabulary. In addition, allowing the learners to discover the meaning of the lexical item/form independently by providing multiple contexts will also help those with low Rote Memory. For both types of learners, these associative learning tasks can be used to prime learners’ vocabulary knowledge before completing higher level tasks such as working with Level 3 texts to write reactions, summaries, or discuss with other students.

Rote Memory falls under a category of aptitudes which indicate Explicit Learning abilities. Currently, the LanguageNation activity below (Figure 4) provides feedback with additional information about the word forms related to a target word, which match an explicit deductive learner who can apply rules which are provided. Hi-LAB includes a test which more directly assesses explicit induction, which could be added to the
LanguageNation battery in the future. Explicit inductive learners have a strength in being able to determine patterns, such as grammatical constructions from examples in the target language. Thus, additional training on word order patterns or verb paradigms can also be conducted using an explicit inductive, rather than deductive, task.

![Figure 4. Fill in the Blank correct answer with additional information presented](image)

**Inhibitory Control**

All foreign language learners, especially at higher levels, need to be able to handle L2 input that is highly variable due to sociolinguistic factors such as different speakers’ accents, dialects, ages, or gender. Learners with high Inhibitory Control ability should be able to handle more variety, while those with a low ability may need additional exposure to the same input, or, alternatively, focus on one accent or dialect initially with an incremental increase in the variety of speech. Thus, as in the activity shown above for filling in the video captions (Figure 1), listening to speakers with varied accents (e.g., European vs. Brazilian Portuguese) may be more difficult for learners with lower Inhibitory Control ability. These learners may benefit from less varied exposure earlier in their training. In other words, they should become thoroughly familiar with one accent, perhaps even one speaker, before moving on to input from additional speakers within that accent and eventually to input from the range of accents that would ultimately be encountered when using the second language.

Inhibitory Control may be especially relevant for cross-training between closely-related languages: those with low Inhibitory Control may have more difficulties with interference (negative transfer) of features that are similar in both languages but not identical, or features that are different and assumed by the learner to be identical, while those with high Inhibitory Control are better able to handle L1 or L2 influences because they will be better able to suppress the L1 and any other previously learned languages and focus on the target language. These issues may extend to differences/similarities in an L2 and L3 as in the case of LanguageNation learners who have Spanish as an L2 and are learning Portuguese. Thus, taking into account Inhibitory Control scores will allow tailoring toward this potential learner difficulty. For features that are similar or identical in both languages and can lead to positive transfer, suppressing L1 or any other previously learned language is not necessary. On the contrary, learners can benefit from the previously acquired knowledge and learn the new language faster because they are already familiar with some linguistic concepts.
For those with low Inhibitory Control, it is recommended that only the L2 is used for communication and activities in order to maintain consistent cognitive activation of the L2 and reduce competition from the L1. Learners with high ability may be able to handle switching among multiple languages to a greater degree. In other words, these learners should be able to hold metalinguistic discussions in the L1 without it negatively impacting their ability to “stay” in the L2 during activities. To that end, activities such as that in Figure 5 which has the learner stay in the target language to reorder sentences into a coherent text may be better suited for a learner with lower Inhibitory Control and the activity in Figure 6 which involves translating and going back and forth between languages may be more suitable for a learner with higher Inhibitory Control.

Figure 5. Reorder the sentence

Figure 6. Correct the Machine Translation
Perceptual Acuity

Perceptual Acuity refers to learners’ ability to perceive sound contrasts that are not in their native language. For learners with low Perceptual Acuity ability, it may be helpful to draw the learners’ attention to the ways in which difficult sounds in the foreign language are produced. For instance, a lesson may point out similarities and differences of the L2 phoneme pronunciation compared to the native language or other known languages and explicitly point out differences in their articulation, thus putting a stronger focus on having the participant articulate the sound, rather than on hearing the difference. Another strategy, as in the activity presented below (Figure 7), is to encourage the students to practice typing what they hear to enforce the connection between the auditory and visual signal, promote phonological categorization, and help the teacher or coach identify any confusion. This may be particularly useful for words with difficult or confusable sounds, helping train learners who need extra practice. Another way to help learners with low Perceptual Acuity may be to provide visual cues (e.g., color coded letters) to learners on specific sound contrasts during activities to help reinforce the salience of those distinctions.

![Figure 7. Flashcards (transcription)](image)

When presenting new material (vocabulary, grammar, content), it would be best to avoid exposing learners with low Perceptual Acuity to “noisy” audio files (e.g., distracting or obscuring background noise, speakers talking over one another); audio files should be automatically tagged on these dimensions if possible. Learners with low Perceptual Acuity ability will likely ultimately require more practice with “noisy” input in order to be able to handle such materials in their jobs, but the use of noisy materials for the purpose of other learning goals (e.g., learning vocabulary or structures) may present too much of a challenge. Similarly, learners with high Perceptual Acuity ability should be able to handle more variety, while those with a low ability may need additional exposure to the same input, or, alternatively, focus on one accent or dialect initially with an incremental increase in the variety of speech.

For those learners with high Perceptual Acuity, they may not need the explicit support that learners with low Perceptual Acuity need. Rather, with sufficient input, they may more easily detect fine grained sound distinctions without further explicit instruction. The key for these learners is to ensure immediate access to floods of input with the relevant sound distinctions that they can pick up on. For those with low Perceptual Acuity ability, including text with audio input may aid learners to process the input, while those with a high Perceptual Acuity ability may not need textual support.
**ADDITIONAL DIMENSIONS**

**Training aptitude**

One question which is often raised with regard to aptitude is whether a learner’s ability can be improved. In recent years, CASL researchers have developed a cognitive training regimen that can enhance an individual’s working memory. In typical implementations, participants spend 30-60 minutes per day for 2-6 weeks, performing a series of cognitive tasks aimed at improving working memory. These tasks can be completed online on a web browser or on mobile devices, such as tablets or smartphones. Successful completion of the cognitive training regimen has been shown to lead to improvements in an individual’s working memory and, critically, to performance on language tasks (e.g., Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2013; but see Sprenger et al., 2013, for a discussion of the limits of cognitive training). In the context of LanguageNation, learners with low working memory could be offered this cognitive training to boost their working memory abilities in advance of engaging with the platform. Because working memory has been shown to be an important ability supporting language processing and language learning (for a meta-analysis, see Linck, Osthus, Koeth, & Bunting, 2013), the hypothesis is that learners with low working memory could complete the cognitive training regimen to boost their working memory abilities, which should also lead to improved language learning (for a review of training studies and discussion of implications for language processing, see Hussey & Novick, 2012).

**Providing feedback on error**

One particular area that may benefit from tailored training is how the system provides feedback on learner error. Feedback on error is a necessary component of successful language learning (Carroll & Swain, 1993; DeKeyser, 1993; for a discussion of online feedback, see Benson et al., 2012; Golonka et al., 2013). The LanguageNation platform currently includes systematic feedback when learners have to indicate comprehension or when learners have to produce language. First, for all activities that do not require the learner to produce any language (i.e., the learner has to indicate comprehension by matching or ordering pieces of information), the system only provides color-coded feedback (green = correct; red = incorrect). For example, in the Matching exercise illustrated in Figure 8, the learner must draw lines matching the topic headings on the left-hand side and the paragraph excerpts on the right-hand side. When a learner clicks on the “check answer” box at the bottom, the lines that are correctly matched turn green, and the lines that are mismatched turn red. Several other activities have this kind of feedback, for example, the Reorder the sentences activity illustrated above in Figure 5.
There are various kinds of feedback given when learners are asked to produce the target language. Activities requiring production include: Correct the machine translation (with or without surrounding context), Fill in closed caption, Flashcards, and Fill in the blank (FITB). In the FITB exercise, the system can recognize when a learner has made a morphological error, a spelling error, or a lexical error, as shown in Figure 9.

In this case, if the learner produces an incorrect form of the word, the system asks the learner to check the form of the word and gives the learner six choices, one of which is the correct answer. In this same exercise, when a learner has produced the correct answer, additional metalinguistic information is provided for the correct word, as illustrated in Figure 10.
Another kind of feedback on learners’ production in the LanguageNation platform is to give the learner the correct answer with all or most of the letters replaced by stars. In some of the exercises, a character or two is revealed if the learner continues to make errors to give the learner hints. This kind of starred feedback is shown in Figure 11.

The effectiveness of LanguageNation’s feedback on learners’ comprehension and production could be related to learners’ aptitudes. CASL researchers have identified three broad categories of feedback on error that relate to learners’ cognitive abilities: Error Identification, Grammatical Explanations, and Recasts. The feedback incorporated in the LanguageNation platform to date includes some features of these three kinds of feedback;
looking forward, there are other features of feedback that could be incorporated and tailored for different learner aptitude profiles.

One type of feedback is Error Identification, which would involve pointing out the presence of an error, rather than providing an overt correction or correct example of some kind. Being able to make use of this feedback requires the learner to recall an utterance and figure out how to correct the error. A learner with strengths in Explicit Learning (Explicit Induction) and Working Memory may do well with this kind of feedback. This kind of feedback is already implemented in various activities in the platform, e.g., FITB and the other production activities. However, in the current platform, the learner generally has visual access to erroneous production. One way to push learners, especially those with strengths in Working Memory, would be for the erroneous utterance to disappear after the learner is told there is an error. This would be an adaptation that would also provide some opportunities to mimic the characteristics of spoken language, i.e., the lack of a visual record to review when one commits an error.

The red/green color coding for errors is an example of an inductive kind of feedback; the learner must figure out why the selection is incorrect with no further information. This kind of feedback may work well for those learners that are strong in inductive learning, either implicit or explicit. Whether a learner with implicit or explicit learning strengths would benefit more from this kind of feedback may also depend on the nature of the task. The Matching exercise (passage comprehension; Figure 8) is meaning-focused and naturally requires less attention to linguistic features, which an implicit learner may perform better at, while in the Reorder the sentences task (Figure 5), the learner is forced to look more carefully at the meanings, forms, and syntactic order of each word, which may require more metalinguistic reflection, a typical strength of an Explicit learner. The starred character feedback (Figure 11) is another inductive kind of feedback; the learner is given starred characters as a hint for the correct word. As some of the letters are revealed, figuring out which word is the correct target word may be suited well for learners with strong Explicit Learning (Explicit Induction). This kind of activity may also be effective as a word learning activity for those with strong Rote Memory, who may benefit from this decontextualized practice with the target vocabulary.

Another common type of feedback is an explicit, overt, grammatical explanation of an error. For example, this might involve pointing out an error specifically, and then providing a rule or other explanation for how to correct it or avoid the error in the future. This kind of feedback may be more effective for learners with strengths in Explicit Learning (Rote Memory and Explicit Induction). The platform currently incorporates a version of this kind of feedback (Figure 10), but the grammatical information provided might be enhanced by allowing the learner the option to click on a link for more information about the error. In the same vein, learners’ abilities to comprehend metalinguistic terms, such as “subjunctive” or “infinitive,” vary depending on many factors; therefore, it might be most helpful to always show examples of the target structure or form along with the grammatical information. For the activities that do not involve learners’ production, for example, in the Choose the correct one activity shown in Figure 12, a learner could also be given some metalinguistic feedback after a certain number of attempts, and/or when the activity is successfully completed.
Finally, recasts are a type of feedback in which the instructor or more proficient partner provides a subtle correction to a learner's error by repeating back something the learner said, but with the correct target language (i.e., fixing the error). This can be effective not only for errors, but also for providing the learner with a more appropriate way of saying something. The learner needs to be able to hold both the original utterance and the recast in the mental workspace and still have some attentional resources available to notice the difference. Learners that are strong in Implicit Learning and Working Memory may especially benefit from this kind of feedback. This kind of feedback is traditionally implemented in a communicative task when a learner and proficient partner are engaged in a meaning-based task. This kind of feedback also commonly occurs in conversation; while recasts are not currently used in the LanguageNation platform, learners communicating with each other through a chat function could be encouraged or instructed to correct each other’s errors by repeating them correctly during synchronous interaction. The ability to deliver implicit feedback such as recasts automatically should also be a priority for the LanguageNation platform.

Language Proficiency

One way that the LanguageNation platform tailors its content toward a learner’s language proficiency is by considering whether target language words are cognates with a learner’s native language, thus making them easier to learn and less targeted in the training; in the case of the Spanish-Portuguese cross-training scenario, the system will also aim to account for the learner’s knowledge of Spanish cognates in order to increase the training’s efficiency by focusing on words unique to Portuguese. Another way to tailor foreign language instruction is to take into account learners’ proficiency in the target language. Global proficiency can be measured by means of the Defense Language Proficiency Test (DLPT) for reading and listening that uses the ILR proficiency scale, or ACTFL-developed tests such as the Oral Proficiency Interview (OPI), Listening Test for Professionals, Reading Test for Professionals, and Writing Proficiency Test that use the ACTFL scale. If for some reason taking proficiency tests is impossible, self-assessment measures in the target language (e.g., can-do statements) can be used to establish the baseline proficiency for learners. This can be accomplished via a needs analysis. An interesting research question that can be answered using LanguageNation data is whether or not students’ proficiency scores on standard tests and self-assessment match initial assessments from the system for how much a student knows.
One way of using proficiency for tailoring within LanguageNation is to ensure that learning materials are at the appropriate level for their stage of development, that is, slightly higher than the learner’s current individual levels so that they learn new information and right at the learners’ levels so that they develop greater fluency. Determining the level of a given text by a human can be very resource-intensive and, as in the case of LanguageNation, not suitable for real-time materials fetching. Rather, the LanguageNation platform needs to be able to automatically rate passages to determine the ILR level of fetched texts. As we discussed earlier (Linck et al., 2014), automatic text leveling for written discourse is available for some high resource languages, such as Portuguese, however, research is needed to expand this capability to low resource languages, such as Somali, and to spoken discourse in any language. The LanguageNation technology is presently capable of performing automatic ILR level detection with written texts in Portuguese and Farsi, though the information is not currently incorporated into the learning platform.

Related to global proficiency is the issue of lexical frequency that needs to be taken into account when selecting vocabulary for instruction (Linck et al., 2014). While very little research has been published on what linguistic features are typical of each proficiency level, there is some research on the relationship between lexical competence and L2 proficiency as measured by the Test of English as a Foreign Language (TOEFL) (Golkar & Yamini, 2007; Meara, 1996; Meara & Jones, 1988; Nizonkiza, 2011). By and large, results from these studies indicate a high correlation between learner’s vocabulary knowledge and TOEFL scores. For instance, in a study of 104 EFL learners at the University of Burundi, the scores on the Word Associates Test and TOEFL correlated significantly and correlations were stronger at lower levels than at higher levels (Nizonkiza, 2011). With respect to the ILR scale, recent work by CASL revealed that the ability to recognize real words at the frequency range between 5 to 50 per million was the biggest differentiating factor between an ILR level 2 speaker and an ILR level 2+/3 speaker in Russian (Golonka et al., 2012). Results from a CASL study (Long et al., 2012) indicated that vocabulary was among the correlates of proficiency that contributed to differentiating between ILR levels 2 and 3 in Russian. Finally, findings from vocabulary studies indicated that the top 2,000 English words typically cover between 80% and 90% of any text and that additional coverage begins to drop rapidly thereafter, with the third 1,000 words adding only about 3% of extra coverage (Nation & Chung, 2009). These findings imply that introducing learners to high-frequency words early on will allow them to better comprehend texts relatively early. With these findings in mind, tailoring in LanguageNation can be done by careful selection of vocabulary to be used in instruction. Specifically, given that high frequency words are a relatively small group of words that occur very often in a language, these words can be introduced to the beginner learners so that, once learned, they can understand as much of the target language texts as possible from the very beginning. Ideally, the system will track learning of the high frequency words to ensure that learners have mastered these and committed them to long-term memory. At higher proficiency levels, instruction can focus on vocabulary from lower frequency bands.

**Learner Language Needs**

In building a learner profile, Tare et al. (2013) also proposed conducting a detailed needs analysis to determine the learner’s language learning goals and the contexts in which he/she would be using the language. This type of interview or questionnaire would yield information about topics or subject areas that would be more aligned with the learner’s ultimate goals. While everyone needs to learn the most frequent target language vocabulary to achieve general proficiency, more advanced and specialized vocabulary could be chosen using topic tracking algorithms. Even providing students with options of topics so that they can choose what interests them more, for example, when completing the Passage Comprehension Matching exercise (Figure 8), may increase motivation and continued use of the platform, thus improving stickiness (Clark et al., 2014).
Assessment

The LanguageNation system aims to assess or challenge learners on its learning targets in order to determine whether a learner can move ahead to the next learning target. The format of an assessment can also match or mismatch a learner’s aptitude profile, and given that the system’s evaluation is essential to learners’ progressing in the learning module, any mismatch could have important consequences. For example, learning of the lexical targets is assessed through flashcard type exercises, which, as we discussed above, are more suited to explicit learners, though these learners should also be challenged to show greater mastery and automatization of learning through more integrated activities. Further, implicit learners may also be learning these words through the exercises, but may not perform well at explicit types of assessment tasks, thereby underestimating their knowledge and limiting their ability to advance through the platform. A more suitable assessment for these learners might be having learners complete a meaning based exercise, where they have to use those target words in context, for example, recognizing target words in a sentence or completing a fill-in-the-blank exercise using a paragraph-long passage.

Further, while it may be pedagogically appropriate to ensure that a particular linguistic concept be presented and assessed before a related concept is introduced, it may not be developmentally reasonable to require “mastery” of that concept as a prerequisite to progress in the system. In other words, some learning targets may have lower expectations of initial success (such as article usage in English), though progress toward mastery can and should be tracked. In addition to evaluation of discrete learning targets, LanguageNation should also strive to provide assessment and feedback on learners’ functional language ability and long-term goals as determined from the language needs analysis. The use of self-assessments and task-based assessments should be investigated for that purpose (Clark, et al., 2013). Employing multiple measures of learner progress can provide a more complete picture of language learning than any single test.

PILOT USER DATA AND PLANNED ANALYSES

Aptitude

Learners enrolled in the Spring 2014 pilot LanguageNation Portuguese training were requested to take three tests from the Hi-LAB battery to test their cognitive aptitudes for Working Memory, Explicit Learning/Rote Memory, and Implicit Learning.

Working Memory was measured using the Running Memory Span task, which asked examinees to recall the last 6 letters from an auditory list of 12-20 letters. This test measures the ability to keep information in memory while updating that information without making mistakes. The test score is the average number of letters correctly recalled across the lists, with higher scores indicating higher working memory.

Explicit Learning/Rote Memory was measured using the Paired Associates task, which asked examinees to study 20 pairs of words, consisting of one English word and one pseudoword, and then recall the pairs during the test trials. This test measures how well a language learner can quickly associate meanings with new, unfamiliar word form. The test score is the total number of correctly recalled pairs, with higher scores reflecting better ability to learn new word pairs.

Implicit Learning was measured using the Serial Reaction Time task, which asked examinees to press a button corresponding to the position of an asterisk in one of four locations at each presentation. This test is designed to measure implicit learning of the underlying sequence, since examinees are not explicitly informed that there is an ordered sequence and are often unaware of any such patterns in the stimuli. The implicit learning score measures the degree of speed-up in response time after repeated presentations of the ordered sequence, relative to a random order; higher scores represent more implicit learning speed-up. This test also yields a processing speed measure which is the mean reaction time from the first random order; lower scores indicate faster processing speed.
Nine learners took part in this testing. Descriptive statistics of their scores are presented in Table 1 and demonstrate their range of ability on these tests. This aptitude information can be leveraged to tailor training using the methods described above.

Table 1. Descriptive statistics of Hi-LAB scores for pilot participants

<table>
<thead>
<tr>
<th>HiLAB test</th>
<th>Minimum</th>
<th>Median</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running Memory Span (out of 6)</td>
<td>0.75</td>
<td>2.45</td>
<td>2.33</td>
<td>3.80</td>
</tr>
<tr>
<td>Paired Associates (out of 20)</td>
<td>1</td>
<td>7</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Serial Reaction Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction Time Speed-up (ms)</td>
<td>-59.75</td>
<td>9.94</td>
<td>7.58</td>
<td>50.88</td>
</tr>
<tr>
<td>Processing Speed (ms)</td>
<td>476.20</td>
<td>573.90</td>
<td>588.20</td>
<td>701.10</td>
</tr>
</tbody>
</table>

Figure 13 shows how the LanguageNation participants’ aptitude scores compare to those of a reference group (n=267) of adult learners aiming for high language proficiency (many of whom have already attained this). Given the small sample size at this point, it is not possible to draw inferences; however, this type of analysis with a larger group of learners might inform how training should be tailored. For example, the processing speed of the LanguageNation participants is slower relative to the reference group, which may suggest that they need more repetitions of input or time to respond in learning exercises.

Figure 13. LanguageNation participants’ scores relative to a reference group of high-level learners, expressed as percentiles

Learner summaries

Many questions can be addressed using learner data from the pilot users of LanguageNation. Recently, two CASL researchers tested out the Portuguese training modules (Participants A & B). We present an initial high level summary of their experiences based on aggregate data records that we received from IBM to give a sense of the types of information that can be gleaned from examining such records. Following that, we suggest potential questions that could be answered with a more extensive analysis of log data.

UNCLASSIFIED
Figure 14 is an example of a learning plan in LanguageNation. The boxes in the plan represent one of three types of learning targets: lexical (specific lemmas), thematic (e.g., colors, days of the week), or grammar concepts (e.g., indefinite articles). Exposure to learning targets can be contingent other learning targets such that the successful completion of a given learning target will allow additional learning targets to become available. Those additional targets may, in turn, be prerequisites for other targets. Figure 14 shows these contingencies. Depending on the target, the “completion” may range from simply being exposed to the learning target (e.g., being presented a short overview of Portuguese pronunciation) to requiring a certain level of success on target “challenges” (activities which can be graded correct or incorrect). Thus, the set of learning targets active for a particular user will depend on their success on targets up to that point. In the current version of the system, two learners with the same plan will eventually see all of the same targets, though the order may differ, as the system usually has several active learning targets from which to choose at any given time.

Because lemmas represent the most frequent type of learning target for the current LanguageNation plan, we focus primarily on the lexical environment of the platform.

Exposure to lexical items

A rich lexical environment can be conducive to language learning (Nation, 2010). In using the system for approximately 10-15 hours, both participants were exposed to a large number of words. In addition to lexical presentation as explicit linguistic targets, such as through flashcards, incidental exposure to additional lexical items in other activities occurs (contributing to the total words seen). For words that are part of the learning plan, the system includes “challenges” (i.e., an activity that can be scored correct or incorrect) to monitor success on those lexical targets. As seen in Table 2, the majority of target words appeared at some point in a challenge, though these words represent a subset of all words actually encountered. In other words, LanguageNation users have the opportunity for the incidental acquisition of lexical items that are not part of the learning plan.
Table 2. Word exposure and challenges

<table>
<thead>
<tr>
<th>Participant</th>
<th>Total words (% challenged)</th>
<th>Target words (% challenged)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>733 (0.62)</td>
<td>470 (0.97)</td>
</tr>
<tr>
<td>B</td>
<td>603 (0.61)</td>
<td>398 (0.92)</td>
</tr>
</tbody>
</table>

Word frequency

Research into lexical learning suggests that focusing on high frequency words is an efficient language learning strategy (Nation & Chung, 2009). The LanguageNation platform relies on a set of system-internal corpora to provide word frequency information which, in turn, influences the lexical items presented in activities. To independently investigate LanguageNation’s lexical environment, we undertook some queries using frequency information from the Corpus do Português (Davis & Ferreira, 2002). For this, we used the 20 million word 1900s corpus, which comprises texts from both Portugal and Brazil across a range of genres. By comparing the words that our researchers encountered during their time using the LanguageNation platform against the words’ frequency in the corpus, we can get a sense of the relative frequency of words across the activities. Because we did not have access to individual word frequency information across the entire corpus, we chose instead to separate words into frequency bands using the frequency per million of words at band boundaries (e.g., 501st most frequent, 1001st most frequent) to determine whether a given word was within that band.

As seen in Figure 15, the LanguageNation learning activities covered words from the highest frequency bands at a much higher rate than the lower frequency bands. In one sense, this is not terribly surprising, as one would expect frequent words to appear more frequently in any given sample of language. Nevertheless, it does provide independent evidence to suggest that learners will receive a good deal of exposure to the most frequent words in the language with only several hours of exposure to the system. This is in line with the pedagogical recommendation of focusing on frequent words before infrequent words for initial learning.
Across targeted words, there was a great deal of overlap between the two CASL researchers in terms of the words to which they were exposed, with over 390 shared lemmas. One of the researchers did receive 79 unique targeted words, though this may be partially attributable to the greater number of words seen overall. However, even though there was considerable overlap between the two learners, this does not mean that the learning experience for each researcher was same.

Because the system tailors the experience based on the user’s success, the number of times that any individual word is encountered may vary between users. As seen in Table 3, the CASL researchers’ experiences varied in terms of average number of times words were shown or challenged, with one researcher being challenged on a single word 23 times.

Table 3. Number of words, challenges, and success rates

<table>
<thead>
<tr>
<th>Participant</th>
<th>Shown words</th>
<th>Shown average (SD)</th>
<th>Challenge average (SD)</th>
<th>Challenge max</th>
<th>Success rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>470</td>
<td>7.32 (10.07)</td>
<td>2.379 (2.43)</td>
<td>23</td>
<td>0.89</td>
</tr>
<tr>
<td>B</td>
<td>398</td>
<td>5.68 (8.77)</td>
<td>1.52 (1.77)</td>
<td>14</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Potential additional analyses

The analyses above give a glimpse of the type of information that can be gleaned from investigating the data captured through the LanguageNation platform. Because the data analyzed for this report were aggregated over the users’ entire experience, there are many types of analyses that could not be performed here which would, however, be possible given the full set of extensive data that is logged in the platform. In particular, the temporal aspect of the LanguageNation experience was missing from the aggregate results; that is, details about the activities and learning targets that the user was exposed to in the order that they occurred. These data would be an area ripe for further investigation because they would give insight into how learners interacted with the system on a moment-by-moment basis across the range of activity types, and could be used to answer questions such as the following:

- Are some activities more conducive to success than others? Does this vary by learning target or learner?
- How does the learning experience differ in terms of the sequence of activities for learners with similar and dissimilar profiles?
- What is the time course of exposures and activities for words that are difficult for a given learner? How does this differ from those words that were easy?
- Are learners with different needs and goals being given learning targets that maximize those needs and goals?

Future analyses of this nature could be conducted with user data from the Spanish-Portuguese cross-training pilot to inform future iterations of the platform.

CONCLUSION

This report provides an overview of the current and future potential for tailoring training towards individual learner profiles on the HLT-enabled adaptive LanguageNation platform. Research into how to best tailor training in order to increase learning will need to be conducted over time, with large sets of learners in order to draw conclusions about which specific tailoring strategies help to improve proficiency. CASL plans to conduct analyses with data from the IBM pilot learners to begin the iterative process of informing more improvements to the LanguageNation platform. Once this tailoring is implemented, it can be evaluated as well and feed back into the system.

REFERENCES


