Learning objectives, methods, and techniques for Persian learners
Attaining & Maintaining Interagency Language Roundtable Level 3

Corey Miller, PhD, Carrie Bonilla, PhD, Evan Jones, MA, Thomas Triebwasser Prado, MA, Ewa Golonka, PhD, Martyn Clark, PhD, Medha Tare, PhD

CONTENTS

Technical Report ..................................................................................................................................... 1
Learning objectives, methods, and techniques for Persian learners ................................................ 1
Attaining & Maintaining Interagency Language Roundtable Level 3 ................................................ 1
Introduction ........................................................................................................................................ 2
Review of literature on acquisition of Farsi and Dari .......................................................................... 2
   General recommendations for course design .................................................................................. 2
   General second language acquisition principles ............................................................................ 5
   Acquisition of Farsi and Dari dialects of Persian ........................................................................ 8
State of the science of Persian instruction in the U.S. ..................................................................... 14
   Tools to rank/grade proficiency in Persian .................................................................................. 14
   CASL tools for Persian ............................................................................................................. 15
   Consultations with experienced teachers and learners ............................................................. 18
   Review of existing Farsi materials .......................................................................................... 19
Discussion: Learning objectives, methods and techniques ............................................................. 25
   Intelligent fetching .................................................................................................................. 26
   Automatic elaborated text generation ...................................................................................... 28
   Intelligent feedback via morphological parsing + DYM spell-correction .................................... 28
   21st century lexical resources ................................................................................................. 29
Conclusion ....................................................................................................................................... 30
References ....................................................................................................................................... 32
Appendix A: Description of Tasks and Targets of the Persian LCP Battery .................................... 36
Appendix B: Sample Grading Breakdown from the DLI Online Diagnostic Assessment .............. 39
Appendix C: Interviews with experienced instructors of Persian .................................................. 45
INTRODUCTION

Attainment and maintenance of Interagency Language Roundtable (ILR) Level 3 in second languages is a well-known challenge for adult language learners. The challenges in reaching such advanced levels of language competency are compounded in less commonly taught, but critically important, languages such as Persian, because these low-density languages tend to be both understudied and under-documented. The purpose of this report is to describe the state of the science of Persian learning and teaching in the US and to recommend learning objectives, methods, and techniques for advanced learners of Persian aiming to maintain Level 3 in reading and listening, with the overarching goal of contributing to the design and content of the LanguageNation Learning Management System (LMS) being developed by IBM and CASL. In this report, which will cover both the Iranian variety of Persian, known as Farsi, and the Afghan variety, known as Dari, we will draw on the following sources to achieve these goals:

- Scientific literature on second language acquisition (SLA) and acquisition of Persian
- CASL corpus linguistics work and development of Persian frequency lists
- CASL work on Linguistic Correlates of Proficiency (LCP) in Persian
- Interviews with experienced Persian learners and teachers
- Systematic review of available Persian materials (textbooks, dictionaries, tools, and other resources)

REVIEW OF LITERATURE ON ACQUISITION OF FARSI AND DARI

Maintaining a level 3 in a foreign language implies a learner has already acquired a high degree of proficiency in the language beyond what even a typical college graduate attains after four years of advanced language and literature study (cf. Long, Gor & Jackson, 2012). Although their language ability will still have many areas for improvement, learners at this level should be able to function in a professional capacity using the language, and ILR Level 3 is often expected of diplomats (Leaver, 2003). In this section, we discuss general recommendations for course design based on SLA literature, as well as general principles of SLA relevant to maintaining high levels of language ability.

**General recommendations for course design**

**Needs analysis**

Effective language programs are those that are tailored to meet the specific needs of program participants and stakeholders. These needs can be at many levels, from identifying the societal needs relevant to developing national language capacity (e.g., Brecht & Rivers, 2005), to selecting appropriate individuals to undergo the training based on their particular aptitude or experience, to collecting diagnostic information on an individual learner to align instruction with learner characteristics (e.g., Cohen, 2003). Identifying needs can be established through a process called needs analysis (see Long, 2005; Tare et al., 2013), the goal of which is to provide information that can help tailor instruction to the learners and the mission based on a careful evaluation of the learners’ real world target language needs.
Needs analysis is defined as a process for the “systematic collection and analysis of all information necessary for defining a defensible curriculum” (Brown, 2009, p. 269). Thus, the first step in the design of any foreign language learning course or materials is to consider what the purposes of the training are and what kind of syllabus will be used to structure the course. Conducting a needs analysis involves questioning the learner(s) for their language learning needs and goals, and researching the kinds of tasks and language the learner will need to be able to accomplish those goals. This research can include observing and/or interviewing people who use the language in the context for which the learners are being trained. In addition, the particular duty requirements of a position (e.g., giving press briefings, creating summary reports from source documents) will provide information on target language tasks. The information collected then drives the content of the course (see Brown, 2009).

Once the clear range of language learning needs and tasks have been identified, the nature of the syllabus should be established. While many traditional language learning materials focus on a grammatical syllabus (e.g., based on starting with simple grammatical concepts, such as subject/verb agreement and moving toward increasingly complex grammatical concepts), alternative, more effective, options for designing syllabi, particularly for learners who need to maintain or enhance an advanced level of language, are based on a learner’s ultimate target language needs. This could take the form of a content-based syllabus (e.g., focused on learning relevant subject matter areas in the target language) or, alternatively, a task-based syllabus (e.g., focused on the accomplishment of specific target language tasks that increase in difficulty throughout the course). In the learning scenario in this report— maintaining a high level of language competency—it will be especially vital to include the kinds of input that involve higher level language processing for advanced learners, such as biographies, plays or television shows, blogs, documentaries, film/book reviews, scientific articles, or newspaper editorials (cf. Ghanoonparvar, 2012; Minuchehr & Mills, 2012).

**Task-based instruction**

It is important to emphasize that it is not only tailoring the content of the course that is crucial, but also incorporating relevant tasks that are directly related to the needs of the learners. Task-based language teaching uses target language tasks, and not inventories of grammatical features, as the basis for syllabus design. For example, if a learner eventually needs to be able to conduct business in the target language, having learners study lists of business-related vocabulary, or reading business-themed newspaper articles with multiple-choice comprehension questions is not sufficient for maximal language acquisition. A learner must engage in business-related tasks (e.g., making a sales call or conducting a meeting) in order to merge their needs with the real-life tasks they may engage in (see Doughty & Long, 2003 for a discussion of task-based approaches in distance learning). At the same time, while the learners must be able to cope with real world target tasks at the end of the instructional sequence and should be given ample opportunity to use language in the context of real world tasks, this does not mean that learners are expected to endlessly repeat real world tasks until they “get it.” In cases where the real world target task might be beyond the current capabilities of the learners, the real world target task is broken down into a set of pedagogical tasks which help scaffold the learner. The important point here is that these pedagogical tasks are derived from real world needs as identified by a careful needs analysis, and are not merely vehicles for the delivery of a grammatically-based syllabus. In the task-based syllabus, drawing learners’ attention to the grammatical features of target language can and does occur, but always in the service of meaningful communication with the goal of developing form/meaning mappings (Doughty & Williams, 1998).

A prototype task-based lesson on giving directions in Korean was developed by Long, Doughty, Kim, Lee, and Lee (2003). A needs analysis with second language learners of Korean revealed a large number of students who intended to visit Korea and would need to understand and possibly give street directions. To prepare learners for this real world target task, a series of nine pedagogical tasks were developed. The first task involved listening to actual samples of target discourse (i.e., street directions given by native Korean speakers in the local context) while the instructor traced the route described on a projected map. For the second task, learners were asked to trace routes on a map themselves, based on increasingly complex authentic direction fragments read by the
This was followed up with another listening task using a more complex map, and several subsequent tasks in which learners worked in pairs to give and follow directions, using transcripts and marked routes at first, but gradually decreasing the availability of such aids. The final tasks involved learners giving directions unaided to the instructor and peers. Learner progress is assessed in terms of how well they can perform the target tasks. Although many language courses include practice with maps and giving directions, of particular importance in this example is that the task was motivated through an analysis of learner needs and the pedagogical tasks were derived from real world target language discourse. Though this particular example was geared for learners at a lower proficiency level, a properly conducted needs analysis would uncover a variety of target tasks requiring advanced levels of proficiency.

SLA research has identified several task features that may influence learner performance on various tasks (e.g., Byrnes, 2002; Long & Norris, 2000; Norris, 2002; Norris et al., 2002; Robinson, 2007; Skehan & Foster, 1999). Some task factors place cognitive or interactional demands on learners, and those demands are hypothesized to affect language learning either positively or negatively. A very useful conceptualization of task design as it relates to language curriculum can be found in Robinson’s Triadic Framework for task classification (Robinson, 2007), in which tasks are classified according to three types of factors: task complexity, task conditions, and task difficulty. “Task difficulty” in this framework refers to the characteristics of the learner, and will vary across individuals given their cognitive ability profiles and other affective factors. Because the syllabus designer is usually only able to control task complexity, Robinson (2010) presents a general framework for task sequencing which focuses specifically on those task complexity variables listed in Table 1. The presence (+) or absence (-) of these variables determines the relative complexity of the task, with the less complex state indicated first for each variable. Thus, for some variables, the presence of the variable makes the task less complex while for others, the absence of the variable makes the task less complex.

Table 1: Task complexity variables adapted from Robinson (2010, p. 257)

<table>
<thead>
<tr>
<th></th>
<th>Task Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thinking</strong></td>
<td>Resource-directing, developmental dimensions</td>
</tr>
<tr>
<td></td>
<td>+/- here and now</td>
</tr>
<tr>
<td></td>
<td>+/- few elements</td>
</tr>
<tr>
<td></td>
<td>+/- spatial reasoning</td>
</tr>
<tr>
<td></td>
<td>+/- causal reasoning</td>
</tr>
<tr>
<td></td>
<td>+/- intentional reasoning</td>
</tr>
<tr>
<td></td>
<td>+/- perspective taking</td>
</tr>
<tr>
<td><strong>Reacting</strong></td>
<td>Resource-dispersing, performative dimensions</td>
</tr>
<tr>
<td></td>
<td>+/- planning time</td>
</tr>
<tr>
<td></td>
<td>+/- prior knowledge</td>
</tr>
<tr>
<td></td>
<td>+/- single task</td>
</tr>
<tr>
<td></td>
<td>+/- task structure</td>
</tr>
<tr>
<td></td>
<td>+/- few steps</td>
</tr>
<tr>
<td></td>
<td>+/- independency of steps</td>
</tr>
</tbody>
</table>

As seen in Table 1, these complexity variables are divided along two dimensions: resource-directing and resource-dispersing. Resource-directing dimensions are those which direct the learner’s attention on the target language linguistic system to meet task demands. For example, talking about events in a shared context (+ here-and-now) will be less demanding than talking about events that happened elsewhere at another time (- here-and-
now); similarly, tasks that don’t require discussing complex reasoning or inferences about the intentional states of other people (not causal reasoning, not intentional reasoning) will be less demanding linguistically than those that do. Resource-dispersing factors, on the other hand, are task characteristics that affect task difficulty but do not necessarily direct learner attention to particular aspects of the language. For example, a task with few steps and sufficient planning time (+ single task, + planning time) will be less demanding than a task with many steps and no planning time.

All of these task complexity factors can potentially be manipulated by instructors or software developers to sequence pedagogic tasks in a manner conducive to language learning. In general, one starts with the least complex version of a task that is reasonable for a given learner. Then, resource-dispersing dimensions of the task are increased. This allows the learner to develop automaticity with their current linguistic resources. Next, the resource-directing dimensions are increased, promoting the restructuring of the learner’s interlanguage through new form/meaning mappings. Principles of input processing (described in a later section) could be particularly relevant here. Note that not all complexity variables will be relevant for all tasks, and, as in the Korean map example above, the ultimate aim is to map pedagogical task demands to real world target task needs in a coherent way (Robinson, 2010).

General second language acquisition principles

As for the content and activities of the course, the maintenance of advanced foreign language abilities necessitates building on the principles that are necessary for initial second language acquisition: that is, sufficient, varied, and comprehensible input, combined with opportunities for output and interaction, and provision of feedback for learners to process and correct errors. The goal for advanced learners is to sharpen their command of the language. Methodological principles for optimization of these elements are listed below, based on Lafford et al. (2007), Doughty and Long (2003), and Chapelle (1998). See also Nielson et al. (2009) for a review of online language course design.

**Input (receptive exposure to the target language)**

- Input should be authentic (culturally, sociolinguistically, and pragmatically) and rich (Lafford et al., 2007). To achieve high levels of proficiency, learners must be able to deal with a wide variety of written and spoken texts.
- Input must be comprehensible (Chappelle, 1998; Doughty & Long, 2003). The goal of language is to convey meaning. As such, the follow principles should be followed:
  - Input should be elaborated, not simplified, to the appropriate development level of the student (Doughty & Long, 2003). Elaboration, rather than simplification, helps to maintain the richness of the language while making it comprehensible.
  - Lexical and grammatical (syntactic, semantic, and morphological) structures can be made salient (e.g., through textual highlighting; Lafford et al., 2007). Learners need to be aware of the language forms used to convey various meanings even if they cannot explain every lexical and structural rule.
  - A variety of resources should be available to help learners understand the input (e.g., dictionaries, glosses, and pictures; Lafford et al., 2007). This helps them become self-directed learners.
  - Multimodal input (video + subtitles, written text + pictures) will aid learners to process input more deeply (Lafford et al., 2007). Multimodal input can increase saliency and comprehensibility.
Output (production of the target language)

- Learners need to produce the target language in various modes (oral and written) multiple times in various contexts as well as in isolation (Chapelle, 1998; Lafford et al., 2007; Swain, 1985). Fluent language use requires practice in using the language.

Interaction (receptive and productive communication in the target language)

- Learners need to interact in the target language to facilitate negotiation for meaning (Chapelle, 1998; Long, 1991), which promotes noticing of target language features – a necessary component for acquisition (Long, 1991; Schmidt, 1990). This interactional adjustment can make input comprehensible.

Feedback (positive or negative feedback to target language performance)

- Learners need feedback on production to assist with noticing of target language features (Doughty & Long, 2003); as well as opportunities to correct errors (Chapelle, 1998; Lafford et al., 2007). Learners need to be able to convey meaning in the language, but achieving high levels of proficiency also demands a requisite level of control over lexical and grammatical structures.

- Teachers can facilitate occasional focus on form to draw learners’ attention to target structures (Doughty & Long, 2003). Activities should focus primarily on meaning, but learners also need to recognize how language forms encode that meaning, especially for more subtle features of the language.

Incorporating these four elements into tasks or activities that learners will engage in throughout a course can take many forms. In the discussion, we will return to these general principles of SLA to make more specific recommendations for the maintenance of advanced levels of the Farsi and Dari dialects of Persian.

Cognitive principles of learning

In addition to SLA principles that are specific to learning foreign languages, it is useful to consider general learning principles, derived from years of research in cognitive psychology and education. Out of the 25 general learning principles identified by a group of 30 experts from different areas of the learning sciences, such as psychologists, science instructors, and representatives from the National Science Foundation, we present four that are particularly relevant to adult L2 learning and supported by empirical evidence (Graesser, Halpern, & Hakel, 2008; Halpern & Hakel, 2003). A variety of activities provided through the LanguageNation platform can be driven by these principles.

Dual Coding Effect

The Dual Coding effect refers to the fact that both the initial encoding of information and the memory for that information are improved when information is presented in different modes (e.g., verbal vs. pictorial), modalities (e.g., auditory vs. visual) or media (e.g., computer vs. paper; Graesser et al., 2008). In a theory first put forward by Allan Paivio, it was postulated that human learning unfolds primarily through verbal and visual associations and these two types of information are processed and stored separately (Paivio, 1969, 1971). Thus, these types of information can serve as an independent representation of a given piece of information, thereby increasing the likelihood of remembering it (Clark & Paivio, 1991; Mayer & Anderson, 1991). Some evidence exists that this principle may be especially helpful for L2 vocabulary learning. For example, Shen (2010) found that dual coding of vocabulary items increased both initial learning performance, and retention performance after a 1-day interval for students learning the meaning and character shapes of abstract Chinese words. In the
context of LanguageNation, this principle can be implemented through the use of text+picture glosses or providing captions for videos.

Desirable Difficulties Effect

The Desirable Difficulties Effect refers to the finding that some activities that make learning initially more difficult can lead to better long-term retention of information (Halpern, 2008; Schmidt & Bjork, 1992). Such difficulties might be instantiated by requiring the learner to search out and organize information, rather than passively receiving information organized by the teacher (Graesser et al., 2008), by alternating tasks, rather than blocking them (Richland, Bjork, Finley, & Linn, 2005; Taylor & Rohrer, 2010), or by presenting new vocabulary in thematic categories such as “beach”, “winter” rather than grouped into categories under a shared hypernym such as “food” or “clothes” (Schneider, Healy, & Bourne, 1998). Although initial learning may be slower and more error prone when such complexities are introduced, long-term retention of the information should be greatly improved. Technology can support this principle by modifying input to make it more challenging (e.g., by omitting key information that the learner must supply).

Testing Effect

The Testing Effect reflects the finding that frequent testing promotes learning (Karpicke & Blunt, 2011; Karpicke & Roediger, 2007). When tests are given frequently, students are encouraged to maintain a study schedule, teachers and students receive important feedback on students’ progress, and students are less likely to forget material (Graesser et al., 2008). In fact, research shows that practice at retrieval is a very important variable in promoting long-term retention of information (Halpern & Hakel, 2003). That is, all practice at retrieving information (regardless whether or not graded by the teacher, or self-tested by students) may help long-term memory. Karpicke and Roediger (2008) found that the attempt to retrieve learned information through testing, rather than passive study, was the critical variable in memory for that information after a delay, as measured by recall of foreign language (FL) word pairs. Such tests should provide the learner with minimal cues and require the learner to generate the answer (e.g., tests of recall, not of recognition; Graesser et al., 2008). This principle can be supported by technologies such as electronic flashcards because the learner is required to attempt recall of a word’s meaning and then receives immediate feedback on his or her accuracy.

Spaced Practice Effect

The Spaced Practice Effect, also known as distributed practice, refers to the finding that learners show higher learning performance and greater long-term retention for material when practice is distributed across several smaller sessions (spaced), rather than completed in a large block (massed; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). In the context of LanguageNation, this principle might be implemented through the use of mobile technologies to encourage practice of material at intervals throughout the day.

High-frequency and domain-relevant vocabulary

When planning the vocabulary component of a well-balanced language course, it is important to consider word frequency and range as the basis for vocabulary selection. Research shows that a learner needs between 95% and 98% coverage of the running words to read a text in English without assistance (Hu & Nation, 2000; Laufer, 1998; Nation & Chung, 2009). One way to get learners to such high level of coverage is to teach them the right words at the right time. High-frequency words with a wide range of use are a relatively small, but critically important, group of words that occur very often in a language. Nation and Chung (2009) argue that because the 2,000 most common English word families1 typically cover between 80% and 90% of the running words in a given text, depending on the type of text and the corpus used to estimate coverage, it is reasonable to teach

---

1 A word family consists of a base word and its inflected and derived forms (Bauer & Nation, 1993).
vocabulary according to frequency of occurrence. Using a corpus of The New York Times, the authors calculated that the first thousand word families from the frequency list covers about 76% of the newspaper’s texts. Then the coverage figures drop rapidly: the second thousand words adds about 8% of coverage, the third and fourth thousands add 3% each, and the sixth thousand adds less than 1%. Taking these findings into consideration, we recommend introducing high-frequency words right from the beginning, but as learners reach higher proficiency, expand vocabulary lists to low-frequency items. Frequency-based vocabulary lists can be created using word frequency dictionaries, if available, or using representative corpora of a given language (see section below on “Corpus Analyses of Persian”). Another group of words that should be selected for instruction is specialized vocabulary (e.g., academic, technical, business, and medical) that meets future language-use needs of the learner, which can be determined through needs analysis. CASL has generated frequency-based and specialized vocabulary lists for Persian from written and spoken corpora that can be used in instruction.

Acquisition of Farsi and Dari dialects of Persian

In this section, we begin with the linguistic characteristics of the Farsi and Dari dialects of Persian, including the writing system they employ. Many of these characteristics make Persian particularly challenging for English-speaking learners. We then discuss the existing literature on the acquisition of Persian by English-speakers, with some attention given to heritage learners.

Typological description of Farsi and Dari

Persian can be genetically classified as a Southwest Iranian language (Windfuhr, 2009). Iranian languages are a part of the Indo-Iranian branch of the Indo-European language family, which includes English. Close relatives of Persian include Kurdish, a Northwest Iranian language, and Pashto, an East Iranian language. Hindi, Punjabi and many other Indian languages are also Indo-Iranian.

Persian is often misconstrued as a close linguistic relative to Arabic because of their common script and overlapping vocabulary. However, Arabic is a Semitic language from the Afro-Asiatic language family and is more closely related genetically to Hebrew than to Persian. Nonetheless, because Arabic is the language of Islam, beginning with the conquest and conversion of the Persians to Islam in the 7th to 9th centuries, the language of the religion began to have profound effects on Persian. As mentioned above, one of the most visible influences from Arabic is the writing system, but in addition to this, many words from the Arabic lexicon have entered the language. By some measures, over half of the lexicon of Persian is borrowed from Arabic (Jazayery, 1970). To a lesser extent, Arabic has affected Persian morphology, such as in the formation of some adverbs (Jazayery, 1970). Adjectives in both languages generally follow the noun they modify (e.g., ‘house red’ is grammatical, not ‘red house’), but both Arabic and Persian superlatives precede the noun they modify, (e.g., ‘biggest pencil’ and not ‘pencil biggest’).

Persian is syntactically referred to as a Subject-Object-Verb (SOV) language, meaning that declarative sentences are most often ordered with the grammatical subject first, followed by the object of the verb and ending with the verb itself. There is some variation in the spoken language and also some flexibility when it comes to adverbial constructions indicating time and place, but SOV word order is the most basic. English, by contrast, is a Subject-Verb-Object (SVO) language. As mentioned previously, nouns precede their modifiers, again differing from English. Persian, unlike many more commonly taught languages, has no grammatical gender and even lacks different words for ‘he’ and ‘she.’ Another unique trait of Persian is that it has retained numeral classifiers, which are appear between the number and the noun it is applied to. There is a generic classifier (ٖ) that can be used with almost any noun, but traditionally, different classes of words (e.g., humans vs. livestock vs. grains and other small round objects, etc.) each have their own classifier. Nouns are not pluralized if modified by a number, whether or not a classifier is used.
Verbs in Persian are of two types: simple verbs consisting of a verb alone, and compound verbs (known as light verb constructions or LVCs) made up of a noun, adjective or preposition (the preverb) and a verb (the light verb). As in English and many other languages, the parts of the verb that indicate the agent of the action show up as suffixes on the verbal stem. Unlike English, however, the direct object of the verb can be expressed either as an explicit and free-standing pronoun or as an attached pronominal form which suffixes onto the inflected verb (i.e., ‘I saw you’ can be either دیدم یو، نورا دیدم نورا with an independent pronoun or دیدم توا with an attached form).

Both simple and compound verbs can have attached object pronouns, but in light verb constructions, the additional information can intervene between the preverb and the verb itself. This means that the verb ‘to enter’ in the phrase ‘I entered the city’ would be rendered as an ezafe construction (ezafe is a linking vowel between compound nouns, noun-adjective and possessive constructions) with an auxiliary verb، واردِ شهرَشدم، and ‘I entered it’ could simply be واردِ شد. These separable phrasal verbs are very flexible and can pose some challenge for learners unfamiliar with their characteristics.

Persian is generally a synthetic language, which means that each individual word can have several morphemes that combine to form the meaning of the word (e.g., ‘cats’ has both ‘cat’ and the plural marker). Some languages, such as Mandarin, are isolating and separate each morpheme into its own word, so ‘cats’ would be split into two words. Others, like Turkish, are agglutinative and in certain situations can combine to form full sentences comprised of only one word. Although it is generally synthetic and generally only has a few morphemes per word, Persian has a highly productive morphology which gives it some level of agglutinative capability. What this means is that parts of words can be combined or recombined in a relatively free manner to create new meanings and these meanings are easily decoded by the highly proficient learner or native speaker. Persian, like German, can often combine nouns to create a complex noun, so ‘disambiguation’ in Persian is a combination of ‘ambiguity’ (ابهام، ‘to remove’) which forms بegriff ‘concept’ (and Klärung ‘clarification’ to achieve Begriffsklärung ‘disambiguation’). This is a technique widely used to build new technical vocabulary and to keep up with scientific jargon. Another example of the agglutinative traits of Persian morphology is a full sentence such as نکوچاندنشان کان. ‘They did not cause them to migrate.’ Such sentences, consisting of a single complex word, are relatively common in the spoken language.

Persian employs a complicated system for showing courtesy and politeness known as تعارف taarof. Beyond the simple distinction that many European languages have for a formal and familiar ‘you,’ Persian also has a system of politeness phrases that are used in certain situations (e.g., along the lines of replying ‘You’re welcome’ when someone says ‘Thank you,’ but far more extensive). These phrases may not have the same meaning as their face value. For example, if an Iranian were to offer someone some of their food with the situationally appropriate equivalent of ‘Help yourself,’ it may not be a true offer and could even be seen as rude to accept because a cultural norm was overlooked. In addition to this, the language itself has distinct vocabulary and styles used in different registers of formality. Generally in Persian, like in Japanese and Korean, a system of honorifics is used; meaning the speaker of lower status (e.g., an employee addressing a supervisor) will use words like ‘sir’ when speaking to a supervisor and ‘servant’ when referring to himself in the same conversation. In conversation with friends and family, normal pronouns would more often be used, but formal speech often employs different vocabulary from colloquial speech. For example, instead of directly saying بنه شما همیشه شما عرض کردم، one might say بنه خدمت شما عرض کردم (lit. ‘The servant (I) presented it for your (formal) service.’), while the speaker of higher social rank or esteem would never have ‘said’ anything, but rather would have been said to have ‘commanded’ (Megerdoomian, 2008, p. 39).
Iranian Persian (Farsi) and Afghan Persian (Dari)

Persian is a pluricentric language, meaning that in different regions, it may be spoken with a different official standard (Miller, Strong, Jones, & Vinson, 2013b), with multiple regional dialects, but there is a common historical written standard (Nawid, 1994, p. 24). The most dominant and widely-spoken dialect of the language is the variety spoken in Iran which is also known as Farsi (Farsi is the Persian word for Persian, like español is the Spanish word for Spanish and français is the French word for French). The dialect of Afghanistan is known as Dari (despite the fact that many Persian-speaking Afghans call their language Farsi) and the Tajik dialect is spoken in Tajikistan, and often written in the Cyrillic alphabet. These are the three major dialects distinguished within the Persian language and each follows the standard dialect of the capital city of its country, that is, Tehran, Kabul, and Dushanbe, respectively. Farsi is spoken natively by approximately 35 million people of the 70 million person population of Iran and similarly, Dari is the native language of about 17 million people out of the 33 million person population of Afghanistan (Windfuhr & Perry, 2012, p. 418). Among native speakers of Farsi and Dari, the two dialects are reported to be highly mutually intelligible (Beeman, 2005, p. 4), but both have clear distinguishing grammatical, phonetic, and lexical characteristics to be described below. However, the differences between Farsi and Dari for learners have been found to be significant necessitating both separate courses or conversion courses between one and the other (Roashan, 2011).

In terms of syntax, spoken Farsi tends to have a highly flexible word order, but Dari tends to be more rigid and a majority of Dari utterances remain verb-final (Nawid, 1994, p. 35). The absolute future tense also varies between the two dialects: in Farsi, an auxiliary verb ‘to want/will’ agrees with the subject of the verb and the main verb invariably follows in the third-person singular past tense form. In spoken Dari, the same auxiliary verb is used, but is always in the third-person singular present tense, with the main verb in past tense and agreeing with the subject (Wahab, 2004, p. 100).

Farsi and Dari also tend to form sentences using modal verbs differently. In Farsi, it is common to use a subordinate clause with its verb in the subjunctive, whereas Dari more commonly uses the modal verb as the final verb in the sentence and the other verb is used in a participle form immediately preceding the inflected modal verb (Nawid 1994, p. 36).

The more colloquial language a speaker uses, the more distinction can be seen between Farsi and Dari, but these are some of the basic and most readily noticeable variations. The two dialects have very similar phonetics, but as can be seen in the example above, there is a lot of variation in vowels. Dari rarely raises [ɒ] to [u] before nasals like Farsi and words ending with a short a in Farsi are generally pronounced [e], but pronounced [a] in Dari (Farhadi, 1955, p. 8; Nawid 1994, p. 29). Dari also retains vowel distinctions that Farsi does not. One example is the long [eː] sound seen above in [meːtɒːnə]. In Farsi, this has merged with [i] and [oː] has similarly merged with [u] (Dari [roːz] ‘day,’ in Farsi is pronounced [ruːz], both written روز) to form a system with only three vowels (Miller 2012). Farsi functionally has three long vowels that can all be written explicitly in the script, whereas Dari has five, two of which, [eː] and [oː], are not distinct in writing. Additionally, unlike Farsi speakers who pronounce both 漕 and漕 as one sound, ʌ/ (with allophones [y] and [ə]; Majidi & Ternes, 1998), Dari speakers pronounce漕 as [q] and漕 as [χ]. The letter漕 has the consonant sound [v] in Farsi and [w] in Dari. Pronunciation in Dari can also vary from Farsi in words with diphthongs and drops initial [h] and final consonants where Farsi does not (Nawid 1994, pp. 31-34).
### Farsi Dari Feature

<table>
<thead>
<tr>
<th>Farsi</th>
<th>Dari</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>گروگان</td>
<td>گراوگن</td>
<td>diphthong retained</td>
</tr>
<tr>
<td>میدان</td>
<td>ماجدون</td>
<td>diphthong retained</td>
</tr>
<tr>
<td>همان</td>
<td>امون</td>
<td>h-dropping and final deletion</td>
</tr>
</tbody>
</table>

On a more general note, Dari intonation is generally more staccato than Farsi, which is widely recognized for its melodic intonations (Nawid 1994, p. 26). Vocabulary differences can cause difficulty in understanding between learners of these two different dialects. Persian as a whole has borrowed heavily from Arabic, but in Iran there is a movement to make the lexicon of the language more Persian. In Afghanistan, Persian is one of two official languages alongside Pashto, which has led to many Pashto words entering everyday Dari speech. Besides borrowings and dearabicization, some words have simply shifted in terms of their semantics. A Farsi speaker who uses the word پچه means ‘child’ but is not specifying the gender, whereas a Dari speaker using the same word is specifically referring to a boy and would use the Arabic word طفل instead for the generic term ‘child.’

### Diglossia in Persian

Diglossia is a situation where two varieties of the same language are spoken in different situations. This generally means a high (H) form (typically used in formal, educational, or official settings) and a low (L) form (characterized by usage among acquaintances, family members, at the market, etc.) are spoken in parallel with one another depending on the social context. In addition, the two varieties are always closely related (Ferguson, 1959). Many languages are diglossic to varying degrees, including English. For example, in Jamaica a form of English known as Jamaican English or Jamaican Patois is spoken in many settings in life as the L variety, with Standard English used as the H form. Arabic is also widely considered diglossic, and mutual intelligibility between many colloquial forms of Arabic and the Modern Standard Arabic taught in schools and used in the media is relatively low (Ryding, 1991). Persian, like these two examples, exhibits traits of diglossia in that the written standard Persian language that is seen in most textbooks and used in formal speeches differs markedly from spoken Persian (Jeremias, 1984; Perry, 2003). We will call the written (H) form “standard” and the spoken (L) form “colloquial.”

Standard is the usual means of written communication and the variety first encountered by most students of Persian because textbooks most often use this form of the language. Standard is almost exclusively used in books, print media, official communications, and other forms of written language. It is only spoken in formal spoken media, official addresses, and in some ritualized forms of language (e.g., in courts). Colloquial differs from standard in many ways. Colloquial exhibits several phonetic properties, such as assimilation, deletion of final consonants in clusters, syllable reduction, and the raising of /u/ to /u/ in nasal contexts (Miller, 2011).

- **Assimilation** [dæst] ‘hand’ → [dæss]
- **Cluster simplification** [tʃem] ‘eye’ → [tʃe]
- **Syllable reduction** [migozæræd] ‘he/she/it is passing’ → [miːɡzæɾe]
- **Raising** [zæbɒn] ‘tongue’ → [zæbûn]

Some alternations also occur in verb conjugations, as can be seen in the above example of syllable reduction. The third person singular present verb ending is generally shortened from /æd/ to [e] and more complex tenses, such as the past perfective, are pronounced almost identically to the simple past tense. These alternations in conjugations and pronunciation are visible even across dialects. For example, Farsi, Dari, and Tajik all shorten the third person singular present verb suffix, similarly to the example above. Dari also has a tendency to simplify consonant clusters and drop final consonants to a greater extent than Farsi (Nawid, 1994, p. 34). Besides this, the syntax in the written language is far more rigid than in the spoken language. In colloquial Farsi, it is relatively common to find prepositional phrases ‘to school’ or ‘from him’ after the verb, even though a
sentence in standard Persian almost always ends in a verb. In certain instances, the direct object or even the subject of a verb can also appear as the last constituent of an utterance in order to express emphasis (Megerdoomian, 2008, p. 42).

Although these might seem like significant differences, standard and colloquial Persian are still highly mutually intelligible. Both are fully functional forms of the language, but one is used predominantly in speech and the other mostly in writing. This often leads to a simplified distinction of calling them ‘formal’ and ‘informal’ Persian, but because both have the capability to be very informal or very formal in parallel ways, this may not be the most accurate way to describe this form of diglossia.

Script

Persian has used many different scripts throughout the centuries, including cuneiform for Old Persian, and the Hebrew and Pahlavi scripts in Middle Persian. The script currently in official use for the Farsi and Dari dialects of Persian is the Arabic script (the Tajik dialect uses a modified Cyrillic alphabet). The Arabic script was adopted for and adapted to Persian when it entered the country as the script used to write the language of Islam: Arabic. The writing system was adopted close to the onset of the ninth century CE, well after the actual Islamic conquest of Persia. The Arabic writing system is written right to left, contains 28 characters and various diacritics, and is technically considered an abjad (Kaye, 1996, p. 747), which generally only explicitly writes the consonants of a word (or in this system, also has characters for long vowels, but not short vowels), whereas an alphabet will have overt representations for both consonants and vowels.

Four consonant characters were added to the system for use in Persian because the sounds do not exist in standard Arabic (پ [p], گ [g], ژ [ʒ], چ [tʃ]). Conversely, Arabic has many more consonant sounds than Persian. Due to the simpler phonology of the language, Persian has one sound for multiple graphemes in several cases (o, h, [z], [s], [t], [tʃ], [r], [r], [q], [q]; Kaye, 1996, p. 746). Persian does have a standardized writing system, but there is variation in some instances (Hashabeiky, 2005). For example, some prefixes and compound words can be written separately even though they are one word, written physically adjacent, but unconnected, or written connected as one continuous word (e.g., ‘he/she goes’ can be written as می‌رود, میرود or می‌رود).

Literature on the acquisition of Farsi and Dari

While a great deal of research has been conducted on the acquisition of commonly taught languages in the US, research on the L2 acquisition of Persian is sparse. Modarresi (2001) discusses the maintenance of Persian from a sociolinguistic standpoint by Iranians living in the US. Expanding beyond the immigrant community, Mohktari (2007) surveyed 166 learners of Persian in the US regarding their backgrounds and language learning strategies and beliefs. The results offer some observations about Persian learners, both heritage and native English speakers, in the US. Participants were beginning (first or second year) students of Persian in 3 universities near Iranian emigrant populations in California and Texas. Most of the participants reported being first language English speakers (45.8%) or heritage learners of Persian (42.8%). Participants reported common motivation of social goals, e.g., getting to know other speakers of Persian and Persian culture better. Many students felt translation and grammar were not very important in learning Persian, which, as the author notes, is interesting because the common method for teaching Persian has typically been grammar-translation. Some strategies they reported as a group to be important were asking an interlocutor for an unknown word, finding different ways to express ideas and guessing in speaking Persian.

Abasi (2012) explored intercultural differences in summarization abilities in a class of Advanced Low (approximately ILR level 2) students of Persian enrolled in a one-year graduate program on special topics in Iranian language and culture. In a dual role as the classroom instructor and language researcher, Abasi discusses
pedagogical techniques throughout a two-week module along with the results of a research study designed to shed light on differences in intercultural rhetoric. The interplay of theme (also known as topic: new information presented in a clause) and rheme (also known as comment: what is said about the theme) differ in Persian and English, which can lead to difficulty in following the information flow in a text (Khedri & Ebrahimi, 2012). The skills covered over the two-week module included higher-level analysis of texts, including identifying main and supporting ideas, graphically representing main and supporting ideas, manipulating texts by removing the main or supporting ideas and collectively or individually discussing and brainstorming the missing ideas. Towards the end of the second week, students had to summarize two editorials covering an international issue regarding Iran’s nuclear program. Students first read a summary in English of the issue, and then they had to produce two summaries in Persian based on two articles written in Persian. One was an original Persian editorial, while the other was originally an editorial from the New York Times that the researcher translated to Persian. Abasi analyzed the students’ drafts and summaries, answers to a reflective survey about the summary writing process, a transcript of classroom discussion of the summary writing, and the researcher’s own notes throughout the teaching module.

Overall, Abasi identified multiple features of the summaries that differed between the original English editorial and the Iranian editorial. The Iranian editorial was shorter than the original English editorial, and generally, students wrote shorter summaries of it, but they took longer to do so, and they committed more lexical and grammatical errors. For instance, regarding the hierarchical structure of ideas, seven of the students differed in how they ordered the events than in the Persian text, while nine of the original English summaries followed the order of the original text exactly. Students also used a higher number of discourse markers in the original Persian text, which the author suggests reflects their attempts to make the text more cohesive, given their general perception that the text was generally unorganized. The author attributes this additional level of text complexity as a positive pedagogical scenario that led to greater task engagement, essential in L2 reading and writing. Additionally, the author notes that the greater number of errors in this task were revealing from a pedagogical perspective to allow him as instructor to see where learners were committing mistakes, while the lesser number of errors in the original English editorial likely reflected the fact that learners spent less time on the summary and had additional time to metalinguistically review their summaries for errors. Abasi led two direct interventions to follow up on the results of the summary writing. First, he brought in a former Iranian journalist exiled in the US to discuss the political culture in Iran, in order to illustrate how cultural differences can influence text cohesion and interpretation. Regarding the summaries more directly, students worked in small groups to discuss strategies they used to deal with the summary writing and ideas for how to handle texts structured in unfamiliar ways. Taken together, these findings illustrate a source of difficulty in comprehending texts for advanced Persian learners with L1 English, as well as some pedagogical techniques to deal with these difficulties in the classroom.

Several studies on Persian language acquisition have looked at heritage learners (Mohktari, 2007; Ramezanzadeh, 2010). While by definition a diverse group, in terms of language acquisition, a heritage learner is generally a person who shows incomplete language acquisition, likely due to having early exposure to a language (in the home setting or via cultural background) that is not the main language of society (Montrul, 2009). Heritage language learners may show greater strengths in phonology and morphosyntax, but weaknesses in vocabulary or other complex grammatical areas that are late-acquired in L1 acquisition (Montrul, 2009). Ramezanzadeh (2010) conducted an ethnographic study of Farsi heritage speakers in the US in order to explore patterns of heritage language maintenance and loss for a minority language often perceived in a negative light due to historical and political differences between the two countries. Megerdoomian (2010) discusses particular linguistic features observed that learners in a Persian for heritage learners program have difficulty with and advocates an explicit inductive approach to teaching these features in the classroom. This approach involves a learner being exposed to examples of the target features and figuring out the rules or patterns to the feature in question. The features Megerdoomian proposes are: Arabic plurals (morphology), bilabial assimilation (phonology/orthography), the ezafe construction (orthography), and light verb constructions (syntax).
Atoofi (2013) also conducted a study of heritage learners in the US, specifically, on the use of markers by the instructors and students to convey affect (e.g., attitude, moods, and emotions) in Persian. The author analyzed thirty-six hours of video recordings and fourteen hours of audio recording in two Persian as a heritage language classes over the course of a year and a half. Overall, the students and instructors used a variety of linguistic strategies to express affect, for example, formal vs. informal pronouns, intonation variances, and repetition. These research articles represent a common characteristic of heritage language research; much has centered on linguistic characteristics of heritage learners or differences between heritage learners and traditional L2 learners, while less focus has been on the effectiveness of L2 instruction for heritage learners vs. L2 learners. Generally though, it is agreed upon that the needs and linguistic competencies of heritage learners are distinct enough to merit targeting them as a group, rather than in a traditional class for L2 speakers (Montrul, 2009). In the context of the LanguageNation LMS, a student’s heritage status should feed into the learner model.

STATE OF THE SCIENCE OF PERSIAN INSTRUCTION IN THE U.S.

In this section, we discuss a review of materials and resources available for Persian learning and teaching. We begin by reviewing tools that can be used to assess proficiency levels in Persian, followed by a discussion of tools CASL has developed for Persian. The last two sections present synopses of interviews conducted with experienced teachers and learners of Persian and a review of available Persian materials and resources.

Tools to rank/grade proficiency in Persian

Linguistic Correlates of Proficiency

CASL researchers are in the process of developing and validating a receptive test of Persian language proficiency targeted at learners between ILR 2 and 3. The goal of this test, part of a larger project called Linguistic Correlates of Proficiency (LCP), is to gather data on different aspects (e.g., linguistic features) of the interlanguage of speakers at varying proficiency levels, in order to identify which features learners of varying levels master and which features remain difficult. Long, Gor and Jackson (2012) presented results of a Russian LCP battery including Russian perception and production tasks that found differences in control of aspects of Russian phonology, morphology, syntax, lexis, and collocations for learners between ILR levels 2, 2+, and 3. Presently, CASL is piloting test batteries for not only Persian, but also Chinese, Spanish, and a receptive Russian battery.

The current Persian battery is in the process of being validated by piloting with L2 learners between ILR 2 – 3 as well as with native speakers (Vafaee et al., 2013). The tests are currently being delivered in a remotely executable format using DMDX (Forster & Forster, 2003), an experiment presentation software program. There are 10 total tasks in the battery, which the test taker completes in four 30-minute chunks. The tasks all involve responding to linguistic stimuli by pressing buttons on a computer keyboard, with the goal of assessing learners’ knowledge of vocabulary, morphology, syntax, phonology, or discourse. The measures of interest are accuracy and, in some tasks, reaction time. The tasks target features of Persian that are typically mastered by, or still posing problems, for English-speaking learners at ILR 2, 2+, and 3. A brief description of the tasks and linguistic targets of the Persian LCP battery can be found in Appendix A.

While the test is still in validation phase, there are multiple structures included in the battery likely to be difficult for advanced learners of Persian with L1 English. For morphology, structures such as plural formation, the accusative morpheme ra, subject-verb agreement, tense and mood (including the subjunctive), copulas, and the causative morpheme, likely remain difficult to acquire. In terms of syntax, light verbs have been noted to be a challenge. As for sound discrimination, phonetic differences between liquids and vowels in Persian and English are targeted in the test battery, as well as stress patterns in varying morphosyntactic contexts. Another particularity of Persian is the phenomenon of ezafe, which can cause difficulties in production, perception and
interpretation. Finally, the test battery includes several tasks targeting vocabulary drawn from varying frequency bands, as well as idiomatic and collocational knowledge.

Defense Language Institute Online Diagnostic Assessment (ODA) Tool

The Defense Language Institute (DLI) Online Diagnostic Assessment (ODA) Tool is designed to assess a learner’s ILR level in listening or reading. Currently, the system has reading and listening tests available for Farsi. The reading assessment for Dari is now running, and the listening assessment is in validation. These tests can be accessed by the public by registering on the website. The assessments consist of relatively lengthy authentic texts or audio clips from various genres, followed by questions on both comprehension and linguistic knowledge. The assessment is semi-adaptive, which means that the quantity of passages and questions learners are given is dependent on how well they are doing (i.e., someone who is more advanced will see or listen to more passages). The system does have a limited ability to “ratchet down” the assessment for lower-level users, but this has a more pronounced effect on the length of the assessment rather than on the complexity of the texts and clips used. In other words, the better a user performs, the longer the assessment will run, but the assessment does not provide the option to save and come back at a later time. Since users have to work through the lower levels in their entirety before they reach higher levels, this draws out the length of the test, which can make it tiring for the user.

In terms of format, the comprehension questions are generally short answer. The linguistic questions have a variety of forms, including multiple choice responses where the user must identify the correct form or grammatical category, short answers where the user must type in the English translation of selected terms and phrases, and questions where the user must identify correct areas within the Persian text (e.g., “Select the object of this verb in the text.”). In some exercises, the assessment prompts the user to translate the Persian word into English, but because some words can have multiple translations, ODA always gives the target both isolated and in context to allow the user to translate it more accurately. The listening assessment allows users to listen to an authentic text (with any imperfections or distortions it has) two times per question before they are required to answer. A modified, rerecorded version is also provided with no background interference for users who want to use it for their own improvement, but this does not count towards the learner’s final score on the assessment.

The breakdown provided upon completing an assessment is fairly detailed compared to other online proficiency assessments. It provides an estimated ILR level based on the user’s performance, the user’s next target level, and goes on to score specific areas such as comprehension of main ideas and supporting ideas, vocabulary, performance within the subject areas of Geography, Military, and Culture, comprehension of specific structural (grammatical) elements, etc. (See Appendix B for a sample breakdown). What it does not include is a more granular breakdown of user error or the ability to see a scored version of the test, which is one of the main complaints about this tool because the user does not know which problem areas to focus on to improve. Another major issue is that because ODA scores answers automatically, responses requiring user input (i.e. short answer questions, translations, etc.) sometimes can be marked wrong when they are accurate, but use vocabulary that is not located in the word bank that the platform searches to grade responses. Although ODA does far better than other automated assessments, it seems to consistently underrate learner levels in comparison to more traditionally conducted proficiency tests because of this trait. This is another reason that more detailed feedback of the user’s answers and what ODA grades as correct would be more beneficial for learner progress towards higher levels.

CASL tools for Persian

In this section, we discuss several tools and resources being developed at CASL in support of Persian language learning. Several of these tools and resources will be delivered in FY14 as part of LanguageNation Objective 6.
CASL has made major investigations into Persian vocabulary in the course of two projects on Lexical Learning for DLI. In 2012, CASL created a corpus of contemporary written Persian, using materials from the Hamshahr newspaper and the Bijankhan Corpus (Miller et al., 2012). This corpus was analyzed and transformed into a general frequency-ranked wordlist and 5 topic-specific lists. The general wordlist contains the 5000 most frequent word families, from most frequent to least frequent. A word family (Bauer & Nation, 1993) consists of all the common inflections and derivations of a given word. We increment the count for a word family with each appearance of a member of that family. To give an English example, we would increment the count of the ‘be’ family each time one of the words is, was, were, be, being was encountered. In Farsi, we increment the count of the word family کتاب ‘book’ each time one of the words كتاب، كتابي، كتابها، كتابهاي is encountered. The choice of the representative for the word family is arbitrary, but is the same as is commonly chosen as the citation form in a dictionary—for nouns: singular; for adjectives: positive form; for verbs: infinitive. We built frequency-ranked lists in five topic areas used at DLI (also known as Area Studies Final Learning Objectives [FLOs]): economy and politics, geography, science and technology, culture and society, military. These topic-specific lists are less likely to include common function words and more likely to include content words found in materials about these subject areas.

In 2013, CASL built a corpus of spoken Persian, based on the Linguistic Data Consortium’s (LDC) CALLFRIEND Farsi corpus and a year’s worth of Raytheon BBN’s Broadcast Monitoring System’s (BMS) Farsi transcripts (Miller et al., 2013a). The CALLFRIEND corpus contains human-transcribed telephone conversations between native speakers of Farsi. BMS contains television programs transcribed by automatic speech recognition (ASR). A key aspect of CASL’s work on this corpus was to identify the parts of the corpus that contained the highest quality speech recognition. Both a general frequency-ranked wordlist and 5 topic-specific lists were created on the basis of this data. The topic-specific lists only contained data from the BMS corpus, since the CALLFRIEND corpus did not have enough topic-specific content. The lists created from the spoken corpus identified word families already covered in the frequency lists derived from the written corpus and allow users to compare spoken and written ranks for particular word families.

For the LanguageNation project, we are combining the spoken and written corpora and have created additional corpora for Blogs, Fiction and Dari. We are building a general frequency list that reflects all of our corpora. The idea is that highly ranked words will be truly well-represented across genres and styles. The frequency list can be used to assess vocabulary knowledge at different skill levels. In fact, an early version of the frequency list was used in the creation of the vocabulary portion of the LCP.

Words in context / Concordance

As discussed above, the corpus analysis we have done to date focuses on word families. Persian, like English and many other languages, has words with multiple senses. These words in some cases have distinct etymologies, or are homonymous; in other cases, they are historical extensions of existing meanings, or are polysemous. In order to explore the uses of words in context, we have built a concordancer—a tool that enables us to see various words in the sentential context in which they appear.

English dictionaries, such as Merriam-Webster, often separate entries that have the same spelling, known as homographs, yet separate etymologies (or pronunciations or parts of speech). For example, the word bear meaning ‘animal’ and the word bear meaning ‘withstand’ come from different Indo-European roots. In contrast, a word can have one etymology, yet have vastly disparate meanings that have evolved over time, appearing as different senses within a particular dictionary entry. An example is the word tongue, whose definitions range from ‘part of the mouth’ to ‘language’.
Persian has words like ﻫﯿﺮ that can mean ‘lion’ or ‘milk,’ with separate Middle Persian etymologies. In addition there are words like زﺑﺎن which can mean ‘tongue’ or ‘language’ which come from the same Middle Persian root. We are hoping that work with a concordancer will assist learners in the identification of different senses of a given word. Also, this will help us to tag data that will ultimately be in service to automatic word sense disambiguation algorithms (e.g., Yarowsky, 1995). The ability to automatically tag senses will allow us to create more accurate frequency lists where homographs are counted separately.

Did You Mean?

CASL has developed “Did You Mean?” (DYM) technology for several languages including Farsi. A DYM contains a confusion matrix developed for a particular use-case. The use-case CASL has developed for Farsi assumes English-language speakers listening to Farsi. In this scenario, learners may well confuse several of the letters that sound the same, such as س and ص /s/ and ﺕ and /t/. In addition, certain Farsi sounds such as ﻋ may be unfamiliar to English speakers, so create confusion among letters with this and similar sounds such as ﻗ and ﺑ. Finally, letters such as ﻋ may be silent, further adding to spelling difficulties. The confusion matrix allows for confusions to be weighted, so more likely confusions would be at one end of the scale and unlikely confusions would be at the other end. This allows the possibilities to be presented to the user from most to least likely.

CASL has employed the DYM to facilitate dictionary lookup. In the use-case described here, a learner hears a Farsi word and wishes to look it up in a dictionary to find its meaning. Using a dictionary equipped with a DYM, the learner can type a best-guess spelling and the DYM will present the most likely dictionary entries. The learner can then explore these to find which one is most likely given the context in which the word was encountered. Thus the DYM provides an error-tolerant means for a learner to look up a word whose spelling is unknown.

Morphological Analysis and Generation

CASL has developed a morphological analyzer and generator for Persian. When used in analysis mode, the tool can provide the morphological analysis of an inflected word. For example, given the word ﺑ ﻫ ﻣ ر ﻓ ﻧ ‘I am going,’ the analyzer will report that this is the first person singular present of the verb ﺑ ﻫ ﻓ ﻧ ‘to go.’ When used in generation mode, the tool can provide any individual form or forms of a word, given its citation form. For example, given the infinitive ﺑ ﻫ ﻓ ﻧ ‘to go,’ the generator can provide the first person past tense ﺑ ﻫ ﻗ ﻧ ‘I went’ or the second person plural subjunctive ﺑ ﻫ ﺑ ﺑ .

These tools may be employed in exercises dealing with the interpretation and production of morphology. CASL has incorporated them into the DYM tool in a tool called Morphologically Aware Did You Mean (MADYM). Like the DYM, MADYM is incorporated into error-tolerant dictionary lookup. When using MADYM, a learner does not have to worry about identifying the citation form of a word. Rather, the learner can enter a best-guess spelling for any morphological form of a word and the MADYM will use the analyzer to identify possible citation forms.

Dialect Resources

CASL is examining 18 dialects and languages spoken in Iran. These fall on a spectrum from essentially accents of Persian, such as Shirazi, to members of distinct language families, such as Armenian. Because they are spoken in Iran, whose national language is Persian, all of the dialects and languages are subject to code-mixing or code-switching; that is, a given utterance may have both dialect and standard Persian parts. CASL is
producing a sketch of each dialect and language, with attention to its particular vocabulary, pronunciation, morphology and syntax. In addition, CASL is collecting audio and text samples of the dialect, many of which feature code-mixing. Finally, CASL is collecting audio materials in which native speakers of several of the dialects are speaking Persian. In many cases, the Persian of these speakers is accented due to the fact that their native language or dialect is not standard Persian. The particular non-native features present in such accents are described.

Consultations with experienced teachers and learners

As mentioned previously, conducting a needs analysis can provide information on job tasks and learner needs that can help inform course design. Though not a comprehensive needs analysis, we provide some insight here from experienced Persian instructors and learners.

Instructors

In the process of compiling this report, we interviewed two experienced Persian instructors to gain insight into how they approach instruction. Both are native speakers of the language and have each taught Persian for many years in various settings in Iran and North America. Both instructors have advanced degrees in applied linguistics. The interviews were conducted using a semi-structured set of interview questions, and focused on topics such as: aspects of Persian that are difficult to learn, teaching the written versus spoken language, heritage learners, recommended resources, and learner motivation. (See Appendix C for detailed interview notes.)

According to our experts, many learners in their courses come to Persian for a variety of social or career goals, but in the case of some heritage learners, not necessarily of their own volition. As with the previous research on Persian acquisition described earlier (Megerdoomian, 2010), both instructors mentioned the *ezafé* construction as being problematic for learners. Other issues mentioned included the subjunctive, spelling difficulties for those without knowledge of Arabic, aspects of phonetics and intonation, and discourse differences (described in detail in Abasi, 2012). The distinction between the written and spoken language was also highlighted, with both instructors suggesting that the use of written dialogues can potentially create confusion for the learners, as they conflate spoken language and written form. This can be a particular problem for heritage learners, who are generally more developed in oral skills and may be tempted to write how they speak. The instructors recommend extensive exposure to a rich variety of written and spoken texts to deal with this issue.

The instructors recommend additional language practice outside of the class, and typically assign homework activities to recycle and consolidate material covered in class or to expose learners to additional cultural and linguistic input (e.g., through episodes of Iranian TV shows). Homework examples included oral assignments in which learners record themselves and post the recordings to an LMS, highlighting the differences between written and spoken modalities. The instructors note that homework tasks need to be appropriate to the level of the learners. Both instructors were familiar with many of the Persian resources (e.g., textbooks, dictionaries, websites) reviewed for this report. They noted that most currently available textbooks fall short in one or more areas, such as out-of-date materials, lack of authentic input, or generally poor pedagogical quality. In addition, they believe that learners at the ILR 2 – 3 level require materials with a greater level of tailoring than are available from any single source.

Learners

In order to gain insight into most common challenges that native English speakers learning Persian might encounter, we interviewed three learners of Persian, who currently use their Persian language skills for work. The interviews were conducted using a semi-structured format and covered a range of topics such as most difficult features of Persian to learn, resources used, and interactions with heritage learners. See Appendix D for the list of questions and detailed interview notes.
All three learners came to Persian via different paths. All have done graduate work and spent some time in a Persian Flagship program. One learner, whose language study included Farsi, Tajik, and Dari, first took Persian language college courses and participated in two long-term immersions in Tajikistan. The second learner first lived in Tajikistan, where he did not take any formal classes, and only upon his return to the US, enrolled in a graduate program in Persian. The third learner studied Persian at DLI and later enrolled in an undergraduate Persian program as a part-time student.

Even though their paths were different, they all mentioned similar features of language as most difficult to acquire. In particular, all three learners mentioned ٰ (object marker) as causing the most difficulty, and two learners mentioned the prevalence of words with Arabic roots and word order because of syntax variation between the written and spoken languages and between dialects. Other difficult grammar points mentioned were ەزێفە, مگر /mæɡær/ ‘an interrogative word to which a yes answer is anticipated,’ habitual past, light verb constructions, verb conjugations, and the use of pronouns. The learner who acquired dialect first while living abroad also mentioned issues caused by the dialect differences that affected his pronunciation and learning vocabulary because each dialect might use a different word for the same concept. He also mentioned that spelling in the Arabic script was much more difficult for him than spelling in Cyrillic.

The learners discussed a relatively small amount of hard copy textbooks, grammars, or dictionaries. In particular, they mentioned two textbooks they used: Stilo and a Tajik textbook that was used as a tool to explain grammar points that the language partners did not have the background knowledge to explain. One learner mentioned a useful textbook on Arabic constructions used in Persian provided by the Flagship program in Tajikistan as well as Thackston’s grammar that was used on occasion. He also noted that in the past, he needed to work with multiple dictionaries simultaneously because no one dictionary would address all of his needs. The learners listed a variety of online resources they currently use – broadcasts (e.g., Radio Free Europe, Radio Farda, BBC Persian and Tajik), personal learning tools (e.g., GLOSS, Transparent Rapid Rote, www.farsi123.com, Google translate), social media (Facebook, blogs), and movies on Netflix. The learners noted that the language used in BBC articles is highly stylized and written in a Western manner and suggested turning to blogs for more authentic, non-westernized Persian discourse. The learners also noted that Google translate is only good if the learner knows how to use it and that it is not effective for beginner learners. One learner uses Wikipedia as a dictionary for more abstract concepts.

When asked if heritage learners in their classrooms affected their language learning in any way, all three learners responded that they viewed the presence of heritage speakers in the classroom as a positive thing. The heritage learners were helpful in understanding dialectal variations and regional pronunciations; however, sometimes they would provide phrases characteristic to a particular diaspora community and not used in Persian in general.

A common theme in the interviews was assessment for Persian, specifically, the Oral Proficiency Interview. The learners felt that using other Persian dialects or having an accent other than standard Tehran Farsi would have a negative impact on the test results.

**Review of existing Farsi materials**

**Textbooks**

In this section, we present the results of a review of several textbooks for teaching Farsi published between 1963 and 2014. A Persian specialist and researcher in Second Language Acquisition reviewed the textbooks with the purpose of examining the state of Persian language materials in the US. From the Persian perspective, the review encompassed the adequacy, accuracy, and comprehensiveness of the language material. The evaluation
from the SLA side included the key principles of successful language acquisition as discussed in Section 1: ensuring sufficient and varied input, opportunities for meaningful output and interaction, and provisioning of learner feedback. Table 3 sums up the criteria for evaluating the existing textbooks for Farsi. Detailed reviews are found in Appendix E.

Persian is a less commonly taught language in the US with language learning materials mostly reflecting earlier methods of language learning, namely, the grammar-translation method (Minuchehr & Mills, 2012). This method relied on translation of isolated words and phrases, practicing grammatical rules, and learning vocabulary lists. These kinds of approaches are typical in less commonly taught languages, which are often characterized by lack of resources and funding, trained instructors, and institutional history and support, compared to the commonly taught languages in the US of French, Spanish, and German (Gor & Vatz, 2009). Rather, for these languages at least, the last several decades have seen the rise (although still lagging behind worldwide English L2 instruction) of the communicative method, which stresses the importance of meaningful interaction and production in language learning.

Overall, the textbooks reviewed here follow similar patterns: isolated, decontextualized vocabulary presented with English translations, and explicit presentation of grammatical structures. The syllabi tend to be grammatically-based: for example, the Persian basic course by Obolenksy (1963) and Wahab’s (2004) Dari text. The method for learners to produce the language tends to follow a presentation—practice—produce model, i.e., presentation of a grammatical structure, practice with the structure through fill-in-the-blank activities or rote production, sometimes followed by production in more open-ended questions. For instance, in Hillmann (2007), a section on Arabic roots in Persian vocabulary, explicit explanations of patterns for Arabic loanwords and phrases in Persian are first presented with isolated examples in Persian, English translations, and phonetic transcription, followed by several paragraphs of descriptions in English of historical and cultural information. Learners are then asked to match words with Arabic roots and synonyms in Persian then work on translations from Persian to English.

Table 2: Criteria used for evaluating Persian textbooks

<table>
<thead>
<tr>
<th>SLA principles</th>
<th>Considerations for Textbook Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
<td>- Is the target language (TL) presented in isolation (e.g., list form with translations), sentences, or in larger contexts (e.g., paragraphs)?&lt;br&gt;- Is the TL presented in various single or mixed modalities: printed text, video, audio, pictures + text, video + subtitles, audio + text?&lt;br&gt;- Do texts include glosses? Are they translations, or are they given in other forms, e.g., multimedia glosses, multiple choice glosses? Where are they located, (e.g., in the margins, end of text)?&lt;br&gt;- Is the TL authentic, i.e., not just texts created for language learners?&lt;br&gt;○ Have authentic texts been modified in any way, e.g., through simplifying or elaborating?&lt;br&gt;- Does the TL present a variety of language contexts and discourse types appropriate to the learners’ needs?&lt;br&gt;- Is there cultural content provided? Is it current? Does it include both Big C, e.g., iconic and historical cultural elements, and Little c, e.g., elements and features of daily life and popular culture? (Tomalin and Stempelski, 1993)?</td>
</tr>
</tbody>
</table>
Some textbooks, especially for the more advanced learners, e.g., Samareh (1993) and Oliai and Motiee (2010), incorporate more texts as the object of study. For instance, Samareh (1993) presents an authentic text, typically from a news source such as a magazine, followed by explicit follow-up discussion of lexical and grammatical points: words and phrases, phrasal verbs, grammatical structures, discourse markers. After the presentation of the lexical and grammatical structures drawn from the text, there are several exercises for learners to practice. For example, in the first exercise learners are given compound words in Persian and asked to split them up into their component parts. In the second one, they are given examples of explicit unattached object pronouns and asked to attach them to the verb. In the last exercise, they are given two sentences to combine which will require a change in mood to subjunctive with the conjunction that is provided.

Oliai and Motiee (2010) follow a more input-based, communicatively-oriented approach. Lessons start with a vocabulary list with Persian words in isolation and English translations. From there, the lesson switches to all Persian. Before reading a text, learners are asked to complete several pre-reading activities. First, they are given words and asked to write down an antonym and synonyms. They are also given compound words and asked to break down into prefix, root, suffix, and meaning. Then they are asked to write down what they know about several cultural places that will be mentioned in the text. In group work, they are asked to imagine what questions they would ask people with such professions as archaeologist, carpet weaver, and carpet salesperson. Next, learners are instructed to write about a similar product to carpet weaving popular in the US and the process of producing that project or engaging in that profession. Immediately prior to reading the text, learners are given the topic of the text and asked to discuss what they know about this topic. The text itself is pulled from a webpage, but it is not clear whether the texts are genuine or written for non-native speakers. Post reading, learners are asked comprehension questions, mostly based on factual questions. Next, in groups, learners are
instructed to look at pictures of four carpets and say which one they like the best, using word lists with
specialized vocabulary, such as motif and curve and border.

Rather than the mostly synthetic approach found in the Persian texts (i.e., bit by bit language teaching based on
grammatical and structural rules), research on second language acquisition has shown that a more analytic
approach reflects how learners’ interlanguage develops (Long, 2011). In other words, a more analytic approach
advocates focusing on presentation of whole language targeted at learners’ developmental stages, and
intervening with reactive assistance on mapping forms and meaning. With respect to the particular needs of
advanced learners in the maintenance scenario, Leaver and Shekhtman (2002) discuss theoretical and
pedagogical approaches in relation to acquisition of advanced language skills at level 3 and beyond. They note
that the decontextualized nature of the grammar translation method, coupled with translation activities and
grammar and vocabulary exercises, falls short of getting learners to advanced levels in the areas of cultural
competence and oral competencies, due to its nature of focusing on language “usage” not “use” (p. 7). On the
other hand, communicative methods, with focus on oral fluency via communicative tasks, role plays, and
authentic texts, often leave the learner with “underdeveloped precision and formal language proficiency.” Others
have noted that the communicative method falls short in the areas of cultural competency, by its tendency to
generate discussions in discussions of their own norms and traditions, rather than the target culture (Gor & Vatz,
2009).

In general, while Persian language structure and lexicon is extensively documented in the language learning
materials, in terms of pedagogy and second language acquisition, the Persian materials reviewed tend to fall
short on offering learners the kind of genuine, elaborated input necessary for acquisition, and opportunities for
meaningful interaction. The vision of the LanguageNation platform is that it eventually will be fully automated,
adapting to learner needs, ensuring that the content and the activities a learner engages with are in line with
current findings in SLA and pedagogy.

Tools and Supplementary Resources

CASL researchers also reviewed available dictionaries, grammars, and online resources for learning Farsi. The
relevant utility and functionality of these tools as envisioned in the LanguageNation platform are outlined
below. Appendix F has a detailed description of each tool and resource, including web addresses and comments.

Dictionaries

Because LanguageNation will be providing learners access to authentic texts, there will invariably be
vocabulary that is unfamiliar. In order to ensure that learners can get the lexical information they need about
unfamiliar words, such as their meanings and pronunciation, it would be good for the system to provide access
to online dictionaries. In this section, we review some of the best printed and online dictionaries currently
available in the hope that this will inform state-of-the-art lexical access in LanguageNation.

Bateni (2008) provides a thorough review of print monolingual and bilingual Persian dictionaries. The most
comprehensive monolingual Persian dictionary is the 8 volume Farhang-e Bozorg-e Sokhan, edited by Hassan
Anvari (Anvari, 2003). It features Romanized phonetic transcriptions, which are easier for learners to interpret
than the short-vowel diacritics employed by some other monolingual dictionaries. One of the better English-
Persian dictionaries is Puya (Bateni, 2009). Despite its large size and the presence of contemporary vocabulary,
this dictionary does not provide sufficient information to enable the user to distinguish among the Persian
translations for a given English word. This necessitates further consultation in a Persian monolingual or
Persian-English bilingual dictionary to ensure that the proper Persian word is employed in whatever production
exercise or activity the learner is engaged in. For example, for the English word run, Puya gives the following
Persian equivalents: دو، فرار، سفر، سواری ... The learner would need to consult a Persian-English bilingual
dictionary to learn that these are ‘run,’ ‘escape,’ ‘trip’ and ‘ride.’ Laufer and Levitzky-Aviad (2006) observe that this is a common liability among L1-L2 dictionaries, and they propose ways of enhancing them to rectify the situation, as shown in Table 4.

Table 3: Laufer & Levitzky-Aviad's (2006) Bilingual Dictionary Plus

<table>
<thead>
<tr>
<th>L1-L2 translations</th>
<th>L2 information (definitions, examples, etc.) about each translation option</th>
<th>Thesaurus-like information, i.e., semantically related words to each translation option</th>
<th>Additional L1 meanings of the L2 translations</th>
</tr>
</thead>
</table>

Among the Persian-English bilingual dictionaries, we have found the best to be Kimia (Emami, 2010). Despite its relatively compact size, Kimia appears to contain most words learners at the maintenance stage might be searching for, provides clear pronunciations in Romanized phonetic transcriptions, and offers example sentences using the words in context. One of the most comprehensive bilingual Dari dictionaries is Dari-English Dictionary produced by the Center for Afghanistan Studies at the University of Nebraska at Omaha (Neghat, 1993). In addition to Romanized transcriptions for pronunciation, this dictionary provides etymologies, a feature seldom found in Persian dictionaries.

Nahvi (1989) is a specialized dictionary containing Arabic loanwords in Persian. As with Arabic dictionaries, this dictionary is indexed on the usually triliteral roots of these words. Hillmann (2007, pp. i-ii) provides compelling reasons why understanding how Arabic roots and loanwords operate is critical for the advanced language learner:

More than half of the nouns, adjectives and adverbs in Persian come from Arabic. Because Arabic vocabulary in Persian does not exhibit the same patterns as original Persian words do, learning how Arabic loanwords and phrases work in Persian can facilitate learning and remembering a significant portion of Persian vocabulary. Knowing how Arabic loanwords and phrases work makes pronouncing many unfamiliar Persian words easier. Knowing how Arabic loanwords and phrases relate to one another in Persian makes possible educated guesses about the meaning of unfamiliar words in context. Knowing how the Arabic element works in Persian can make the difference between going beyond the threshold vocabulary necessary in general reading and listening comprehension or remaining dependent upon dictionaries.

Online dictionaries offer learners potentially faster lookup than print resources. As discussed above, CASL’s MADYM technology makes lookup possible despite learners’ potential use of misspellings or morphologically complex forms. CASL has implemented this technology using two older Persian-English dictionaries available in digital form, Steingass (1892) and Haim (1936). One of the most comprehensive online Persian monolingual dictionaries is www.vajehyab.com. It provides access to three large, albeit dated, dictionaries: Dehkhoda (1966), Mo’in (1984) and Amid (1964). One of the features of this resource is the inclusion of contextual examples, often stemming from classical Persian poetry. The site www.farsi123.com provides both English-Persian and Persian-English bilingual dictionaries. The learners interviewed found this resource particularly useful because it provides contextual examples. It is not clear what, if any, printed resources this site is based on. Another site, Wordreference.com, has a discussion forum for Indo-Iranian languages, including Persian, where speakers can comment and discuss definitions of words, which is particularly useful for translators and professionals who work in the target language.

Thesauri
As discussed above, thesauri can be used to enhance L1-L2 dictionaries. In addition, they can be employed in vocabulary-building tasks. Fararooy (1998) is a Persian thesaurus based on Roget’s model. FarsNet (Shamsfard et al., 2010) is a Persian version of WordNet (Fellbaum, 2005), a computerized thesaurus.

Reference grammars

Reference grammars can be useful to the advanced learner as a supplemental resource. They provide a means to reference unfamiliar structures encountered in receptive activities or structures that the learner is aware are causing difficulties in oral or written communication. Lambton (1984) is the classic grammar of Persian in English. Originally published in the nineteen fifties, its Persian examples are somewhat dated, but the core of its information remains accurate. Lambton includes sections on colloquial and classical usage as well as practice reading passages that will prove difficult for contemporary learners. One of the highlights of Lambton is a detailed section on the Arabic element in Persian; however, it may prove daunting to learners without prior exposure to Arabic. Thackston (1993) is a more accessible and contemporary reference grammar. It has thorough sections on the colloquial and classical languages. However, Hillmann (2012) criticizes the artificiality of many of Thackston’s Persian examples. As discussed in the section below on Intelligent Fetching, we propose ways that the LanguageNation platform can be used to ensure that authentic materials, curated to the learner’s needs, are used. Yameen (2010) is a grammar of Dari with the original Persian and an English translation on facing pages. Despite the name, this work does not appear to discuss colloquial Dari, so it might just as well be talking about Persian in general.

Verb conjugators

The site http://www.jahanshiri.ir/pvc/pvc.php provides a Persian verb conjugator. It requires that users enter a verb in the infinitive, and then presents a table containing all major verbal inflections in all persons and numbers. http://persian.nmelrc.org/courses/PRS506/PVCList.html provides conjugation information for a set of verbs accessible by list. A useful feature of this latter conjugator is the ability to examine compound verbs (also known as light verb constructions). In addition, this conjugator distinguishes between colloquial and standard versions and provides audio for some forms. Audio is especially helpful for ascertaining the correct stress of conjugated verb forms, which in contrast to most nonverbs, do not have invariable final stress. A liability of both of these conjugators is the inability to enter a conjugated form in order to find the infinitive. This feature could be made possible by incorporation of a morphological analyzer such as CASL has developed.

Speech Technology & Pronunciation

The site http://www.farsireader.com/ provides access to a text-to-speech (TTS) engine for Farsi by Alborz Technologies. Users can enter up to 300 characters of Persian text and hear it spoken by a computerized voice. The voice is intelligible and the system pays attention to unwritten features that affect the pronunciation, such as ezafe. Such a system could be employed to pronounce dictionary entries (see section on 21st Century Lexical Resources below) or read texts aloud. At www.forvo.com learners can hear pronunciations of many Persian words in Farsi and Dari. These are actual recordings by native speakers whose exact geographic origins are available on the site. In many cases there are several utterances of the same word by speakers of different gender, socioeconomic status, geographical origin and age, which allows learners to become accustomed to pronunciation variation. Users can add new words to be pronounced.

The Raytheon BBN Broadcast Monitoring System (BMS) monitors satellite television in both Farsi and Dari. It provides a transcript by means of automatic speech recognition (ASR) and a translation of that transcript using machine translation (MT). Its indexing capabilities enable learners to search for terms of interest in Persian or English. The accuracy of the ASR depends on the type of programming; news shows appear to be much more reliable than children’s programming, for example (Miller et al., 2013a).
Machine Translation

Google Translate provides language detection and translation of Persian into English and vice versa. Users can click on words in the translation to see alternate translations. While this service is certainly not perfect, it often helps learners get a gist of Persian material, or a possible head start when writing Persian (by entering the desired English equivalent).

DISCUSSION: LEARNING OBJECTIVES, METHODS AND TECHNIQUES

In this report, we have endeavored to present the state of the art in Persian instruction for second language learners. This has included a review of the nascent SLA literature related to Persian, as well as interviews with learners and teachers, and an extensive review of printed and online learning and reference materials. In this section, we will propose activities that could be developed in a Learning Management System (LMS).

From the SLA perspective, approaches for Persian instruction should follow the general principles and techniques outlined in the earlier section of the report. That is, building an appropriate curriculum involves conducting an analysis to determine the learners’ specific needs for the language and then using this analysis to create pedagogical tasks that step students closer and closer to mastery of the language needed to fulfill their goals. Tasks should be designed that take into account complexity and sequencing appropriate to the students’ level. Importantly, focus or noticing of grammatical forms should occur in the context of meaningful activities. The following table summarizes these recommendations as relevant to language maintenance instruction.

Table 4: Concise recommendations relevant to maintenance instruction

<table>
<thead>
<tr>
<th>Goal</th>
<th>Recommendation</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provide rich target language input</td>
<td>Select authentic texts</td>
<td>• Choose texts written by and for native speakers;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Make use of target language corpora;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Provide texts from a range of sources</td>
</tr>
<tr>
<td>Provide comprehensible input</td>
<td>Consider student’s current proficiency level when selecting texts</td>
<td>• Select target texts for appropriate text complexity and text difficulty;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Text-level identification tool</td>
</tr>
<tr>
<td>Elaborate texts</td>
<td></td>
<td>• Build in redundancies to the original text (e.g., add synonyms, explanations, paraphrasing);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Incorporate interactive multiple choice glosses or textual + pictorial glosses</td>
</tr>
<tr>
<td>Occasionally focus on form</td>
<td></td>
<td>• Make lexical and grammatical structures salient;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Input enhancement (e.g., highlighting, underlining, graphics, animation, font size/color, audio repetition)</td>
</tr>
<tr>
<td>Utilize resources to aid comprehension</td>
<td></td>
<td>• Dictionaries, glosses, pictures for reading tasks;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Utilize tools such as DYM, parser, concordance;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Electronic gloss/annotation to electronic dictionaries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Grammatical information and conjugators as supplementary tools</td>
</tr>
<tr>
<td>Engage learners in interactive tasks using the target language</td>
<td>Use tasks, not texts, as the unit of analysis</td>
<td>• Use task-based language instruction;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Incorporate simulations, virtual worlds, tutorials.</td>
</tr>
<tr>
<td>Provide opportunities for practice in the target language</td>
<td>Use tasks, not texts, as the unit of analysis</td>
<td>• Engage learners in production, not only comprehension checks of the input, e.g., use open-ended activities</td>
</tr>
</tbody>
</table>
Maximize interaction

- Task-based activities or discussion in synchronous text chat;
- Messaging/discussion board posting in asynchronous communication;
- Virtual reality or collaborative games.

Ensure that learners notice and correct their own errors

Provide negative feedback on error

- Can be computer-generated, peer mediated, instructor/coach mediated, native speaker mediated;
- Can be synchronous (given immediately) or asynchronous (given after some time has lapsed);
- Can be explicit (e.g., you made an error in verb agreement) or implicit (e.g., teacher repeats or models the correct utterance without explicitly discussing the error).

Ensure exposure to vocabulary most likely encountered in the real world

Select for instruction mid-low frequency vocabulary with a wide range of uses

- Develop mid-low frequency vocabulary lists based on corpora.
- Incorporate mid-low frequency vocabulary in activities.

Ensure exposure to vocabulary needed to understand texts on specialized topics

Select for instruction specialized vocabulary geared toward student needs

- Conduct needs analysis to determine students’ future language needs.
- Use target language corpora or dictionaries to create specialized vocabulary lists.
- Incorporate specialized vocabulary in activities.

Facilitate long-term retention of information

Maximize cognitive principles of learning

- Provide information in redundant formats.
- Do not make initial learning too easy.
- Spread activities across different lessons.
- Test frequently by recall of information.

One of the great advantages of the contemporary situation for the advanced language learner is the plethora of authentic in-language material available on the internet. There are two principal things that an LMS can do for the advanced student wishing to exploit the internet for learning purposes: suggest relevant and appropriate materials, and automatically elaborate the text to enable the learner to understand it better. Once target texts, audio or video have been identified, and learners commence interacting with and responding to these materials, intelligent feedback must be provided so that learners are on a trajectory of continual improvement. At all times, advanced lexical resources must be available so that learners can attack and digest complex authentic materials. Each of these potential LMS features will be described below.

**Intelligent fetching**

The amount of material present on the internet is both its greatest advantage and one of its greatest challenges. Finding relevant and appropriate material based on a user’s needs is a problem confronted everyday by search engines like Google. In the language-learning maintenance scenario, the search for materials can depend on any combination of the following factors:

1. Level (e.g., ILR)
2. Topic/Genre/Style
3. Contemporaneity
4. Dialect
5. Richness in relevant general or topic-specific vocabulary
6. Richness in particular morphological forms or syntactic constructions
7. Specific search terms
Ideally, the LMS would first gather information on learners’ language learning needs relevant to the above factors via a needs analysis process (see the section on ‘General recommendations for course design’ above). The LMS system, with possibly varying degrees of user input, would then be able to execute searches, which depend on various forms of human language technology (HLT), such as automatic text-leveling, topic modeling, corpus analysis, morphological and syntactic analysis, and dialect identification, in addition to general principles of web searching. Intelligent fetching will help provide rich target language input to learners. We will examine each of the factors described above with respect to the Persian maintenance scenario treated in this report.

**Level**

In the maintenance scenario, users will need to have materials at their level. They may also want to warm up with materials below their level, or challenge themselves with materials above their level.

**Topic/Genre/Style**

Users may want to find materials in a topic of interest or that is required for their job. Topics may be broad like “economics” or narrow like “Iranian nuclear negotiations.” Both to maintain an advanced level of competency and to be versatile on the job, learners must be exposed to a range of genres, such as technical reports, dialogs and religious exegesis. Learners should also be competent in a range of styles from informal to formal. Due to the diglossic nature of Persian, this is particularly important.

**Contemporaneity**

Advanced language learners need to be able to handle language from different time periods. Farsi exhibits many changes since the 1979 Iranian revolution (Keshavarz, 1988; Modarressi, 1993; Paul, 1999). Even if language learners are only dealing with present-day materials, they will invariably contain references or excerpts from earlier periods in the language, and the learner should be able to interpret them and ideally be able to identify them as marked; e.g., possibly connoting irony, political affiliation, etc.

**Dialect**

As discussed above, Persian has two major varieties written in the Arabic alphabet: Farsi in Iran and Dari in Afghanistan. Learners may want to focus on one or the other of these dialects. In addition, it is possible that learners will want to focus on smaller or nonstandard dialects, such as Shirazi, Esfahani, etc.

**Vocabulary Richness**

Some texts are richer in vocabulary of interest than others. Such “rich texts” may provide a concentrated and efficient way for learners to focus on particular vocabulary. Learners may want to specify certain tranches of vocabulary or topic-specific words they are interested in focusing on. By using the greedy algorithm (e.g., François & Boëffard, 2002), texts particularly dense with such words could be identified.

**Morphological/Syntactic Richness**

Analogously to finding texts rich in particular vocabulary, it would be advantageous for the system to be able to find texts rich in particular morphological or syntactic phenomena. This presupposes the ability to perform morphological syntactic parsing in Persian.

**Specific Search Terms**

In any fetching scenario, the ability to search for particular words would be required. Morphological analysis could be incorporated into this process, to optionally allow for query expansion including morphological variants of the specified words (whether baseforms or not).
Automatic elaborated text generation

Once suitable materials have been identified by such a search, the learner is confronted with a text that may be difficult (e.g., if the ILR level was chosen to be challenging, or if automatic text leveling was not performing optimally, or for many other possible reasons). As discussed earlier, target language input provided to learners should be comprehensible; thus, while reading authentic texts, learners may need help comprehending semantic and syntactic aspects of linguistic input through some type of text modification or the use of resources aiding comprehension. Research has found that *elaborating* text is superior to simplification because, in addition to aiding in comprehension, all the linguistic features learners must acquire remain intact rather than being excised as in simplification (Long & Ross, 2009; Yano, Long, & Ross, 1994). Elaboration can take many forms, but the one we will focus on is to leave the text intact, but to provide unobtrusive supplementary information that enables learners to garner the full meaning of the text.

Nerbonne et al. (2001) describe a web-based system that allows users to click on a word in a foreign-language text and see its morphological analysis, its definition in an online dictionary, and other examples of the word in context. The output of the morphological analysis includes a citation form which enables the word to be found in the online dictionary. The principles described by Nerbonne et al. (2001) are present in customized language-learning (or specialist) websites such as the Perseus Digital Library (http://www.perseus.tufts.edu/hopper/), principally for Ancient Greek and Latin. For Persian, Connie Bobroff has created an elaborated text based on Jamalzade’s story “A Day in Rostamabad of Shemiran” (http://www.lib.washington.edu/static/public/neareast/yekruz/index.html).

While these sites are invaluable, we believe that on-demand automatic text elaboration would enrich the language learning experience enormously. This might be accomplished by extending some kind of automated annotation functionality. For example, a plugin could be developed that would automatically annotate a webpage that has returned from a customized search as described above. The annotated text would look the same as the original text, except each word would be a hyperlink, optionally accessing a morphological analysis, definition (in both monolingual and bilingual dictionaries), and/or words in context from other materials. Extending this idea, an elaborated version of the text could be produced through a similar process in which some subset of words in the text (e.g., those targeted for elaboration by some other process, such as in consultation with the learner model) are replaced with the target word and its synonym or monolingual definition, such that *dwelling* would be replaced with *dwelling* (*house*) in the displayed document. Thus, the elaborated version of the test would contain all of the information in the original text as well as these additional paraphrases. One could also imagine providing increasingly elaborated versions of a text each time a learner clicks on a “help” button or similar. Giving learners choices or providing scaffolding aids when needed or by request (e.g., after failing an attempt or finding a particular task too difficult) supports the Desirable Difficulties Effect discussed above, which refers to the fact that making learning initially more difficult can lead to better long-term retention of information. Ideally, text elaboration or other scaffolding comprehension aids should be provided when the learner needs them.

Intelligent feedback via morphological parsing + DYM spell-correction

Whenever the language learner is required to respond to a query posed by the LMS, whether in the form of a single word or an essay, intelligent feedback can be provided to enable learners to improve their skills. CASL’s Did You Mean? (DYM) technology can propose a ranked list of words corresponding to a possibly incorrect form. The rankings reflect the distance between the learner’s word form and each candidate correction, as determined by use-case-specific confusion matrices. For example, we have developed a confusion matrix for Persian as heard by English speakers that accounts, for example, for confusions arising from the presence in Persian of phonemes that do not exist in English.

If a user selects a particular candidate correction, the LMS can provide feedback based on the particular confusion or confusions that may be responsible for the divergence between the original and corrected form.
For example, if a user proposes the word ﺑﺎگ, the system might suggest the word ﺑﺎغ ‘garden.’ If the user selects this correction, the system could say something to the effect that the learner may have confused the letter گ /g/ with the letter غ /ɣ/, and provide audio samples contrasting the two.

Morphological parsing can be incorporated into this framework as follows: morphological variants of a proposed word form can be suggested that may better fit the morphological or syntactic requirements of a given context. The difference between the learner’s original proposal and the corrected form can be analyzed by the system to provide brief targeted instruction. This type of feedback on error could presumably occur within any otherwise meaning-based activity in which the learner is interacting with the system. The feedback could range from merely providing the correct form, to repeating the sentence with the correct form (potentially highlighted) included, to providing an explicit metalinguistic description of the error (e.g., “You proposed the second person singular form, when the verb (or context) required the use of the second person plural form.”).

21st century lexical resources

The LMS should ideally be outfitted with lexical resources that will promote learning at the maintenance level. Because learners at this level will often be challenged with new words and forms, they need access to quick and efficient lexical resources that will uncover the meaning of unfamiliar words and phrases. These resources should take advantage of contextual information to isolate the most likely relevant senses. Lexical resources in the LMS should be both bi-directionally bilingual (e.g., between Persian and English and vice versa) and monolingual (e.g., all Persian). According to the Dual Coding Effect discussed earlier, providing information in redundant formats (e.g., text and picture) will facilitate long-term retention of that information. Therefore, dictionaries should allow audio (either oral or from recordings) and written access, in order to serve queries emanating from audio/video material and written material. These characteristics and others will be treated in the following sections.

Pronunciation

Dictionaries must provide accurate multidialectal pronunciation information in order to ensure that learners associate the correct pronunciations of a word with its spelling and meaning. Traditionally, dictionaries have employed a variety of sometimes ambiguous and conflicting symbols and diacritics to communicate pronunciation information, and all too often these are baffling to learners and go unused. To remedy this problem, the LMS should provide text-to-speech (TTS) output for each pronunciation. In the case of Persian, this output should be provided by clearly marked buttons for both Farsi and Dari, to allow learners to acquaint themselves with the differences and in order to ensure that they can recognize the word when spoken in either dialect.

Audio/Oral Input

When learners encounter words in an audio context, the dictionary lookup problem is compounded. Traditionally, learners must attempt to transcribe what they heard in order to be able to look a word up. The challenges of Persian spelling (e.g., non-representation of short vowels, existence of multiple letters for the same sound, silent letters) add to the difficulty. While the Persian Did You Mean (DYM) can mitigate this problem, we believe the ultimate solution will be to enable audio lookup. This could have two forms. In one form, the learners can cut the word they are interested from the audio and input this to the dictionary which can use automatic speech recognition (ASR) to look it up. In the other form, the users can speak what they heard into the microphone and ASR can look it up. The latter will be useful in the case where the user cannot easily find or excerpt the word from the audio stream.
Traditionally, dictionary words are indexed by citation forms. Citation forms for different parts of speech vary by language. For example, in French and Spanish, verbs are indexed by their infinitive, in Latin by the first person singular present tense and in Arabic by the third person singular past tense. Learners may not always know the citation form for a given unknown word they wish to look up. To mitigate this problem, a morphological analyzer can be used to transform a given form into the appropriate citation form to enable dictionary lookup.

**Word Senses and Collocations**

Dictionaries vary greatly by the extent to which they indicate the ways in which disparate senses of a given word can be used. Given the state of Persian lexicography (Bateni, 2008), disparate senses of words are too often given as a list with no further contextual information. To remedy this, concordance access can be provided to enable learners to see the various contexts in which a word is actually used. Word Sense Disambiguation (WSD) can be employed to isolate the most likely senses used in a particular context.

A related concept is that of collocations. In Persian, many verbs are expressed using light verb constructions (LVC) that consist of more than one word, possibly with intervening words. Compare the English phrasal verb “wake up,” which can have intervening material, as in “wake me up.” Dictionaries should enable learners to input collocations (possibly inflected).

**Bidirectional bilingual and Monolingual dictionaries**

Even at the advanced level, learners can benefit from bilingual dictionaries. They need access to both Persian-English and English-Persian dictionaries, depending on whether they are involved in a production or perception task. Persian-English are useful when a learner is attempting to read or listen to Persian material, and English-Persian are useful when a learner is attempting to write Persian or respond in-language to a particular query.

Laufer and Levitzky-Aviad (2006) discuss methods of enhancing bilingual dictionaries with monolingual dictionary information. Along with Jenpattarakul (2012), they underline the importance of monolingual dictionaries to the advanced learner. Such dictionaries are often more comprehensive both in available words and senses for a given word. In addition, by immersing the learner in the foreign language, monolingual dictionaries promote learning by restating unknown material using known material, a technique similar to that described above for elaborated text (Long & Ross 2009).

**Synonyms and Vocabulary Building**

Lexical resources should enable learners to improve both vocabulary depth and breadth (Nagy & Herman, 1987), that is, both to deepen their knowledge of words as well as to expand their vocabulary. FarsNet (Shamsfard et al., 2010), a Persian version of WordNet, can be employed to present vocabulary networks to learners, adaptively pruned according to the learner model. These networks will allow learners to explore synonyms, hypernyms, etc., and LanguageNation can provide full lexical information for each word learners desire to learn more about.

**CONCLUSION**

In this report, we have presented the state of the science of Persian learning and teaching in the US and provided recommendations for learning objectives, methods, and techniques for advanced learners of Persian aiming to
maintain ILR Level 3 in reading and listening. These recommendations are grounded in SLA and cognitive principles of learning that are well-established and based on empirical research. We have also presented and evaluated available resources for learners of Persian, including those developed by CASL. The overarching goal of all of these efforts was to contribute to the design of the LanguageNation platform. We have introduced a number of ways in which LanguageNation can be elaborated to address the problems of Persian maintenance. These methods make maximum use of the internet, state-of-the art second language acquisition techniques, and human language technology. One of the hallmarks of this program is the ability for learners to work with the system on their own schedule, with an adaptive system taking the place of the traditional classroom and instructor, resulting in a convenient and economical method for language maintenance. It is our hope that learners equipped with LanguageNation will maintain their Persian at a higher level and in less time than previously existing methods.
REFERENCES


Studia Iranica


Imagery and verbal processes


Spoken Dari.

Sayd, M. (2010). A Brief Study on the Challenges that Farsi to Dari Conversion Poses at DLIFLC. Persian LEARN.


APPENDIX A: DESCRIPTION OF TASKS AND TARGETS OF THE PERSIAN LCP BATTERY

Task 1: Lexical Decision Task (LDT)

In this task, participants must decide if a word heard is a real Persian word or not. Results from this 60-item task assessing learners’ ability to distinguish words and non-words will show the degree to which they have mastered two features of Persian: negatives and agentive nouns, i.e., nouns that indicate the person or thing that performs the action a verb describes (e.g., paint - painter). In Persian, there are different ways to make negatives and agentive nouns, as illustrated in the following table with varying allomorphs in bold:

Table 5: Negatives and Agentives

<table>
<thead>
<tr>
<th>Negatives</th>
<th>Agentives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persian</td>
<td>Phonetic</td>
</tr>
<tr>
<td>بی‌مزه</td>
<td>bimæze</td>
</tr>
<tr>
<td>لا مامد</td>
<td>naomid</td>
</tr>
<tr>
<td>غیر قانونی</td>
<td>yeʃreʃanuni</td>
</tr>
<tr>
<td>ضد آب</td>
<td>zedeæb</td>
</tr>
</tbody>
</table>

The formation of these structures is mostly idiosyncratic; therefore the LDT is more of a test of advanced lexical knowledge.

Task 2: Explicit Task for ezafe

Learners see two-part Persian compounds on the screen, some of which require a connecting vowel, known as ezafe, between the two words, and some which do not. They hear the words pronounced with no ezafe and must decide if what they heard is correct or incorrect. For example, آب گوشه فرنگی /ɑbegowʤefærængi/ ‘tomato juice’ requires ezafe, but آب سبی /absib/ ‘apple juice’ does not. Results from this 40-item task will show the degree to which learners have mastered the use of ezafe, whose use or absence in some may be idiosyncratic. Even advanced learners find the use of ezafe challenging (Megerdoomian, 2010).

Task 3: Explicit Task for Plurals

Learners see Persian singular words on the screen. Then they will hear their plural forms and indicate whether the plural form they heard is correct. Results from this 30-item task will show the degree to which learners have mastered Persian irregular plurals. For example, منظر /maŋæɾ/ ‘view’ turns to مناظر /maŋæɾæɾ/ ‘views’ but چنگل /jangal/ ‘jungle’ cannot turn into چنگلگان /ʃæŋægel/ ‘jungles.’ Learning irregular plurals in Persian, which are borrowed from Arabic, is a challenge for learners, even at the advanced levels.

Task 4: Aural Grammaticality Judgment Task (GJT)

Participants hear a sentence and decide if it is grammatical or not. Results from this 80-item task assessing learners’ ability to identify sentences as grammatical or ungrammatical will show the degree to which they have mastered five features of Persian: the accusative morpheme -ræ, subject-verb agreement, tense and mood (including the subjunctive), and copulas. These five grammatical features pose difficulty for advanced learners.
Task 5: Collocation and Idiom Decision Task

Learners see a collocation, idiom or proverb on the screen with several parts underlined. They must choose if one of the underlined portions is inappropriate, or if the collocation, idiom or proverb is correct as is. Results from this 100-item task will show the degree to which learners have mastered proverbs and idioms in Persian. The appropriate and correct use of proverbs and idiomatic expressions aids fluency and distinguishes advanced from less proficient learners in any language. An example of an English proverb might be:

\[
\text{a bird in the hand is better than two in the bush}
\]

Following is a Persian proverb with almost the same meaning as the above English proverb. In this case, underlined portion 3 contains a word that is synonymous with the word that appears in this proverb, حلوای ‘sweet’.

Task 6: Vocabulary Translation Task

A participant is given a Persian word and must choose between three English words for a translation. The focus of this 100-item task is primarily vocabulary items sampled from frequency bands in a corpus of contemporary written Farsi compiled at the University of Maryland’s Center for Advanced Study of Language (Miller et al., 2012). The corpus combined (i) the Hamshahri Collection, Version 2, containing over a million types and 145 million tokens from articles on a variety of topics published between 1996 and 2007 in a Teheran newspaper, and (ii) the Bijankhan Corpus, containing 76,000 types and 2.4 million tokens, also from Iranian newspapers and other current written texts.

Task 7: Naturalness Task

Learners hear an utterance and are asked to decide whether what they heard sounded natural. Results from this 50-item task will show the degree to which learners have mastered six important features of Persian phonetics and phonology. The features included in this task are:

- Stress (plural): the stress must be on the plural maker morpheme, e.g., /ketabha/ ‘books’
- Stress (superlative): the stress must be on the superlative morpheme, e.g., /bozorgtarin/ ‘the largest’
- Stress (clitic): no movement of stress onto the possessive clitic, e.g., /masheen man/ ‘our car’
- Stress (negation): the stress must be on the negative morpheme, e.g., /bi mazeh/ ‘tasteless’
- Vowels: the quality of Persian round vowels [o] and [u] have no secondary articulation as in English, e.g., /oman/ ‘Oman’
- Liquids: the ability to distinguish natural Persian [r] and [l] from the English retroflex \( r \) and dark \( l \), e.g., /rævef/ ‘method’

Task 8: Acceptability Judgment Task

Participants are given a verb in Persian and must pick one of three light verb choices. This is a 100-item test. Light verbs can vary among a small set of verbs including those meaning ‘to do,’ ‘to hit’ and ‘to eat.’ It is not
necessarily clear from the semantics of the light verb construction which light verb should be used, so students must learn this.

**Task 9: Visual Grammatical Judgment Task**

Participants see a sentence and decide if there is an error of one of four regions of the sentence. They also have an option to say there is no error. Results from this 100-item task assessing learners’ ability to identify sentences as grammatical or ungrammatical will show the degree to which they have mastered the same features of Persian targeted in the aural GJT: the accusative morpheme ɿ, subject-verb agreement, tense and mood (including the subjunctive), and copulas. It also targets an additional feature: the causative morpheme which distinguishes causative verbs like پوشاندن/پوشاندن/ ‘to dress someone’ from basic forms like پوشیدن/پوشیدن/ ‘to wear.’ These grammatical features pose difficulty for advanced learners.

**Task 10: Aural Collocation and Light Verb Judgment Task**

Participants hear a sentence and decide if it is grammatical or not. Results from this 100-item task assessing learners’ ability to identify sentences as grammatical or ungrammatical will show the degree to which they have mastered light verbs and other collocations. These kinds of verbs are difficult for second language learners of Persian (Megerdoomian, 2010).
APPENDIX B: SAMPLE GRADING BREAKDOWN FROM THE DLI ONLINE DIAGNOSTIC ASSESSMENT

Reading Assessment

Language: Farsi
Date of the Diagnostic Session: 1/22/2014
Name: Evan Jones

Based on your performance in this ODA session, your ILR proficiency level estimate is 2+ (Current Level).

Your goal is to work toward proficiency level 3 (Target Level).

Note: The primary purpose of ODA is to provide you with formative feedback - feedback to help you in the learning process. The ILR level estimate you are given here is intended to function as a reference for charting your progress toward higher proficiency. You may or may not receive the same level at an official test.

The goal is to work incrementally toward your target proficiency level, by learning more about the content areas and the lexical, syntactical, and discourse aspects that you have not yet mastered. The following is a detailed diagnostic feedback on your performance.

Content Questions*

Linguistic Questions**

*Content Questions are all the questions about the meaning of a text, events, details, ideas and arguments.
**Linguistic Questions are those about key vocabulary, sentence structure and relations between ideas.
1 Performance Report – Current Level 2+:

Content Questions

Main Ideas
understand the main ideas of unfamiliar factual texts
understand the overall point of view of the author
understand simple opinion texts

Supporting Ideas
understand the major factual details of unfamiliar fact-based texts
understand the simple supporting arguments of a general discussion of ideas
evaluate which details are important and which ones are less important
understand some complex facts that are not part of daily or familiar occurrences

Linguistic Questions

Vocabulary
understand some vocabulary having to do with unfamiliar subjects
understand some abstract vocabulary

Subject Area Breakdown

Environment: environmental issues (e.g. protection and conservation) 6/7
Politics: foreign policy

Structure

understand some complex grammatical relationships within sentences

Structural Feature Breakdown

relate low frequency adverbs to the verbs they modify in some complex sentences with embedding (e.g., رئیس جمهوری، اصلاح، گریز

identify adverbs that relate to verbs in the subjunctive mood (e.g., ایران اقتصادی را در پی گیری سیاستی غرب سیرانه کرده

distinguish and differentiate both direct and indirect objects in some complex sentences (e.g., ایران اقتصادی را در پی گیری سیاستی غرب سیرانه کرده

Discourse

understand some complex relationships between sentences and paragraphs

understand how some idiomatic cohesive devices are used to connect ideas

Discourse Feature Breakdown

understand the role of cohesive phrases that relate ideas to each other by comparison, conclusion, contradiction, etc. (e.g., (اما از سوی دیگر

relate pronouns to nouns that are modified by phrases (e.g., چپ‌کش مذکور در قانون از

(ناظرت دولت خارج شده و شهرداری ناظرت بر ان را به عهده می‌گیرد)
Performance Report – Target Level 3:

Content Questions

Main Ideas

Correct / Total

6/6

Supporting Ideas

understand the major supporting evidence of the main argument

make inferences about ideas that are indirectly conveyed

identify the arguments that specifically support opinions about abstract subjects of a general nature

read between the lines

Linguistic Questions

Vocabulary

understand general vocabulary dealing with unfamiliar subjects

understand general abstract vocabulary

Subject Area Breakdown

Politics: political issues (e.g. democracy, corruption, human rights) 3/7

Society: social issues (e.g. poverty, crime...) 6/7

Security: security issues (e.g. terrorism, trafficking) 6/7
Politics: government systems (e.g. structure, administration, state, local, political parties, elections) 6/7

Culture: social structure and relations (e.g. family, class, ethnicity) 6/7

Culture: religion 3/7

Structure 4/6

understand complex grammatical relationships within sentences

Structural Feature Breakdown

identify group reference of subjects/objects across paragraphs (e.g., 2/3

identify adverb relation with verb in subjunctive mood units (e.g., 1/2

identify the function of emphatic words and phrases (e.g., 1/1

Discourse 6/6

understand complex relationships existing between sentences and paragraphs

understand the general cultural shaping of ideas through the use of idiomatic cohesive devices

Discourse Feature Breakdown

identify the role of conjunctions in completing, linking/ or elating complex ideas (e.g., 1/1

associate commonly used generic terms (i.e., 2/2

identify the author’s attitude by focusing on the use of words and idiomatic expressions 1/1

© 2014 University of Maryland. All rights reserved. May 2014
understand the role of cohesive phrases that relate ideas to each other by comparison, conclusion, contradiction, shifting, transitioning etc. (e.g., اما آن می‌دانست‌)<br>

identify the role of complex conjunctions that qualify or condition arguments (e.g., که (جوابی گه)
Teacher X

Meeting with Dr. X to discuss the needs of Persian learners

1) What are the most difficult things to teach/learn?

a. In terms of grammatical constructions, compound verbs are problematic, but often more so at lower levels. Problems can persist into the higher levels where students will, for example, use the wrong auxiliary verb in a compound structure. For example, انجام دادن instead of انجام کردن is a frequent mistake and such things can fossilize quickly in L2 learners.

Other less common phrasal verbs are also problematic at higher levels such as قائل شدن, which, despite what many students accept as a rule, actually takes a direct object, even though it is a weak verb with the auxiliary generally used for the passive. This is one of few unique cases that can still trip up higher level learners.

Ezafe is an issue at all levels, but one of the most obvious issues at levels 2-3. Especially in dictation students can continue to have issues distinguishing a final –he from ezafe. One example follows:

خودرو حامل سوخت/ خودرو حامله سوخت

The first and correctly transcribed phrase is ‘a vehicle transporting fuel’ whereas many students will misinterpret the sound of the ezafe to be a final –he which gives the reading ‘a vehicle pregnant with fuel.’ This becomes more of an issue in writing since the two phenomena are homophones. The subjunctive has also been problematic in the past for Dr. X’s students because their L1 often does not have such constructions.

b. In terms of phonetics, a student’s L1 always influences their pronunciation at some point or on some level. Issues specific to English speakers are uvular stops, on occasion velar fricatives, geminate consonants, distinctions between long and short vowels and specifically long and short ‘a.’ Intonation and the rhythm of speech are often very unnatural in L2 speakers due to L1 influence and these things are not taught explicitly in many classrooms.

The script can be a challenge as well because of the homophones in the script. Multiples of certain letters present challenges for those learners that do not have a strong background in Arabic, which is the source of these spellings.

c. On a related note, one of the hardest things to teach/learn is the distinction between the spoken and written languages. American students, due to how they learn the language and how it is most commonly taught, speak as they write, which is unnatural and awkward for native speakers. The lack of emphasis on the difference between the written and spoken language is the reason that none of the existing textbooks can really be considered comprehensive.

d. Dr. X also says that the proficiency tests used to measure how far a student has come are often skewed or flawed. Such tests as the OPI are as much a test of personality as of linguistic proficiency and are very subjective. It is also a very small sample of a student’s linguistic capability and may not be the best representation because the sample is so limited. For these reasons she disapproves of the use of such examinations.
e. Discourse particles, such as emphatic *ke* at the end of a sentence may be difficult for students to internalize and produce.

f. Flexible word order in Spoken Persian: It’s difficult to convey to students that there are certain words which can be moved around, while others can’t.

2) How do you recommend handling the issue of the written vs spoken language with your students?

First, the fact that there is a difference needs to be emphasized. Dr. X describes them like two parallel languages, one written, one spoken, which both have all the same capabilities for formality and informality. You can be very formal but still use the syntax and vocabulary of the spoken language. The differences in register exist in both parallels of the language, but this is not emphasized in most curricula. On top of that, the language used needs to be authentic and situational and is very difficult to replicate in the classroom.

According to Dr. X, students need to be exposed to both forms of the language in a natural way, using real language samples, not scripted examples from text, so they can become accustomed to the characteristics of each. She suggests making it tangible so the students can really see the difference between the written and spoken forms. Such things as writing dialogues are ineffective because they create confusion by using the written form of what is supposed to be an oral text. What she suggests is to use samples of listening material or video and having transcriptions of them for students to use as a supplement if needed. As long as the system (for transcription?) used in the classroom is consistent, students will adjust and adapt to make use of it. Opportunities to practice are also key, specifically at lower levels, but 2-3 can also benefit from the chance to use what they have learned and reinforce it.

3) What are the particular needs of the heritage student?

Heritage students are generally at a 1+/2 in speaking/listening, but 0+/1 in reading/writing. Because of this, the lack of exposure to written Persian needs to be addressed and often these learners also have less exposure to higher registers of the language. Unlike L2 learners who have more exposure to written Persian and speak as they write, heritage learners are the opposite and often write as they speak.

The tonality/intonation of heritage speakers often sounds non-native because they are often bilingual and their other language interferes. Their oral proficiency often makes them overconfident and they overlook their own literacy issues.

Dr. X says that ideally at lower levels heritage speakers would be split from other learners to address their particular literacy and grammar needs directly. L2 learners often grasp the structure of the language better than heritage learners, whereas heritage speakers can intimidate L2 learners with their oral proficiency. This kind of performance in the classroom can skew the instructor’s perception of the students’ skill level because of the grasp of grammar of L2 students or the oral proficiency of heritage students.

4) Why do students come to Persian and how does that affect their motivation?

Students in undergraduate programs often do not have strong motivation for why they study Persian. Dr. X finds that the students who know why they are studying are often the best students and it is far easier to perform needs analysis for them and develop the most appropriate curriculum because they often need to use the language for some specific purpose. These students also generally enter her program at a higher level than beginner.

Because of the nature of the language and political climate surrounding it, students often come into Persian with the intent of working in the federal government or translation and interpreting. Some also come in because of academic/regional interests, or some interest in the literature and culture. Yet other students come in to try to
gain some familiarity with the culture of friends or family, and some are pressured into classes by family members. These students can be less successful in class because it was not their own motivation that brought them to study the language.

Students from fields such as biology and agriculture and even human rights do not come to study the Iranian dialect of Persian. This is because if they are interested in the region, they cannot go to Iran, but they can go to other Persian-speaking countries where international NGOs can function. Due to this, however, they would more likely tend towards lessons in Dari or Tajiki than Farsi because it would be more applicable to their line of work.

5) What textbooks do you recommend?

Historically, teachers of Persian in the US have not had a background in language pedagogy, SLA, methodology, linguistics, or any field related to teaching a language. UMD was the first institution to have teachers with SLA and applied linguistics backgrounds working as their main Persian instructors. Many of the textbooks, however, were not designed by this group of specialized instructors or were not designed by someone with a background in methodology. This limits the effectiveness of the instructional materials. At lower levels for structure’s sake, Dr. X would consider having some kind of supplemental textbook, but at the higher levels (i.e., 2) she would limit its use and focus on authentic materials and would cut out textbooks altogether at 3 and above.

As previously mentioned, none of the existing texts address both the written and spoken varieties of Persian. Anousha Shahsavari at UT Austin is working on materials to teach the spoken language better (called The Persian of Iran Today, with a supplemental video series called کافه دنده) and although it is still scripted, the language is more natural than previous textbooks. Outside of this, Dr. X has no suggestions for textbooks and considers the vast majority of them inadequate.

6) What dictionaries do you recommend?

For print dictionaries using Persian as both the source and target languages, Dr. X recommends فرهنگ سخن and for Persian-English print dictionaries recommends Bateni’s dictionary پیا, which unfortunately does not have examples in context. In terms of online dictionaries, she suggests www.farsi123.com and Dehkhoda’s online dictionary (www.loghatnaameh.org) which both have examples in context.

7) What websites do you recommend?

a. For learning the language?

UT Austin has good online reference materials for the grammar, and also now some supplemental videos to focus on the spoken language.

b. For practicing the language?

For practicing, Dr. X recommends Jadid Online, Gushe, Man o To, and BBC’s news/talk show programs. She is a huge fan of talk shows, call-ins, and Nobat-e Shoma on BBC which all feature natural samples of the language because they are not scripted. DLI’s online materials could also be used for problem-solving practice.

8) What kinds of homework or exercises work best to supplement this?

Dr. X tries to assign 2-3 active hours of work outside of class for every hour spent in class. She intends the work to be meaningful and demands it from her students, but tries to control it at lower levels so it isn’t
overwhelming. She strives to recycle activities, vocabulary, and lessons from the day in the homework to reinforce it.

Dr. X assigns not only written, but also oral homework assignments to support the distinction between the two parallels. She requires students to record themselves speaking using ELMS or the university’s online blackboard system. The students are not supposed to read from a script while doing this. Dr. X stresses that this type of homework could be used for all levels of learners and in all registers, as well as that it reinforces the parallelism of the language.

Teacher Y
Meeting with Dr. YY to discuss the needs of Persian learners
15.11.13

1.) What are the most difficult things to teach?

a. most difficult grammatical constructs:

The accusative postposition rA was what Dr. Y first listed as one of the most difficult grammatical structures to teach non-natives. He specified that errors with this persist even into higher levels. He also mentioned the subjunctive, ezafe, and how exceptions to such rules add to the difficulty.

b. most difficult phonetic issues:

Dr. Y mentions that many students pronounce the uvular plosives غ and ق as [g].

c. other:

In regards to this section, he brought up discoursal differences between English and Persian and cultural awareness to such issues as the sociopolitical atmosphere of Iran, taarof, and indirectness. He also brought up the difficulty of teaching language functions, such as with different instances of taarof, because it is so challenging to find contextual samples of the necessary speech events to illustrate these language functions.

2.) How do you recommend handling the written language/spoken language issue with your students?

Dr. Y cited the divide between the written and spoken languages as the hardest thing for students to come to terms with. He also noted that he has never taught a group of absolute beginners, but if he could, he would immerse the students entirely in spoken Persian for the first four months with no introduction to the script to simulate how Iranian children learn. He accepted that because he is working in an academic environment, this may not be feasible, and because these are adult and not child learners some explicit instruction may be required in addition to the implicit instruction of the immersion environment. Down the line, he would hope to introduce the written language in a way that the students would be able to decode the differences. He would also implement assignments for both the spoken and written languages.

3.) What are the particular needs of the heritage student?

Heritage learners have good spoken capability, but it is generally at a low register and employs a very limited vocabulary. Heritage students need to work more on the written language and in some situations ‘unlearn’ the spoken language. Dr. Y has noticed that heritage learners often seem impervious to error correction and he thinks it might be more efficient to correct student errors when the student initiates the discussion of his or her errors.
4.) Why do students come to Persian and how does that affect their motivation?

Often students with genuine interest in the region or language or literature are the best motivated students. Some are motivated by career opportunities, but their motivation does not include learning the culture behind the language, rather only extends up to the point where they believe they can meet their own linguistic goals. Other students are pressured by family to take classes or required to have a language to graduate.

5.) What textbooks do you recommend?

Dr. Y samples from the AZFA textbook series, American Councils’ Panorama on Iran texts, among others, but any of these texts alone is not very useful or up-to-date. The professor describes these textbooks as ‘pre-scientific era’ (something we should not include in the final report) and says the pedagogy in them is decades outdated. Dr. Y believes that what Persian needs is a well-developed series of textbooks from beginner to advanced that can function as a core text for a course.

6.) What dictionaries do you recommend?

He suggests Sokhan and Dehkhoda, but remarks that they are not helpful for L2 learners because they are Persian-Persian monolingual dictionaries.

7. What websites do you recommend?

Dehkhoda is the only site that Dr. Y mentioned.

8.) What kinds of homework/exercises work best?

Dr. Y commented specifically on homework for the spoken language and said that he likes to use episodes from various Iranian TV series because they provide exposure to both the language and the culture. On top of this, the students can be prescribed tasks based on their level. For example, a student can be given a 30 second clip and be asked to describe a certain character’s appearance or a particular interaction, etc.

9.) What kinds of things would you like to see in a Learning Management System (an online instructional system like Blackboard) for Persian?

He would like to see the LMS have Persian functionality, meaning that students can input Persian text without the issues of text or punctuation not being portrayed right-to-left.

11.) How and when do you introduce the script to students?

He has never taught beginner students.

Additional notes

Regarding spoken/written differences: if you don’t get this write, you will speak in a bookish manner which causes you to be stigmatized by native speakers. Dr. Y believes that whatever is written should be in written form (rather than attempting to mimic the spoken language in script). He admits that others (including his wife/colleague) disagree with him. He mentioned a kind of exercise that could be done whereby a written sentence is presented: AyA be tehrAn rafte id? And this sentence must be transformed into spoken form by the student, e.g., tehrAn rafli?
Particular needs of heritage: good spoken, need reading and writing practice. They need more diverse lexical resources, they often write what they hear, require some amount of unlearning. He said they can be “impervious to correction”, whether or not they are motivated. Also, he said their situation can be considered an example of “fossilization”.

Regarding correction in general, he said it is better to correct as the need arises. There is a right time when the student is ready to absorb the correction (perhaps when he explicitly asks about something?). Every learner has his or her own “internal syllabus”.

Some students today have career (e.g., gov’t work) motivations. Those students are not necessarily interested in the culture.

Dr. Y supplied several text books: Zarghamaian, Oliai, said they are not necessarily up to date, and may be “pre-scientific.” Textbooks are devalued for tenure, so we don’t get as many as we should.

Dr. Y said certain television series are good for teaching, e.g., sAkhtemAne pezeshkAn (The doctors’ building, a satirical series depicting the day-to-day trials and tribulations of a middle class man), gives students insight into the spoken language in daily use.

Regarding taarof, Dr. Y said it was a kind of “phatic communion” and referenced Malinowski.
APPENDIX D: INTERVIEWS WITH ADVANCED LEARNERS OF PERSIAN

A: learner of Persian

A started studying Persian at the age of 20 with a private tutor (an American grad student) while doing his undergraduate degree at the University of Arizona in Tucson and then took two semesters at the beginner level at the university. After this he estimates he was not even at an ILR 1, but went to Tajikistan for 9 months on an intensive immersion program. For the first semester this program focused on Farsi in the classroom, but Tajiki was used in the homestays, and the second semester he opened up his coursework to include Farsi, Dari, and Tajiki.

After this program, he spent approximately a month in Afghanistan where he got more exposure to the Dari dialect and then went back to UA for a year of study at the advanced level. He simultaneously took a semester of Russian and Turkish due to the influence of those languages on the Persian spoken in Tajikistan. He graduated in December 2007, worked for a month in Tajikistan over the summer, then the following fall semester in 2008 began the graduate Flagship program at UMD. This academic year was 100% Farsi in class, and then he spent another academic year in Tajikistan for the second year of Flagship where he focused on the spoken Tajiki dialect since there was no pervasive and natural way to gain exposure to spoken Farsi.

1) A doesn’t mention many particular grammar points that he had difficulty with other than that he didn’t understand –rA usage his first year abroad and learned by exposure and observation (we should try to have the LMS emulate an immersive environment, i.e., be in the target language). He noted that on occasion because of syntax variation between the written and spoken languages as well as between dialects, he had issues with word order. In the same vein, he remembers that in longer utterances he would often drop verbs at the ends of sentences.

Talattof did a good job early on in A’s studies providing him with practice on problematic phonemes, but words in similar patterns are still difficult in some situations for him. One example is پهنازان/پژوهشگان (refugees/researchers) and another is موافق/موافق (agreed, successful) which is problematic for many students.

He said he had issues with formality initially and addressed everyone with more formal pronouns. Taarof and related courtesy phrases are still a challenge in part because they are more widespread in Farsi than Tajiki. Light verb constructions (and knowing the correct auxiliary to use at a certain time) were still issues for him as well. A mentioned that the prevalence of Arabic words was an issue, but he had one teacher in Tajikistan who specifically addressed this issue and helped him learn the basics of root morphology so using Arabic constructions was both made easier and more productive.

One issue with assessments for Persian that A mentioned is that he feels most OPI examiners, who are generally testing for Farsi, not the other dialects, will mark you down for using other Persian dialects although they are the same language. Sometimes even just having an accent other than standard Tehran Farsi can have a negative impact on the test results.

2) For A, the written/spoken parallel wasn’t an issue until he went abroad and even then he went to a country where the spoken language closely resembles the written language anyway OR is so different it’s practically a different language. Generally he used more formal language and avoided colloquialisms, especially Farsi colloquialisms because his Farsi exposure was limited to the classroom and he spoke Tajiki outside of class.
3) At UMD, classes were almost half heritage students and all speaking partners were heritage or native speakers of Farsi. Through these people, A gained some exposure to Farsi colloquialisms and did not consider these students a hindrance to the learning process. A, though, was also more interested in the mechanics of Persian as a whole and not just Farsi because the Tehran dialect is useless in Dushanbe and many heritage students there could not handle switching between dialects.

4) He came to study Persian because he was a Middle East Studies major in undergrad and wanted to take a language. He was not interested in Hebrew and thought Arabic would result in slower, less visible progress, so he had to choose between Persian and Turkish. He almost went to Turkey instead because he could not find a study abroad immersion opportunity for Persian, but happened upon American Councils’ program in Dushanbe and proceeded from there. He chose this route in the wake of September 11th and the beginning of the war in Iraq figuring he could potentially get work in the government.

5) A only mentioned one textbook that he used: Stilo, a two-volume hardcover. This book starts with romanizations and focuses heavily on phonetics and the spoken language in the earlier stages. Outside of personally motivated practice, A had no exposure to the script until halfway through the first semester using this textbook. Flagship in Tajikistan also provided a textbook on Arabic constructions used in Persian intended to help students see the patterns in Arabic morphology. He also thinks he used Thackston’s grammar on occasion. He brought all of these textbooks to Tajikistan and left them there because the programs he participated in were lacking even resources such as this.

6) A did not generally use a dictionary stateside and eventually started using a small Hippocrene pocket Persian dictionary, but realized eventually that it was a poor resource because in the translation for ‘about,’ it had only which means about in the prepositional sense, but he was using it to mean about in the sense of ‘approximately.’ A used the Haim dictionary and a Tajiki dictionary that included Perso-Arabic script in its entries while abroad, but notes that he often had to work with multiple dictionaries simultaneously while translating because some were only Persian-English or vice versa and he also had to deal with different dialectal challenges.

7) Until www.farsi123.com came along, A used very few online resources for learning Persian. In terms of websites used for practicing the language, A uses google, Radio Free Europe for both Tajiki and Farsi, blogs, and BBC. He notes, though, that BBC articles are highly stylized and written in a Western manner, so the materials are almost formulaic and not necessarily the best samples of the language. For more native Persian language material he would find a blog which is more likely to exhibit stylistic and non-westernized Persian writing.

8) A thinks that of the programs he did, most did not have much homework. Some had minor drills, but he generally took on translations of texts from Persian to English as homework or work outside of class and thinks that was most beneficial for him. He also made an effort outside of class to stay in an immersed environment, which he emphasized was most beneficial for him.

B: learner of Persian

In 2009 at the age of 23, B started studying Persian when he moved to Tajikistan. He had no background in the language previous to this. While in Tajikistan he lived with a host family and worked in an environment where Tajik Persian was used for most communications. He had no formal classes, but did arrange for a native Tajik language partner who knew some English with whom he would meet 3 times a week for about an hour each session. This informal learning was reinforced with his hosts and studying from textbooks in his own time. The second year he lived in Tajikistan he started meeting an Afghan language partner instead, but his work and living situation had not changed. The sessions were about 1.5 hours each, still three times a week, and this is when he finally started learning the script.
B used his language partners as a field linguist would use consultants and elicited structures from them instead of using them solely for practice. After his two years in Tajikistan he returned to the US, maintained his Persian over the summer by listening to and reading the news, and in the fall of 2011 began the graduate Persian Flagship program at UMD which was entirely in Farsi. These were B’s first formal classes in Persian. During this time, B used Tajik media as a resource for work he did writing for a think tank. After the year in grad school, he spent 6 weeks as a technical translator and now uses Persian at work.

1) B immediately mentioned –rA, the direct object marking postposition, when asked what grammatical challenges he encountered while learning Persian. He says Ezafé can still trip him up in reading, for example. Early on he says the SOV word order was also cumbersome as was getting used to conjugating verbs for inanimate plurals in the singular. He also said that he felt that he overuses pronouns, making him more noticeably a non-native speaker, so the pro-drop features of the language were also of concern for him.

B specified that the ꞉ (ɣ) and Ꞛ (q) were the hardest consonants to learn, whereas he thinks his vowels in general are off because he has studied all of the three main dialects and is the most familiar with the least widely spoken dialect, which has vowels somewhat different from Farsi or Dari. On top of that, he speaks a somewhat more esoteric non-standard regional dialect of Tajiki because of how he structured his language partner sessions. In addition, he now works specifically on different dialects of Farsi spoken within Iran, which he thinks will affect his vowels further.

B said the dialect continuum itself was problematic specifically because he was using non standardized or less commonly spoken variants most often. Because of the dialect issue, he also said it was difficult to learn new vocabulary because sometimes each dialect uses a different word. Spelling in the Arabic script because it’s not phonetic is an issue he still has, but says spelling in Cyrillic is less difficult. Many of the TLTI words are troublesome for B.

2) B feels most comfortable with the written language by far. Outside of a sociolinguistics class for Flagship which discussed register changes and the spoken language in depth and some comments made here and there by the UMD instructors, all of the spoken language he learned through language partners. If he were to do it again, he would try to get a more balanced emphasis on the spoken language in relation to the written language.

3) B, overall, viewed the presence of heritage speakers in the classroom as a positive thing because it was beneficial for his own speaking and listening skills. He prefers to be on the lower end in terms of skills in the classroom because the students at a higher level give him something to move towards and create a better learning environment for him. However, one disadvantage he did note was that sometimes he would learn something incorrect from the heritage speakers, something used only in the diaspora community and not used in native Persian.

4) He came to Persian because he thought he’d be able to find work if he had this language. Also, while he was in Tajikistan, he either had to learn out of necessity and he forced himself to learn to speak as the natives on the street would. His motivation to maintain it is professional as he uses Persian at work on a daily basis.

5) He has used no Farsi or Dari textbooks, but had a Tajiki textbook that was helpful for him and he used outside of his sessions with his language partners. He used the textbook as a tool to explain the technical end of the language that most language partners didn’t have the background knowledge to be able to discuss or explain.

6) He used only one physical dictionary and it was specifically for Tajiki. Currently he uses farsi123.com and google translate almost exclusively. He asserts that google translate is really only good if you know how to use it and for beginner students it may be less effective. B also mentioned that he has starting using Wikipedia as a dictionary for more abstract concepts.
7) B used the University of Texas at Austin’s site that reviews Persian grammar and breaks it down on a series of pages for brushing up when he needed it. He also uses GLOSS because it gives him the option to select a level and modality to work on and has a lot of material. For practicing the language, B uses Tajik news sources, khamenei.ir (the Supreme Leader’s website), BBC, BBC’s Persian TV, and youtube and facebook for the more colloquial language and registers.

8) Having not had formal classes, B didn’t encounter much homework, but says that of what he did have, he most benefited from preparatory work for class discussions where in class he would be told the topic of what would be the next session’s discussion in class, he would go home and do some background research, look up domain-specific words, and then make use of that in class. He also found translation and/or summarization a good exercise because for him it’s a measure of whether or not he actually understands a text.

9) Other comments: B says if he could do it differently, he would want more structure (and more exposure to Farsi) early on than he received since he didn’t have formal classes. He also says that a formal lesson on triliteral Arabic roots and a basic introduction to Arabic morphology would be very helpful in learning and retaining groups of borrowings from Arabic used in Persian. In terms of vocabulary, he mentioned that learning decontextualized lists of vocabulary has not been incredibly effective for him and he has started using google searches to find examples of new lexical items in context, which is more helpful for him in terms of learning usage and retaining the vocabulary.

B also said increased feedback would have helped him develop his Persian better when he was in Flagship. He seemed to think that he got more feedback from his Flagship language partners than he did from the instructors. He also brought up that he thinks many of the widely used proficiency tests, specifically the OPI, are inaccurate and unreliable because they ask for presentations on the spot, which most normal people would need preparation time to be able to provide. He says the OPI testers also often give too much credit for using technical terms and other 25-cent words, while they will mark off points if the examinee sounds too much like a different dialect of the Persian language.

C: learner of Persian

C started learning Persian at DLI in October 2005. She enrolled in a full one-year Persian course. Upon graduation, she scored a 3/3/2 (Listening, Reading, and Speaking) on the DLPT. Between July 2007 and Spring 2011, she was not formerly enrolled in Persian courses, but she did enroll in 2-week refreshers courses every six months or so to prepare for the DLPT to maintain her foreign language proficiency scores. In 2009 (?), part of her job became to teach these courses. Throughout this time period, she maintained her original DLPT scores, except for a slight dip upon taking DLPT 5 for the first time, when she scored a 2+/2+. In the Fall of 2011, C had begun working at CASL and enrolled as a part-time student at UMD in the Persian Flagship Program. She took an OPI in 2012 and scored an Advanced High (ILR 2+). She uses Persian for work at CASL, and, besides the language class per semester she has been taking at UMD, she also engages in regular conversation practice with a language partner.

C noted that especially for her DLI formation, there is a heavy emphasis on vocabulary and less so on more formal properties of the language. In order to maintain her language abilities, she engages in conversation practice and continues to enroll in more content-based courses, such as a course on Contemporary Iranian Female Writers on the UMD campus.

1) For most difficult grammatical points, C mentioned –rA (object marker), magar, and verb tenses, especially the habitual past. She also mentioned that words with Arabic roots are easy to confuse. She had trouble initially producing ꝱ and ꝝ, although they are not difficult to produce. There was some trouble initially distinguishing the different letters that express s, z, and t in early stages.

2) C said the classes at UMD were helpful to help sort out the differences between spoken and written Persian, a topic which was more confusing to her at DLI. Her instructors at UMD have been helpful by pointing or highlighting errors in spoken vs. written discourse.
3) At DLI, there was only one heritage learner that C recalled, while at UMD usually there is around a 50/50 balance between heritage speakers and L2 learners. They have been helpful to her for understanding dialectal variations and different regional pronunciations.

4) After taking an aptitude test and being given a choice of languages, C did not know what Persian was, so decided it would be a good language to learn. Since then, she has been highly motivated to continue improving.

5) Parsi, a Persian grammar, Kimia (Farsi-English), Farhang-e-Moaser Persian-Persian, Lambton

6) Steingass and Haim; Google Translate

7) GLOSS, Transparent Rapid Rote, BBC Persian, Radio Farda, movies on Netflix

8) Now, C is writing essays for her UMD classes, which she is enjoying to challenge herself to find the appropriate register. She also has always enjoyed doing projects, such as researching a topic or creating a product and presenting it to a class.

9) When asked what she thought the LMS should be able to do, C said there should be a module early making a distinction between the different trouble modalities, specifically between written Persian and spoken Persian.
APPENDIX E: REVIEWS OF SELECTED PERSIAN INSTRUCTIONAL MATERIALS

Title: Spoken Persian
Author: Serge Obolensky, Kambiz Yazdan Panah, Fereidoun Khaje Nouri.

Spoken Persian was developed for the Foreign Service Institute’s language training courses and for this reason focuses on the variety of Persian (Farsi) spoken in Tehran. The introduction to the text specifically states the intent to treat only the spoken language and not the written one saying “[l]earning to read should not keep the student from the primary goal of learning the language. Progress in the language may be hindered by trying to read too soon.” This probably refers to the issues many native English-speaking students encounter when they learn that spoken and written Persian are markedly different, but whatever the reason, this text does not directly introduce the Perso-Arabic script, although the script is used on occasion.

This textbook suggests that lessons be conducted with a native speaker to learn Persian most accurately. Each lesson in this textbook contains three parts: new material, a grammatical explanation of a new phenomenon, and lastly a drill. There are no true assessments or keys to exercises included in the textbook. Glosses of new vocabulary appear where necessary and the lessons are reduplicated in Perso-Arabic script following the Romanized lesson. These are presented in their written form and do not always match up to the spoken forms presented in the Romanized lessons and for this reason may mislead some students trying to simultaneously learn to read (‘zadam’ was written together as the Romanization for "زدم و" which is written separately in standardized written Persian. Another example is ‘nemidunam’ which was coupled with the standardized spelling ‘داﻧﻢﻧﻤﯽ’ which would indicate a long [aa] and not a [u]). Some of these lessons, though, do a better job of presenting spoken forms in Romanized lessons than other textbooks do addressing the spoken language at all.

The textbook includes a relatively detailed account of the phonetics, stress, and intonation used in Persian, although the diacritics and system of notations for these is rather complicated and makes the texts that they are added to quite distracting and confusing. Nonetheless, few other sources breach this aspect of the spoken language if they even address the spoken language in the first place. If students can work through the complex annotations, they will have a better idea of accent and intonations in Tehran Persian.

Some forms are either still closer to the written language but being taught as spoken, or else have changed in the Persian currently spoken in Tehran since the time of publication of this textbook. Since the book was developed for the Foreign Service in 1973, and the Foreign Service has not had much interaction with Iran since shortly thereafter, the fact that this has not been updated is understandable.

Spoken Persian is not a comprehensive grammar, but does present the most common aspects of the language, e.g., the most common verb tenses and verbal constructions used in spoken Tehran Persian (including a common spoken passive form ‘they kicked him out’ for ‘he was kicked out’), numerous useful vocabulary items pertinent to everyday life or work (bimaristan/marizkhane, detailed kinship terms), and some variations in address for different registers.

This text is unique in that it draws attention to how to address those to whom you want to show respect, which is an integral part of professional Persian. For example, in general conversation, someone could use ‘goftan’ for ‘say.’ However, when speaking at a higher register, this text notes that the speaker does not use ‘goftan’ for things s/he has said to the person demanding respect, but rather uses ‘arz kardan’ for themselves and ‘farmudan’ for things said by the addressee. Similar to this, in general conversations, typical pronouns can be used, e.g., ‘man’ for ‘I,’ ‘to’ for ‘you’ (informal singular), or ‘shoma’ for ‘you’ (formal or plural). At a higher register,
though, ‘man’ can be heard as ‘bande,’ literally meaning ‘servant,’ and ‘jenab aali’ for a respected male addressee or ‘sarkaar’ for a respected female addressee.

One criticism is that this textbook, despite claiming to teach the spoken language, still uses forms relatively more similar to the written language. For example, in the book, to be polite one could say ‘che goftid’ for ‘What did you say?’ and ‘chi goftid’ in more familiar situations, but does not note that this Tehran Persian often changes this to ‘chi goftin.’ Similarly, in Tehran, [t] is often elided or dropped (e.g., ‘dasse raasse’ instead of ‘daste raaste’ or ‘kojaas’ instead of ‘kojaast‘ as used in the text) and long [aa] often turns into [u] before nasals (e.g., ‘tamum’ instead of ‘tamaam’ as listed in the text). ‘Khoshvakht’ being labeled as ‘ﺧﻮﺷﻮﻗﺖ’ instead of ‘ﺧﻮﺷﺒﺨﺖ’ is a common misreading even for native speakers, so the important thing is that such a common word used when being introduced to someone new at least is pronounced correctly. There are some vocabulary words that are given definitions that may not only be misleading, but just incorrect. One example is the word ‘گیلاس’ which is glossed as ‘glass.’ Although it may sound like the English word ‘glass,’ it actually means ‘cherry.’ As far as this text is concerned, such definitions are few, but it should be noted that they do exist.

Grammatically, this text still uses Persian that is somewhere between standard written Persian and the spoken Persian of Tehran. For example, the prefix used on subjunctive verbs is predominantly ‘be,’ but introducing the subjunctive form of a basic verb as ‘bekone’ instead of ‘bokone’ would sound noticeably foreign and strange for a native speaker of Persian from Tehran. Additionally, Persian is a language that does not have equivalents for the articles ‘a’ or ‘the’ which the book notes, but the text also states that in more informal spoken Persian, a stressed suffix ‘-e’ can be added to nouns to make them definite. In fact, this suffix does occur, but in a much more limited context. It does not mark general definiteness, but rather is used to specify or clarify a definite noun that has already been referred to in conversation. Although neither this nor the case of the subjunctive are entirely incorrect, they are not accurate descriptions of spoken Tehran Persian, which is what the book tries to address, and examples of these inaccuracies are prominent throughout the textbook.

Title: An Introduction to Persian
Author: W. M. Thackston

Thackston’s An Introduction to Persian uses the written language, but occasionally notes colloquial pronunciations as well. Later sections address in detail the classical and archaic usages of the language, and then the spoken language, respectively. Further explanatory appendices include information on problematic particles and other expressions such as times in the calendars more often used in Persian-speaking regions. Near the end, the author includes samples of contemporary and classical prose along with a glossary.

The writing style is accessible to most readers, however linguistic vocabulary would be required for terms like ‘enclitic,’ which Thackston never defines, but uses often. This author uses terms for the tenses such as ‘remote absolute’ and ‘past narrative’ to describe the function of the tenses instead of using designations better known through more commonly studied languages. All forms of the conditional, which tends to be problematic for non-native speakers, are very well addressed in this text, including written and spoken differences. Thackston includes exercises at the end of each section to review previous material, but there is no answer key provided. The sentences generally require translation from one language to the other and clearly are designed to reinforce knowledge not only from the current lesson, but also draw on previous lessons in terms of grammar and vocabulary.

This textbook is thorough in its treatment of the language, especially handwriting, phonemes, syntax, Persian and Arabic pluralizing mechanisms, and verbal phrases in the language. It also has good accounts of the functions of enclitics (keshvari bozorg/keshvare bozorgi) and affixes (zadamesh) on nouns and verbs, different count nouns and their areas of use, precise kinship terms (havu: a second/rival wife, pesarkhaleh: mother’s sister’s son), and other features of Persian that do not exist in English. Javad sadam (Javad hit me…this never happens, too ambiguous even in context. ‘Zadamesh’ could work, though)
Excellent notes on danestan and shenakhtan in terms of their use with/out the durative prefix. ‘Know (sthg)’ vs ‘realize,’ ‘know (so)’ vs ‘recognize.’

**Title:** Intermediate Dari  
**Author:** Jonaid Sharif

This book reads right-to-left to accommodate the wide usage of Perso-Arabic script, which it explains in a paragraph at the left end of the book for those expecting it to read left-to-right. This book was intended for those learners who are heritage speakers of Dari or already have some proficiency in Persian, ideally as an aid to those heritage speakers who were raised entirely in a diasporic community. The text is presents new grammar points and has exercises, although not at clearly regularly intervals. These lessons and their exercises have many sources and may include short stories, matching, multiple choice, grammar-oriented drills, or vocabulary exercises. What is different in this in comparison to other similar texts is that at the end of the book there is a key included for all of the exercises so the reader can check their answers and correct any mistakes there may be. After that portion of the text, there is an appendix which includes romanizations of words written in script and a glossary at the very end of the book.

This text often contains explanations in Dari, which is presented in the standard written language, but also includes examples of spoken Dari which are written in romanized Latin script. If written text appears in terms of explanation or instructions and not as examples of the spoken language, there is nearly no way to distinguish it from standard written Farsi. The most common distinguishing factor in writing is vocabulary choice; whereas one word may be common to Farsi, a different word borrowed from Arabic or a Persian synonym may be more common to use in Dari, but educated native speakers and advanced learners with the appropriate exposure would likely understand both lexical selections. The practice of romanizing examples of spoken Dari is most useful for distinguishing the short vowels in words, specifically the vowels that differ from the Farsi dialect, because in the spoken language, pronunciation is what predominantly differentiates the two dialects. The author also presents the grammatical distinctions between the Farsi and Dari dialects throughout the text.

Source texts used in exercises seem authentic, whereas many other texts will specifically for the purpose of being used in the book. Intermediate Dari uses various source materials, as well, such as stories, poems, recipes, and jokes, which expose learners to a much less limited range of language in comparison to many textbooks. The glosses for vocabulary listed after source text is listed in order of the vocabulary in its native script, but not along the same line, which makes for a somewhat disjointed reading experience. Another characteristic of this text that made reading a challenge was that tables don’t have column heads, so it’s difficult to tell if the columns are comparing dialectal differences, synonyms, antonyms, or something completely different, but the author does make an effort to address and present all of these things in some way.

This text describes the influence of Arabic on Dari in detail. It includes such things as synonyms that now exist in Dari where one word is of Persian origin and the other is from Arabic. In the text, the author also goes into an explanation of Arabic verb and participle forms and some commonly seen patterns of broken plurals. This familiarizes the reader with Arabic roots and their morphology which is incredibly important considering the sheer volume of borrowed Arabic words and terms in Persian. Knowing the relation of words with the same root can help a learner understand the meaning and retain new vocabulary faster.

Knowing Arabic and Persian synonym pairs helps further distinguish Dari from Farsi in that some common vocabulary in Dari is from Arabic loanwords, while the Farsi equivalent is an original Persian word. The author points out Dari’s similarities with the Tajiki dialect and even highlights some words of Turkic and Pashto origin.
used in Dari. In general, the vocabulary used in this book is rather varied and extensive and the author includes the connotations for some lexical items to help the learner use them properly.

This book presents several important grammar points specific to Persian that help greatly with wordbuilding, such as the creation of causative verbs from standard verbs (which differs only very slightly from Farsi), and collocations. Additionally, the author addresses the challenges of ‘ke’ which is often used as a conjunction, but colloquially can have numerous other meanings. The productive process of building new words is also touched on when the author points out a newly created word (‘calorieshomaar,’ someone who counts calories) and how often new words such as this are created that cannot necessarily be found in dictionaries but can be understood.

Other things that are not easy to comprehend with only dictionaries as a resource are culture-specific phrases, which are widely used in all dialects of Persian. This book provides an account of many of the most common phrases, some of which vary slightly from their Farsi counterparts, and also describes the situations in which such phrases would be used. The texts used in individual lessons get progressively more challenging and abstract, both signs of higher proficiency texts, as the book goes on with some of the later texts discussing ancient history or Sufism. In sum, this textbook contains useful exercises for more proficient learners of all dialects of Persian, but would be most useful to those with a basic knowledge of Dari or higher proficiency in Farsi.

Title: Beginner’s Dari  
Author: Shaista Wahab

This Beginner’s Dari book begins with a brief introduction of Dari, its history and usage in Afghanistan, and some basics about the writing system. This introduction also includes a description of how to use the book, including a description of the five sections of the book. The first section addresses the alphabet, while the second and third are entitled ‘Rules’ and ‘Grammar,’ respectively, and both address grammatical constructions of Dari. The fourth section is basic vocabulary and the fifth includes commonly used phrases. Following this is an English-Dari glossary at the end of the text.

The first section splits the alphabet into sections based on words with similar shapes. It verbally describes the strokes and directionality used to write each letter, but does not have step-by-step diagrams included in other textbooks meant to teach this writing system. Fortunately, it does include vocabulary that uses the target letters after each subsection. The second and third sections both address Dari grammar, although section 2 is listed as ‘Rules.’ These just seem to be language-specific grammar points, whereas the ‘Grammar’ section contains aspects of Dari present in other languages such as the different verb tenses.

The fourth section focuses on vocabulary in a variety of everyday settings, basic objects, times and months (different from the standard Western calendar and based on the astrological calendar), numbers, weather, and locations in a city. This section is rather short and the vocabulary is limited to very basic nouns and adjectives and does not touch on verbs much. The fifth section on common words and phrases is far more useful in terms of actual vocabulary. This section begins with a brief description of introductions in Afghanistan and the differences in greeting males or females, which is valuable information for those not familiar with the division between the sexes in the Middle East. The vocabulary presented in this section presents such things as phrases to get a taxi, check in at a hotel, and kinship terms, which are far more precise than in English and in some cases differ from Farsi, and some other commonly used phrases listed alphabetically by their English translation. Following all of this is a glossary of vocabulary terms that actually contains some vocabulary items not presented in the lessons preceding it with a slightly expanded inventory of terms that would be of use in day-to-day exchanges. These include some basic illnesses, expanded numerical terms, and more prepositions.
At the beginning of this text, the author states that she is basing the text on Dari as spoken in Kabul, the capital of Afghanistan, but she does not follow Kabuli pronunciation in her glosses and in fact almost exclusively includes formal standard pronunciations instead of colloquial Kabuli pronunciations. For example, glosses often show romanized pronunciations with ‘v’ as in va ‘and,’ while speakers in Kabul pronounce this word with a ‘w’ sound instead. The romanizations in many cases do not represent the actual pronunciation (khw‘andi ‘you read,’ but the ‘w’ is a relic of older forms of Persian and is no longer pronounced) or are not consistent (bero and birau for ‘Go!’). Indeed, the system for romanizations is never directly addressed and the reader is left to infer that an apostrophe before a vowel is to signal a long vowel in pronunciation. This when compounded with the typos in the text can make the learner’s task more challenging.

There are some issues with this text that may hinder the learning process. Specifically, the author lists a plural marker –el/-یﻞ as a common pluralizer when in fact this is not a productive plural marker. This is in fact just the ending of a borrowed Arabic word using a broken plural form. Oddly enough, the textbook never mentions broken Arabic plurals and the many forms they may take, even though they are quite often used in Persian. The most common Arabic broken plural –aat/-ات also was not listed when pluralization was discussed. The text is also unclear when a preposition is used in script because it is written attached to pronouns, which may cause difficulty for learners unfamiliar with the script.

This book makes mention of grammatical gender twice, but the first section entitled ‘Gender’ says that the gender of a noun can be distinguished in context (e.g., we know that ‘father’ is male, biologically). However, the second time gender is mentioned in the section discussing pronouns, the author explicates that there is no gender distinction whatsoever in the language, meaning no distinct pronoun for ‘he’ instead of ‘she’ even. Mentioning this overtly in the section dedicated to gender distinctions could have been more effective and less redundant when addressing the pronouns.

The pronouns themselves led to an introductory set of verb paradigms: first the present tense, then the past, and so on. The paradigms themselves are inaccurate in the second person singular forms (you), which present the same conjugation as for the second person plural (you all). In some situations, this form is used to signal formality, but there is also a variant conjugation restricted only to second person singular that is not mentioned, except for in a footnote linked to the past tense paradigm later on. The present tense also lists the third person singular ending as null, which it is in the past, but not the present. Finally, the verb stems change for certain verbs in the present tense, which makes the present tense harder for learners initially. It may have been more efficient to present the past tense first since the past tense stem is clearly visible in the infinitive. The future tense is presented well because it shows the difference from the future in Farsi.

Although the textbook presents khubtarin/خﻮﺑﺘﺮﯾﻦ ‘goodest’ (which happens very rarely among native speakers and only as a slip of the tongue) as the superlative of ‘good’ instead of behtarīn/بھﺘﺮﯾﻦ ‘best,’ the text is good at addressing words that are probably of Pashto or Turkic origin in Dari, such as the words for university, street, window, and chair, that are not used in Farsi. It also has a brief section on the uses of the suffixes –ak and –cha, which can be used in either a diminutive or a pejorative sense, that is, to denote something small and therefore dear to the speaker, or small and therefore below the speaker. The lessons are not clear that if a noun follows a number greater than one, the noun still does not take a plural form, but it does list d‘ana/دﺎﻧﮫ as a count noun. There are other count nouns that go without mention in the text. The author also emphasizes that double negatives are necessary in Persian, which is important for English-speaking students to learn.
Title: Spoken Dari  
Author: Mustafa A. Sayd

With this text, the author intends to introduce the reader to spoken Dari through authentic conversations held between native speakers and transcribed phonetically in the native script. It was intended for use by those who already have a background in the written language and were interested in using this book in tandem with a set of CDs to explore the spoken form of Persian in Afghanistan. In the introduction to the book, the author provides a grammar section early on, showing the alterations between the written and the spoken languages.

This compilation focuses on the Dari spoken among Afghans and uses authentic texts. The author claims that these conversations have been transcribed accurately in the native script so that the reader can not only hear, but also see the differences between the spoken and written language. The author does do this to some extent, but there are spelling inconsistencies and some transcriptions that reflect the written form as opposed to the spoken form, so the transcriptions are not entirely reliable as accurate representations of the spoken language. On the other hand, the speakers on the recordings do represent multiple dialects of Dari giving the reader more exposure to variations in the spoken language. The transcribed selections include mostly speech from a northern speaker, a speaker from Kabul, and a speaker from Herat, but the CDs also include examples from Hazara and Turkmen speakers of Dari and both male and female speakers were used.

The discourse starts out at an intermediate level, but rises to more abstract political discourse in later selections providing a challenge for the learner. With each translation an English translation is provided that the author tries to keep as true to the word-for-word Dari transcription as possible so the reader can get a sense of word usage. Grammatical footnotes and vocabulary lists are provided in some selections as needed, along with a Dari-English glossary at the end of the book. There are no exercises included in the selections beyond following along with the recording, so there is no option for feedback.

Title: A Dictionary of Common Persian & English Verbs with Synonyms and Examples  
Author: Hooshang Amuzegar

This is a dictionary that specifically presents verbs. It is organized in a way that if a reader opens the dictionary from the left, they are presented with the English-Persian section of the dictionary and if they open from the right, the Persian-English section. The dictionary has over 10,000 verbs and contains target language synonyms in the Persian-English section, as well. The author includes some short vowels to help with pronunciation where it may be problematic. In the introduction he also claims that entries associated with a certain preposition are included, although there are many instances where prepositions are not included. This leaves the reader to assume that many verbs take a direct object whereas the object may require a preposition. Despite this, the dictionary is still an excellent resource for many verbs that can be very challenging to find in other dictionaries, especially more technical or high-register verbs which are so often missing from more traditional Persian dictionaries.

Title: Arabic in Persian  
Author: Azim Bayzoef

This text was developed by one of the instructors for the Persian Overseas Flagship program in Dushanbe, Tajikistan before the program was discontinued. As such it is not readily available to the public; however, it is a very useful resource, which should be available to the USG. This textbook seeks to describe the various usages of Arabic loanwords and their function in the Persian of today. The primary language of this text is Persian, but
the book is organized from left to right and parallels to examples are provided in English in some instances to help illustrate grammatical points.

Generally the material is presented within a structured framework similar to the way it would be presented in Arabic language courses. This means that tables of verbal patterns are introduced by their Arabic pattern, number, and meaning. The breakdown of all forms of Arabic loans, from participles to phrasal borrowings, is clearly described and helps introduce the reader not only to the structure of the Arabic words, but also to the triliteral root system as a whole. The goal of this is that the student would be able to process new words from one root family based off of this knowledge (i.e., knowing that علماء, علم and علم are all interrelated based on the Arabic root علّم, the three root consonants that always appear in the same order within this family).

Another aspect of this textbook that makes this textbook stand out is that the author provides tables of homographs and homophones of both Arabic and Persianate origin for the reader's comparison. This is useful for those times when a word may have two distinct pronunciations and therefore two different meanings, but are written identically. Distinguishing such words is generally problematic for L2 Persian learners. There are no exercises in this book as it was intended entirely to be a reference for advanced Persian students.

Title: Persian Verbs
Author: Azim Bayzoef

Also written by Azim Bayzoef, this textbook, like the text on the function of Arabic in Persian, was developed specifically for use in American Councils for International Education's Persian Overseas Flagship program in Dushanbe, Tajikistan and is not readily available to the public. All the same, this textbook is worthy of discussion because it focuses specifically on Persian verbs and is presented mostly in Persian for advanced learners. The author breaks verbs down into four categories.

The first group that the author delineates encompasses are what Bayzoef calls simple verbs and others call strong verbs. These verbs are the basic, single-word verbs that are often characterized by suppletive present tense stems that differ from their past tense stems. Bayzoef includes a breakdown of these verbs, their stems in the present and in the past, and even gives the reader a table of typical sound changes that occur in the present tense stems to help the reader predict stems when they don’t know. The second group is comprised of verbs that have separable prefixes that syntactically function different than the first group because of these prefixes, but otherwise share the same stems as their first group counterparts. Bayzoef dubs the third group compound verbs, but others also call them weak or light verbs, and they are always two-part words where the first part is often a noun or adjective and the second part is one of the verbs from the first group and functions as such. The author discusses the syntactic complexities of these verbs in detail and provides a good account of them that would be useful for any level of L2 learner. The fourth group of verbs is made up of verbs that can function either as a simple verb with a predictable stem or can also be separated regularly be separated into a nominal form and work as a compound verb.

This description can become complicated because of the detail and the fact that the author occasionally uses non-standard terminology for verbs (i.e., ‘indirect’ instead of ‘causative, ‘prefix’ instead of ‘preposition’). However, the textbook does thoroughly cover the form of the Persian verb, both in the written and spoken language, addressing many problem points for L2 learners such as word choice in different registers. For example, Bayzoef thoroughly explains the subjunctive and all of its forms, which is historically challenging for native English speakers. Explanations are in Persian, but parallels are provided in English in many places for comparison’s sake where possible.
Feedback from exercises is lacking, but they are nonetheless included throughout the book. For example, stories are included for reading practice throughout the text and each includes an accompanying exercise based off of the lesson or grammatical description immediately preceding it. Additionally, though, Bayzoef includes a section on idiomatic verbs where one of the words is a body part because these are widely used in Persian and are necessary for advanced learners to know. The tables include synonyms in less idiomatic speech. At the end of the text, the author has compiled a twenty page glossary of common verbs.

**Title:** Iranian Panorama (4 volumes)
**Author:** Atefeh Oliai

This series of textbooks was written and designed by Atefeh Oliai, the former academic director of American Councils for International Education’s Persian Overseas Flagship program. The author is a certified OPI (oral proficiency interview) examiner and compiled these textbooks with the ACTFL and ILR language rankings in mind. This series covers four years of Persian instruction working up from absolute beginner, to 0+/1 in the second volume, 1+/2 in the third volume, and 2+/3 in the fourth volume. Because it was developed specifically for American Councils’ overseas Persian programs, this series is not freely available to the public.

These textbooks progress through the ILR levels through each volume. Lessons generally begin with new vocabulary and synonyms before a main reading or listening portion in each lesson. These lexical exercises can be as simple as a list of key words for that lesson, or as complex as a challenge to break down a new word into its constituent prefix(es), root, and suffix(es) to help learners learn to dissect words. The feedback for these exercises is intended to be provided in class discussions, though, so an independent learner would have little to rely on. There are questions to be answered before the passage in each lesson, then questions following the reading and discussion questions to complete the unit. These prompts in the lower levels are more direct and have more explicit answers, gradually becoming more abstract and opinion or argument-based in higher level volumes.

The texts are authentic, generally written, and the instructor provides recordings of a native speaker reading each written excerpt that forms the basis of each lesson in an audio format so students can both read and hear the text and any unfamiliar vocabulary. Although the textbooks are intended to address students within a certain range, texts near the end of the each volume tend to be just beyond the upper limits for that textbook. For example, some texts at the end of the fourth year book are an ILR 3+/4, despite the book only being intended for 2+/3 students. The exercises accompanying lessons do not have any automated feedback because they are intended to be discussed in class among the students with the teacher facilitating and moderating debate. This is intended to improve speaking skills alongside reading and listening and also to improve learners’ ability to express and support a position in the L2, which is important during oral proficiency interviews. The topics discussed in these textbooks focus heavily on cultural and socio-political themes and do not equally cover different domains of the lexicon (i.e., scientific vocabulary will rarely be encountered using these materials).

**Title:** Intermediate Persian: A Grammar and Workbook
**Author:** Saeed Yousef

This book is an overview of advanced topics in Persian grammar with exercises after each grammatical structure for students to complete in writing. A key is available at the end of the book with the answers to the exercises. The authors note that the purpose of the book is to provide accessible grammatical explanations with exercises, for use in independent study or as part of a Persian class. Grammatical topics are explained in English, followed by examples in Persian with English translations. Exercises generally consist of the following types:
Title: Persian Vocabulary Acquisition  
Author: Michael Hillman

This textbook provides a pedagogical approach to acquiring new vocabulary in modern Persian. The author does more than just present vocabulary lists by giving the reader the tools to acquire a word family, not just individual tokens. For example, Hillman describes a tendency among a certain class of words such as how to build a participle out of a verb and then presents an exercise where the learner is asked to practice this skill. Hillman begins with vocabulary that was loaned from either English or French because this is the vocabulary that is most familiar to American students, and he moves on from there. He thoroughly addresses Persian word families and even has entire tables of some common words with many derived forms, such as the رسدین ‘to reach’ and many related words. Similarly, he addresses the identification and parsing of Arabic word families that have been borrowed into Persian and introduces the student to the triliteral root system so students are better able to deduce at least which part of speech new Arabic words are. Hillman did not provide an answer key to the exercises, but does reuse word families throughout the book, so it is likely that many forms are attested throughout the text.

Title: Persian Listening  
Author: Michael Hillman

This book is dedicated to listening to different kinds of Persian discourse, including jokes, monologues, telephone calls, radio broadcasts, songs, poems, and films. Each unit starts by presenting a list of vocabulary words in isolation that will be needed in the unit. The, there are several exercise for students to complete before and after listening to the text. These may consist, among others, of identifying topics heard, matching words with English translations, identifying problematic words heard in the audio, answering comprehension questions, and working with translations of the transcribed text.

Title: Persian Reading and Writing  
Author: Michael Hillman

This text begins with an extensive, explicit overview of the Persian writing system, accompanied by exercises for students to practice. The next chapter starts with some basic Persian texts and exercises, followed by a section on Persian handwriting. The final chapter continues with more Persian texts and exercises. A sample unit starts with presenting vocabulary in isolated lists with Persian and English translations. Next, there are several exercises students can complete before and after reading a text. These exercises, among others, may take the format of reading comprehension, grammatical exercises such as identifying certain word forms in the text, or translation activities. There are also grammatical explanations scattered throughout each chapter, the topics of which are based presumably on the reading.
Title: Persian Grammar and Verbs  
Author: Michael Hillman

This text follows a similar pattern to the other Hillman texts. The focus on this book is the grammatical structures of Persian. Each section is based on a part of speech, such as imperative mood verb forms, or adverbs. Each unit also has exercises scattered throughout, which may, among others, take the format of reading comprehension, grammatical exercises such as identifying certain word forms in the text, and translation activities.

Title: Persian Language (4 volumes)  
Author: Ahmad Saffar Moghaddam

This is a series of four textbooks developed and published by the Iranian government’s Council for the Diffusion of Persian Language and Literature. This series is designed to be the textbooks for four years of courses in Persian starting from absolute beginner and working up to an advanced level. The textbooks are organized from right to left, even at the lowest level. The authors make an effort to use as much Persian as possible, including in directions before exercises and in grammatical explanations, not just in the example texts. The amount of Persian used in comparison to English rises as the lessons progress until in the final book almost no English is present. These textbooks include standard short exercises and drills in listening, reading, and handwriting. The grammar is explained with as little English as possible and at higher levels use virtually no English.

This series is unique because it offers exposure to various domains of language, such as geographic, cultural, and economic terms in addition to texts that are descriptive, informative, handwritten personal letters, proverbs, and poetry. In vocabulary lists, which are included in each lesson, short vowels and diacritics are written for instructional purposes and occasionally even the stress in the Romanization. Vocabulary distinguishes part-of-speech where necessary. For example, some words can be used as either a noun or an adjective and this textbook, unlike other resources, makes this explicit. The books also address passive and active verb pairs in light verb constructions, usage of indirect verbal phrases which are widely used in Persian, and the difference in registers used to convey politeness. Lastly, these textbooks directly introduces the Tehran variety of the spoken language throughout the series, including standard sound changes that occur in conjugations, the pronunciation of the direct object marker in speech, and other common alternations. Consonants that are dropped in speech are also highlighted in red so it is clear for the learner.
Title: AZFA (Persian Language Teaching, 4 volumes)
Author: Yadollah Samareh

This is a series of textbooks developed by Professor Yadollah Samareh of the University of Tehran to teach the Persian language from the beginner level (volumes 1 and 2) to intermediate and advanced (volumes 3 and 4, respectively). These materials are entirely in Persian and are accompanied by an instructor’s guide. The intermediate and advanced level textbooks draw entirely on authentic texts from magazines, newspapers, books, and poetry. Most of these texts appear in the standard written language, but some short portions, specifically dialogue taken from stories, is written in the spoken language. Lessons begin with a text, which contains explanatory footnotes for grammar, vocabulary, and uncommon forms as needed, followed by content and comprehension questions. New words are explained in the target language using appositives with a breakdown of morphemes where necessary. Such breakdowns lead to explicit instruction of verbal paradigms and descriptions of the uses of certain prefixes and suffixes that can be used by the learner to accelerate acquisition of new vocabulary.

Each lesson also includes an analysis of the text to help explain the plot of the story or, more commonly, to help a student understand the meaning or context of a poem. Poetry is written in script to introduce the learner to different styles of written Persian, but is also rewritten in prose later in the lesson for clarity’s sake. In addition to grammatical notes, the analysis of the text may include such information as linguistic information and background on the connotations of certain words, as well as points on stylistics and register used in the text, making it very rich in information. Dr. Samareh also includes lists of synonyms and antonyms to help build the student’s vocabulary with association to words that have already been learned and often assigns a composition at the end of each lesson. After the lessons, the author has included a key for all of the drills throughout the book to allow the student or instructor to get feedback. The glossary at the end of each volume includes all short vowels and diacritics where needed.

The instructor’s guide is aligned to each lesson throughout the first two books in order to explicate for teachers exactly what the topic and grammatical points of each lesson are and how they were intended to presented in a classroom setting. As noted, this is only for initial instruction in the language and the first two volumes of the series. The guide contains only a few brief notes on how to proceed at the intermediate level and no explicit suggestions for the advanced level.
APPENDIX F: REVIEWS OF SELECTED PERSIAN REFERENCES MATERIALS

**Online Tools**

**Name:** Jahanshiri.ir  
**Web Address:** [http://www.jahanshiri.ir](http://www.jahanshiri.ir)  
**Stats:** [http://www.statscrop.com/www/jahanshiri.ir](http://www.statscrop.com/www/jahanshiri.ir)  
**Comments:**  
- The website has a Persian verb conjugator. Verbs may be entered in either Persian or Latin script. Grammatical person and tense is marked.
- Short vowels are not available for verbs entered in Persian, however the user may simply re-enter the verb in Latin script in order to find the pronunciation for each conjugation.
- Users must be sure to enter verbs like خواستن as khwâstan (infinitive) while the ‘w’ is not pronounced and may often be forgotten in transliterations.
- There are a few other resources on the website, including a Persian phrasebook, short lessons on grammar and other aspects of Persian, and a Tajiki transcriptor (converts Latin script to Tajiki Cyrillic).

**Name:** Forvo  
**Web Address:** [http://www.forvo.com](http://www.forvo.com)  
**Stats:** [http://www.statscrop.com/www/forvo.com](http://www.statscrop.com/www/forvo.com)  
**Comments:**  
- Users may search for Persian words in order to hear their pronunciations. Pronunciations are recorded by users.
- Users may sign up for free to record pronunciations for words in their native language.
- Forvo currently has the pronunciations of over 10,000 Persian words.
- Audio quality can vary a great deal from file to file.
- Forvo does not offer its own definitions of words, but does have automatic translation powered by Google Translate.
- Dialectal information can sometimes be inferred from the location of the user who pronounced a given word (their location is displayed on a map). Of course, this is not necessarily reliable since users may not speak the dialect of the area where they recorded the pronunciation (as one extreme example, the person who recorded the pronunciation for حداقل is in Uppsala, Sweden).
Name: Joint Language University (JLU)
Web Address: http://jlu.wbtrain.com/sumtotal/jlu2.0/HOME/index.asp
Comments:

- **Access:** users can register at https://www.wbtrain.com/SumTotal/JLU2.0/registration/index.asp, however, they are asked to select which organization they are affiliated with from a list that consists of government agencies, the military, and a few foreign governments (the U.K., New Zealand, Australia, and Canada).

- **Languages:** Persian Farsi, Dari. The online catalog contains sections for Area & Culture Training, Language Training, and Web Based Assessments. Within Web Based Assessments (sic), there are sections for Reading Skills and Listening Skills (there are other sections such as General Language that currently contain no items).

- **Appropriate for:** students between ILR levels 1+ and 3. The inventory of assessments includes:
  - Dari: 39 assessments at 1+/2, 40 assessments at 2+/3.
  - Persian Farsi: 20 assessments at 1+/2, 20 assessments at 2+/3.

  The reading and listening assessments consist of comprehension questions about authentic texts and audio clips. The assessments are self-guided, and only provide students with feedback on whether or not their answers were correct; they are not provided with a plan or milestones for increasing their level.

  The quality of the audio clips is variable. Many of them contain relatively high levels of white noise, or are somewhat faint.

  The description of each assessment contains a recommendation for what level the user should be in order to use it (e.g., “The passages and questions are appropriate for ILR proficiency level 3.”) The assessments score the user’s answers, but don’t diagnose their level. In fact, the description of each assessment contains a disclaimer such as “Successfully completing this Assessment Object does not guarantee that you can perform consistently at proficiency level 3.”

Name: Headstart2
Web Address: http://hs2.lingnet.org/
Comments:

- **Languages:** Persian Farsi, Dari.

- **Access:** users can register by clicking at the web address listed above. They are not required to state an affiliation with any organization. The site does request access to the user’s camera and microphone, and says that users who grant access may be recorded.

- **Appropriate for:** military personnel who need to rapidly attain elementary proficiency in Persian Farsi or Dari in order to meet basic daily needs and handle a limited set of critical situations.
UNIVERSITY OF MARYLAND CENTER FOR ADVANCED STUDY OF LANGUAGE

- Users answer questions in each module to determine whether or not they mastered the language skills and vocabulary in that module. However, the assessment is restricted to “GO” or “NO GO” (pass/fail). The student’s skills are not analyzed, and they are not given a tailored plan to address their personal problem areas (other than to presumably do and repeat all of the modules if necessary).
- The courses seemed to be designed to rapidly train users in a very limited vocabulary set of particular use in military situations. This manifests itself in examples that include but are by no means limited to:
  - Later modules require users to select the correct translation for phrases such as "ما دنیال مواد بمب سازی هستیم" ("We are searching for bomb-making materials.").
  - Grammatical instruction is deliberately limited; for example users are only taught how to give basic, informal commands in the تور form (تور is similar to tú in Spanish and other Romance languages).
  - There are virtual characters that pronounce key phrases and words, and move their mouths in order to show the approximate mouth shape needed to pronounce the items in question correctly. Many of the characters are dressed as U.S. military personnel.
  - Simplified transcription of key words and phrases is provided even up through the later modules, e.g., “مترجم کجاس؟” for مترجم کجاس؟
- The major differences between formal and colloquial Persian are de-emphasized, or perhaps not addressed at all. Headstart2 could hypothetically be used by a new, non-military learner of Persian, but he or she would probably have to move on from Headstart2 to a more advanced curriculum within a few weeks to a month (approximately).

Name: Global Online Language Support System (GLOSS)
Web Address: http://gloss.dliflc.edu/Default.aspx
Comments:
- Languages: Persian Farsi, Dari.
- Access: users can access the lessons by visiting the address above, with no registration required.
- Appropriate for: students between ILR levels 1 and 3.
- There is no assessment on this webpage per se. When one clicks on the “Diagnostic Assessment” tab, they are directed to the DLI’s ODA.
- Like material available on JLU, the material on GLOSS consists of authentic texts and audio clips that users must examine, and then answer comprehension questions. The user’s answers are scored, but the user doesn’t receive a diagnosis of their current ILR level. In fact, the user accesses material by specifically selecting the level of material they want to see.
- The main page states:
  - “In-depth Feedback accompanies all motivating tasks, providing learners with thorough explanations and tutoring just like an attentive and experienced teacher would do.”
  - However, this in-depth feedback seems to consist of built-in textual content, rather than any kind of system that analyzes and provides strategies for addressing the user’s problems areas.
- Some of the audio content suffers from the same lack of quality as the material on JLU.
Online Dictionaries

Name: Vajeh Yab (واژه یاب).
Web Address: [http://www.vajehyab.com/](http://www.vajehyab.com/)
Comments:
- Simultaneously searches for words in dictionaries/categories including لغت‌نامه‌های دهخدا (Dekhoda), فرهنگ علی مینایی (Moein), فرهنگ عمید (Amid) (a Persian-English dictionary), وارزه‌نامه آزاد, اصطلاحات, اصطلاحات عامیانه, and زبان و لحجه.
- Having results from Dekhoda, Moein, and Amid side by side allows the user to decide how thorough or complex of a definition they want to read. The definitions in Dekhoda are phrased in a more complex manner, and are accompanied by contextual examples from classical literature, while the definitions in Moein and Amid tend to be more straightforward.
- Selecting a particular result takes the user to that result’s page, which also loads a computerized voice (text to speech) that pronounces the word upon pressing the play button. The pronunciations presented are relatively rapid and unnatural, but at least give the user an idea of what short vowels are contained in the word.
- Vowel points are included in results from Dekhoda and Moein, while Amid gives pronunciation of Persian words in Latin script.
- Highlighting a word within a result triggers the appearance of a new pane, in which the definition of the highlighted word is displayed. The definitions displayed in this manner appear to be from Moein. This function is not recursive; if the user wishes to find the definition of a word in the new pane, then they must enter it into the search engine.
- Users can suggest Persian words to be added to the dictionary.

Name: Farsi123
Web Address: [http://www.farsi123.com](http://www.farsi123.com)
Comments:
- English-Persian/Persian-English dictionary.
- When users look up English words, Farsi123 offers Persian translations with contextual examples (sometimes phrases, sometimes full sentences) for clarification. For example, when one looks up “fan,” one can use the contextual examples to easily distinguish between the Persian word for “fan” as an electric appliance and “fan” as a devotee of something.
- Computerized pronunciations of English words are available, but the same is not true for Persian words.
- No vowel points are available for Persian words.
Name: Aria Dic
Web Address: http://www.ariadic.com/
Stats: http://www.statscrop.com/www/ariadic.com
Comments:
- English, French, Italian, Spanish, German, and Swedish to/from Persian dictionary.
- Context is not provided for words, which can be troublesome, or at the very least increase the amount of time needed to pin down a correct word. For example, when a user looks up the word “plant” they are given the words (factory/plant”), (plant” in the botanical sense), (to station”/”to plant”) and (”to plant” as in seeds) with no way of immediately distinguishing between them. The user therefore has to spend some time using a process of elimination to deduce that is probably the word for “plant” in the botanical sense.
- No vowel points are available for Persian words.

Name: Farsi Dictionary
Web Address: http://www.farsidic.com/
Stats: http://www.statscrop.com/www/farsidic.com
Comments:
- As with Aria Dic, context is not provided for words.
- No vowel points are available for Persian words.
- Farsi Dictionary also offers machine translation powered by Google, but this produces mixed results. For example, the sentence

“If you see my sister, tell her that I borrowed her car.”

becomes

“اگر شما خواهر من، او را بگویم که من مایلین ای قرض گرفته است.”

= ”If you my sister, I say her that I has borrowed her car.”
Name: Persian Online Dictionary
Web Address: http://www.aryanpour.com/
Stats: http://www.statscrop.com/www/aryanpour.com
Comments:
• English-Persian/Persian-English dictionary. Also offers Spanish, German, Swedish, Arabic, and French to Persian, but not vice versa.
• Some basic words like كتاب seem to not be available. Will check later to see if this was a temporary glitch.
• Vowel points are not available for Persian words.
• Includes a link to a forum on the Persian language and culture, though it does not seem to be very active.

Digital Dictionaries of South Asia

Name: Digital Dictionaries of South Asia (DSAL)
Web Address: http://dsal.uchicago.edu/dictionaries/list.html#persian
Stats: (simply shows information for Uchicago.edu) http://www.statscrop.com/www/uchicago.edu
Comments:
• Has two digitized, searchable dictionaries: Hayyim and Steingass.
• Hayyim:
  o Citation: Hayyim, Sulayman. New Persian-English dictionary, complete and modern, designed to give the English meanings of over 50,000 words, terms, idioms, and proverbs in the Persian language, as well as the transliteration of the words in English characters. Together with a sufficient treatment of all the grammatical features of the Persian Language. [Teheran, Librairie-imprimerie Béroukhim] 1934-1936.
  o The digitized Hayyim is based off of a printed version from ca. 1934, and as such does not contain many modern words (e.g., هسته‌ای, “nuclear”).
• Steingass:
  o Citation: Steingass, Francis Joseph. A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature. London: Routledge & K. Paul, 1892.
  o The digitized Steingass is based off of a printed version from 1892, and as such presents many word definitions that would be considered archaic, highly specialized, or simply not phrased in a modern manner.
  o It sometimes presents unexpected (presumably obsolete or dialectical) forms of words that are still in use. For example it gives the typical phrase ﹒bravo” as ﹒په ﹒په, which of course a typical learner would not think to look for.
The dictionary includes pronunciations of terms that do not always coincide with the “standard” Persian spoken in Tehran. For example, instead of goftan, it gives guftan, which coincides with the pronunciation commonly used in modern Tajikistan and Afghanistan. Steingass may have been basing his pronunciations on the Persian spoken in British India.

Print Dictionaries

Title: *The Large Persian-English Dictionary.*
Author: Solayman Haim (also “Soleyman” or “Soleiman”).
Publisher: Farhang Moaser Publishers, Tehran.
Year: 2009.
Pages: 1141.
Number of Entries: 50,000+.
Comments:

- Preface claims that S. Haim was the first lexicographer to introduce abbreviations for the parts of speech in Persian (which are used in this dictionary).
- Preface states that unlike past dictionaries, which tended to emphasize literary words and de-emphasize or ignore everyday, popular words, S. Haim has attempted to create a balance of both.
- Each Persian headword is accompanied by a transliteration into the Latin alphabet as a pronunciation aid.
- Many entries contain straightforward contextual examples of headwords. Author mentions in the preface that lexicographers in the past tended to use only quotes from classical literature

Author: Abbas Aryanpur Kashani and Manoochehr Aryanpur Kashani.
Commentary:

- Includes separate section on etymology
- Includes a so-called “pioneering Guide to Punctuation in Persian.”
- The English – Persian side is partially illustrated.
- On the English – Persian side, short vowels are not available for Persian words.
- Context for Persian words is also not always available on the English – Persian side.
- The two factors above require a reader to consult the English – Persian side, select a Persian word, and then look up that word on the Persian – English side to learn the word’s pronunciation and verify that the word makes sense in the context in which the reader wishes to use it.
- Parts of speech are not available for Persian headwords on the Persian – English side. The definition of course makes it clear what part of speech the headword is, but it’s not explicitly tagged.

Grammars

Title: A Grammar of Contemporary Persian.
Author: Gilbert Lazard.
Publisher: Mazda Publishers.
Year: 1992 (but see first point below).
Pages: 301.
Comments:
- This grammar is designed to be read by people that are extremely well-informed on grammatical terms. Concepts are defined comprehensively and at length, and are illustrated with plenty of contextual examples in Persian.
- Though this text is excellent for a person who aims to learn about the finer points of Persian grammar, it may seem overwhelming to the more casual learner who simply wants a quick answer to a specific grammatical question.
Title: Persian Grammar.
Author: A.K.S. Lambton.
Publisher: Cambridge University Press.
Year: 1979.
Pages: 339.

Comments:

- Comprehensive grammar that aims to give readers the ability to not only understand modern written Persian, but also be able to delve into classical texts. However, the text may seem daunting and extremely dense to someone who is not comfortable with grammatical terminology.

- There are areas in the text where literary forms seem to be given precedence over common forms, whether this was Lambton’s intention or not. For example, the present tense conjugations of **بودن** (مبین) are presented first, followed by those based on the root **بوم** (می‌باش) followed by enclitic forms (–ام) before arriving at those based on the root **هست** (می‌باشم). The system of organization is logical, but a Persian learner might mistakenly conclude that **می‌باشم** is preferable to **هست** by virtue of being mentioned first by Lambton.

- Includes ample information on the use of Arabic in Persian. Some of it may be beyond the needs of the casual student of Persian, but one particularly valuable section addresses the concept of the Arabic triliteral root and how to derive nouns from them.

- Only about two pages are devoted to spoken Persian, and even these focus largely on the phonetic changes like **آن** → **اون** that take place when shifting from written to spoken Persian.

- Some of the orthographic rules given are out of date, though still technically correct. For example, Lambton states that **آن** and **به** are usually joined to the word they precede such as in **یعنی آن** which nowadays would normally be written as **یعنی آن** or **آن** which would be written as **آن**.