PERSIAN (FARSI)

Persian is a member of the Indo-Iranian branch of the Indo-European family of languages. At present, Persian is the language of Iran, where it is almost universally spoken by everyone, including those areas of Iran where the people have a different mother tongue (Azerbaijan, Kurdistan, Luristan, etc.). A dialectal variant, often referred to as Dari, is spoken in Afghanistan, where both Persian and Pashto are considered the official languages, and where all educated persons speak it. Another dialect, similar to Dari, is Tajiki, the language of Tajikistan.

On the surface, Persian seems simple in terms of formal grammar: there is no gender either on nouns or pronouns, no noun inflection, no adjectival agreement, and a morphologically poor case system. The verb and tense system, on the other hand, is quite complicated, with many tense, aspectual, and mood forms that must be learned. The Persian verb is generally composed of compounds called light verb constructions, which are idiomatic and arbitrary and must be individually learned. For example, the combination of two words which have independent meanings, such as *zemin* 'ground' and *xordæn* 'to eat', yields the light verb *zemin xordæn* 'to fall'. The combination of the words *daraz* 'long' and *kešidæn* 'to pull' gives you the verb *daraz kešidæn* 'to lie down'. Even though there are a number of simplex verbs in Persian, it seems that these are slowly disappearing from the language, to be replaced by light verb constructions. Certainly, Tajiki and Dari are more conservative in this regard, as these dialects make colloquial use of simplex verbs that are familiar to Iranians only through poetry.

The syntactic structure of Persian is idiosyncratic. Whereas linguists refer to languages as being head-initial or head-final, depending on the position of the verb and direct object, Persian displays properties of a mixed language. For example, like the strictly head-final, subject-object-verb (SOV) language, Korean, the verb in Persian is always last in the sentence. But, the expectation is that SOV languages maintain their head-final status in all phrasal constructions. In Korean (and in almost all other head-final languages, including Japanese and Turkish), prepositional phrases have the order Noun + Preposition. Korean relative clauses are pre-nominal; that is, the clause precedes the head noun, and, as expected, lack a complementizer (that, which) in the traditional sense. The verb-final Persian does not follow these generalizations. Persian prepositional phrases have the head-initial order, Preposition + Noun,

and relative clauses are post-nominal; the head noun precedes the clause, and there is a lexical complementizer.

Another issue for Westerners learning Persian is that the vocabulary is not borrowed from the Romance languages; cognates are rare. On the other hand, many lexical items are borrowed from Arabic, and (educated) Persian speakers develop judgments about the acceptability of Arabic templatic forms, even though they are not native to Persian grammar.

Although the Persian phonetic inventory is different from Arabic, Persian is written using Arabic orthography. There are a number of Arabic characters which have lost their distinct sound. Thus, even though borrowed words are spelled with the original characters, those that denote sounds outside of the Persian inventory are pronounced as a homonym to other characters. For example, the *th* sound in 'three' [θ], does not exist in Persian, but a borrowed Arabic word is spelled with the character denoting [θ], while it is pronounced as [s]. None of the Arabic emphatic phones are pronounced in Persian, and the [q] and [γ] have fused to one sound which undergoes phonological alternation. The emphatic [H] as well as the laryngeal 'ayn' sounds also do not exist in Persian, although many words are spelled with these characters. Learning Persian spelling therefore often requires knowing Arabic root form of the word. In addition to the Arabic symbols, Persian has added four characters to its alphabet, to denote the native sounds [p], [č], [g] and [ž], which do not exist in Arabic. It is expected that speakers who learn Persian after having studied Arabic will find these divergences initially confusing. At the same time, a fore-knowledge of the Arabic lexicon and templatic forms will give learners a considerable advantage in vocabulary especially in more formal registers. A question in the initial questionnaire will identify subjects who have a background in Arabic, so that the effects of prior knowledge of Arabic on learning Persian can be assessed.

PHONOLOGY

Teachers of Persian as a Second Language have noticed consistent difficulties encountered by English speakers in the acquisition of Persian phonology and phonological rules.[1] A number of phonological items have been developed to test these claims. The tests are designed to measure

[1] According to the literature, this is in stark contrast to heritage speakers who have good knowledge of the phonetic and sound patterns of the language, at levels that second language (L2) speakers might never attain.

level of attainment, as well as to probe for patterns in the learning of specific areas and features. Test items include knowledge of stress patterns of the language (for derivational and inflectional morphology), and morpho-phonological changes in various contexts (e.g., vowel epenthesis, vowel changes, and assimilation of consonants). Subjects are asked to contrast certain phonemes in order to test their perception of sounds that may be difficult to distinguish for native speakers of English. More specifically, these include unfamiliar phones such as /q/, /ɣ/, and /x/, word final consonant clusters that are impossible in English, constraints on phonological sequences such as the ban on onset consonant clusters, and phonological processes in the grammar of Persian such as place assimilation of nasals and devoicing assimilation. All these tests include phonological items that may be challenging to English speakers because they do not occur in English or because they violate a rule of English phonology. Finally, production of lexical items allows testing of both vocabulary knowledge and the ability to produce certain Persian sounds. Native speakers are included in the pilot sample to provide baseline data.

1 Word-level stress

Persian differentiates stress patterns for derivational and inflectional morphology. Word-level stress falls on the last syllable in Persian. Since derivational morphemes are treated as part of the word, the stress shifts to the last syllable of the derivational suffix. Inflectional morphemes, on the other hand, remain outside the domain of the word, and stress does not shift. For instance, stress on the word *kučík* 'small' falls on the last syllable of the word, as shown with the accent on 'i'. The superlative suffix on adjectives is a derivational morpheme and receives the main stress in the word as in *kučiktærín* 'smallest'. However, pronominal clitics are inflectional morphemes, so adding them to the word does not create a stress shift; the stress remains on the last syllable of the original word: *kučíkešun* 'their small one'.

In addition, certain focus morphemes take the main stress in a word. For example, the negative particle is stressed when it is attached to a verb. Hence, the verb *xundém* 'I have read' becomes *nɛ́xundæm* 'I haven't read'.

It is predicted that lower proficiency speakers will not have acquired the (albeit, tacit) derivational vs. inflectional morphology distinction with respect to stress placement and may not have internalized the relationship between stress and focus particles.

1.1 Paradigm elicitation

This task uses the Paradigm Elicitation (PE) task, as described in Section 2. Subject are given a word and asked to add a derivational, inflectional, or focus morpheme. The new word produced by subjects will show a shift in stress with the addition of the derivational plural or superlative morpheme or a focus negation morpheme, and no shift in stress when the inflectional superlative morpheme is added. Since this is a production test, learners' knowledge of where to place stress in a multi-morphemic word can be tested.

There are four subtests, each using a different morpheme. Two of the subtests ask subjects to add a derivational suffix (the animate plural morpheme in subtest 1, and a superlative marker in subtest 2). A the third subtest asks for the addition of an inflectional morpheme (the pronominal clitic). The last subtest assesses stress placement with a focus morpheme (the negative prefix on verbs). Each subtest is described below.

1.1.1 Animate plural. Subjects hear a singular noun referring to an animate entity and are asked to give the plural form. Since all words in Persian can be made plural by adding the generic *-ha* suffix, subjects are instructed to avoid using the *-ha* plural form; this forces them to use the animate morpheme *-an* and its allophonic variants for this task. Since the plural is a derivational morpheme in Persian, subjects are expected to place the stress on the last syllable of the word, i.e., on the plural morpheme. This test does double duty as a morphology test, since –an has two allomorphs. There are an equal number of words that take each allomorph. Word frequency was also controlled for.

Table P1: Sample items for the animate plural subtest of the stress placement test

Persian prompt	IPA	English gloss	Expected response	IPA of Persian Plural
کارگر	kargær	worker	کارگران	kargæ**ran**
نویسنده	nevisænde	writer	نویسندەگان	nevisænde**gan**
زلزله زده	zelzelezæde	earthquakevictims	زلزله زدەگان	zelzelezæde**gan**
زورگو	zurgu	bully	زورگویان	zurgu**yan**
روحانی	ruhani	cleric	روحانیان/ روحانیون	ruhani-**un**/ruhani-**an**
خائن	xa'en	traitor	خائنان/ خائنین	xa'e**nin**/xa'e**nan**

1.1.2 Superlative. The subject hears an adjective and is asked to produce the superlative form. This suffix is derivational and triggers stress shift. Number of syllables, with an equal number of one-, two- and three-syllable adjectives, was controlled for in the adjectives.

Table P2: Sample items for the superlative suffix subtest of the stress placement test

IPA	Persian	English gloss	Expected response	Syll
Særd	سرد	cold	særdtærin	1
Xošhal	خوشحال	happy	xošhaltærin	2
mehræban	مهربان	kind	mehræbantærin	3

1.1.3. Pronominal clitic. Subjects hear a list of Persian nouns and are asked to change each word into the phrase "our _____" by adding the possessive ending *–man* or *–mun* (both acceptable pronunciations of the 1st person plural morpheme). Subject are to place the stress on the last syllable of the original word and not on the clitic, since pronominal clitics pattern with inflectional morphemes. Length and the number of morphemes used were controlled for; no compounds were used in this test. This is a double-duty test as nouns that end in a vowel other than a [i] permit the insertion of the syllable [-*ye*-] between the word and the plural, as in *mæɣaze-**ye**-man* 'our store'. This syllable is not permitted in nouns that end in a consonant or [i], as in *sandali-man* 'our chair' not **sandali-**ye**-man*.

Table P3: Sample items for the pronominal clitic subtest of the stress placement test

Prompt	IPA	Gloss	Expected response	Transcription	Syllables
فرش	færš	carpet	فرشمون / فرشمان	færšemun / færšeman	1
پدر	pedær	father	پدرمون / پدرمان	pedæremun / pedæreman	2
مغازه	mæɣaze	store	مغازه مون / مغازمان / مغازه یمان	mæɣazæmun / mæɣazeman / mæɣazeyeman	3
همسایه گان	hæmsayegan	neighbors	همسایه گانمون/ همسایه گانمان	hæmsayeganemun / hæmsayeganeman	4

1.1.4. Negation. Subjects hear a verb and are asked to provide its negative form. Stress should shift and fall on the negative morpheme. Tense, presence of auxiliaries, person, agreement, and whether the conjugated verb is a simple or light verb were controlled for.

Table P4: Sample items for the negation subtest of the stress placement test

Prompt	IPA	Expected response	Persian	Verb	Tense	Person
خوابیدم	xabidæm	næxabidæm	نخوابیدم	Simple	Past	1s
تلفن کرده بودیم	telefon kærde budim	telefon nækærde budim	تلفن نکرده بودیم	LV	PastPerf	1p
می شنید	mišenid	nemišenid	نمی شنید	Simple	PastCont	3s
نامه می نویسید	name minevisid	name neminevisid	نامه نمی نویسید	LV	Pres	2p

2 Vowel change

The Persian negation morpheme is usually pronounced /*næ*/, but when it appears before the progressive marker, it is pronounced /*ne*/. The goal is to test whether subjects know the rule for the allophonic alternation in this negation morpheme.

2.1 Paradigm elicitation

Paradigm Elicitation is used for this test. This test is run in conjunction with the Stress placement test and uses the same items as the negation morpheme subtest, described above. Subjects are given a list of Persian verbs and asked to provide the negative form of each. If there is a progressive marker *mi-* on the verb, they should use the morpheme *ne-* with the mid-vowel /e/, and if there is no progressive marker, the *næ-* form with the low vowel /*æ*/ is required. There are 28 items, split equally between verb tenses that require the *ne-* form and those that require the *næ-* form. Tense, person and agreement, and whether the verbs are simple or compound are controlled for. For sample items, see table P4.

3 Consonant clusters

Constraints on consonant cluster formation in Persian are examined, especially when they do not match constraints found in English. The goal is to determine whether learners have internalized constraints on consonant clusters, and whether English cluster rules affect the results.

In contrast to English, Persian does not allow consonant clusters in onset position (i.e., at the beginning of the word). This can be seen in words that have been borrowed from English or French. For example, *kelas* ‘class’, *perænses* ‘princess’, and *kerem* ‘cream’ have all entered Persian from French (*classe*, *princesse* and *crème*, respectively), but the pronunciation has been

modified by the insertion of a vowel in between the initial consonant sequence, as shown in bold in the words, *k**e**las*, *p**e**rænses*, *k**e**rem*. English permits many cluster combinations in onset position whereas this is completely prohibited in Persian. Many speakers elide the vowel between two initial consonants that are common onset clusters in English. For example, the sequence [f-r-] is a common English onset (Fr<u>ench fr</u>ies); a vowel between this sequence in Persian, as in, *forušga* 'store', cannot be deleted, **frušga*. This test assesses the transfer of English phonology to the pronunciation of Persian words. It is predicted that many English speakers will fail to break up the onset cluster in familiar borrowed words, and will delete the vowel between consonants that form a cluster in English.

In coda position, the situation is reversed, and Persian is more liberal than English in allowing coda (i.e., word-final) consonant clusters. Many of these are impossible in English. For example, the words *qof<u>l</u>* 'lock', *zæ<u>hr</u>* 'poison', and *bæ<u>br</u>* 'tiger' contain the word-final clusters *-fl, -hr,* and *-br,* respectively, which are prohibitted in English. Persian also has coda clusters which are prohibitted. The segments [t] and [p] are not permitted as a cluster sequence, and neither are the familiar English codas –sts and –sks (as in the words, lists and asks). It is predicted that learners will have difficulty both in pronouncing words with non-English codas, and in recognizing illegal coda clusters.

TASK: Ban on onset consonant clusters
Task-type used: Picture-word elicitation

Picture-Word Elicitation task is used. Subjects see a picture and an English translation of the target word. They pronounce the target word and their responses and reaction times are recorded.

Design

Two subtests are used to assess learners' knowledge of constraints on word-initial consonant clusters. The first subtest consists of only borrowed lexical items where the initial onset is a consonant cluster in English. It is expected that learners will fail to insert a vowel for resyllabification in borrowed lexical items where the consonants in the initial CVCV sequence are an onset cluster in English. Note that in Persian, the vowel is inserted between the first two consonants in most cases. The exception is the word initial consonant sequence [s] plus a stop.

In this case, a vowel is inserted before the two consonants, as in *eski* ‘ski’, and *espanya* ‘Spain’. Three questions are asked in this design:

1. Do learners resist making a cluster, even when shown the English translation?
2. Where is epenthesis applied, before the cluster or between consonants?
3. Which vowel is epenthesized? Are there effects of vowel harmony?

The items in the second subtest consist of native Persian lexical items in which the two consonants of the initial CVCV sequence may form a cluster in English. The hypothesis is that subjects may delete the vowel in the first syllable between the consonants when pronouncing the word if the two consonants frequently appear as onset clusters in English.

Items

In the first subtest, items are borrowed lexical items where the initial onset is a consonant cluster in English. These clusters include *bl, br, dr, fl, fr, gl, gr, kl, kr, pl, pr, sk, sp, st, tr, kr*. Number has been controlled for.

Table P5: Sample items for the borrowed lexical items subtest

English gloss	Persian test item	Expected Response	English Cluster
blouse	بلوز	boluz	bl...
France	فرانسه	færanse	fr…
brother	برادر	bæradær	br...

In the second subtest, items are native Persian words with an initial CVCV sequence, where the first two consonants form a legal cluster in English. These items are controlled by matching them against a consonantal sequence that would be illicit in English.

Table P6: Sample items for the Persian lexical items subtest

English gloss	Persian test item	Expected Response	English Cluster	Condition
tall	بلند	boland	bl…	Possible in Eng.
rice	برنج	berenj	br…	Possible in Eng.
tree	درخت	deræxt	dr…	Possible in Eng.
violet color	بنفشه	benæfš	*bn	Impossible in Eng

fountain	فواره	fævare	*fv…	Impossible in Eng
beggar/poorman	گدا/ فقیر	geda/fæqir	*gd.../*fq…	Impossible in Eng

TASK: Coda consonant clusters

Task-type used: Picture-word elicitation

Picture-Word Elicitation is used in this test. Subjects see a picture and an English translation of the target word. They pronounce the target word and their responses and reaction times are recorded.

Design

The words used in this task all contain coda clusters that are illicit in English. One of the following responses is expected:

1. Learners will epenthesize a vowel between coda consonant clusters that are illicit in English.
2. Learners may overgeneralize the ban on onset clusters in Persian to all clusters, and may disallow even those coda clusters that are licit in English.

Items

The items in this test consist of Persian words that end in a word-final consonant cluster. Whether the cluster is legal or not in English has been controlled for. For instance, *noql* 'almond candy' does not contain a valid coda cluster for English, whereas the cluster in *ræft* 'he went' is licit in English, as in the word 'soft'.

Table P7: Sample items for the coda consonant clusters test

English gloss	**Persian**	**Expected response**	**Incorrect response**
lock	قفل	qofl	*qofel
lecture	نطق	notq	*notex
ceiling	سقف	sæxf	*sæxef/*sæxfe
poison	زهر	zæhr	*zæher

TASK: Consonant clusters

Task-type used: Lexical decision

Lexical Decision is used in this test.

Design

In this test, subjects hear only non-words, half of which are possible words in Persian, and half of which are impossible for Persian because they contain illegal onset or coda clusters. Reaction time of the subjects' responses is recorded. Subjects are to recognize that each item is not a Persian word. The prediction is that subjects will take longer to say 'no' to items that obey Persian phonological rules, than to those that violate a constraint. The goal is to determine whether subjects have internalized the impossible onset and coda clusters in Persian.

Items

The following three conditions are controlled for:

1. impossible word containing an illegal coda cluster
2. impossible word containing an illegal onset cluster
3. pseudo-word (i.e., a non-existent but possible word)

Table P8: Sample items for the consonant clusters test

Non-word	Problem	Condition	Feature tested
mosks	…sks	illegal word	coda
skæp	sk…	illegal word	onset
sæxuræk	NONE	pseudoword	

4 Assimilation

A common phonological process is assimilation, where one sound is influenced and changed by a neighboring sound. There are a number of instances of assimilation in Persian, where a sound changes its characteristics based on the features of the preceding or following sound. The effect is that the pronunciation of the word does not match its orthography, creating difficulty for

second language learners who learn the language with the writing system.[2] It is predicted that L2 speakers will often fail to learn the assimilation process, retaining the written form as the phonetic output.

One example of assimilation in Persian is devoicing. Coda consonants become devoiced when immediately preceding or following a voiceless consonant. For instance, the word for 'horse' is written *æsb* but is pronounced *æsp*. This occurs because the final consonant /b/ is influenced by the preceding voiceless consonant /s/ and loses its voice feature to become the voiceless consonant /p/.

Another phonological operation is place assimilation of nasals**.** In Persian, the nasal sound /n/ that precedes a bilabial (i.e., a sound produced by putting together the two lips such as [b] or [p]) assimilates for place of articulation and is pronounced as the bilabial nasal /m/. An example is the word for 'lazy' which is written as *tænbæl*, with a coronal nasal /n/ preceding the bilabial [b]. This word is pronounced [*tæmbæl*], that is, the nasal is now pronounced as a bilabial [m].

TASK: Assimilation

Task-type used: Picture-word elicitation

Picture-Word Elicitation is used in this test. Subjects see a picture and an English translation of the target word. They pronounce the target word and their responses and reaction times are recorded.

Design

Two subtests have been devised for this test to evaluate the learner's knowledge of assimilation patterns in Persian. The first subtest assesses devoicing of consonants in coda position and the second subtest examines place assimilation of nasals.

Items

In the first subtest, items consist of phonetic sequences in which one segment undergoes devoicing. The items are verbal and nominal elements that include both simple and compound

[2] Interestingly, heritage students, generally learn the correct pronunciation and are less influenced by orthography.

nouns. The second subtest is comprised of words that contain /n/ appearing before a bilabial, a phonetic environment that triggers place assimilation. These are contrasted with an equal number of words with /n/ appearing before non-bilabial sounds (such as dentals or velars) where assimilation should not take place.

Table P9: Sample items for devoicing subtest of assimilation

English Translation	Persian test item	Transcription	Expected Response
horse	اسب	æsb	æsp
ceiling	سقف	sæqf	sæxf
record / make a recording	ضبط کردن	zæbt kærdæn	zæpt kærdæn

Table P10: Sample items for place assimilation subtest of assimilation

English Gloss	Persian test item	Transcription	Expected Response
Monday	دوشنبه	došænbe	došæmbe
lazy	تنبل	tænbæl	tæmbæl
stepmother	زن پدر	zæn pedær	zæmpedær
tooth	دندان	dændan	dændan
sugar cube	قند	qænd	qænd
prison	زندان	zendan	zendan

5 Phonological patterns distinct from English

It is predicted that English speakers will have trouble with phonemes that differ from English. One example is the coda cluster [ng]. The sequence produced is a combination of the velar nasal /ŋ/ followed by the velat voiced stop /g/. In Persian, the final /g/ is pronounced (e.g., *jæŋg* ‘war’ not *jæŋ). This differs from English where the final /g/ is never pronounced (e.g., bæŋ, not *bæŋg for “bang”, and sIŋ, not *sIŋg for “sing”). Another phenomenon is the insertion of a glide (i.e., /y/ sound) between two vowels in English, which does not occur in Persian. For example, in English the word ‘naive’ is pronounced /nayiv/ but in Persian the word *rial* ‘Rial’ is pronounced without a glide – not *[riyal] but rather [rial].

TASK: Phonological patterns

Task-type used: Picture-word elicitation

Picture-Word Elicitation is used in this test. Subjects see a picture and an English translation of the target word. They pronounce the target word and their responses and reaction times are recorded.

Design

Two subtests have been devised for this test to evaluate the learner's knowledge of phonological patterns that are distinct in Persian and English. The first subtest tests the pronunciation of /g/ following a /ŋ/ in coda cluster sequence and the second subtest examines the insertion of an intervocalic glide.

Items

The words for these tests are interspersed among all the other production tests. Items that contain these specific features are marked and tabulated separately. They do not interfere with the results of the other task.

6 Persian phonemes

Persian contains a number of phonemes that do not exist in the English repertory of sounds. Several tests have been developed to evaluate subjects' ability to recognize and pronounce these sounds. Samples of the sounds that a native English speaker may have difficulties with are the laryngeal /ɣ/ as in *ba*ɣ 'garden', the velar sounds /x/ as in *xærčæng* 'crab' and /q/ as in *qu* 'swan', and the rolled /r/ as in *ruban* 'ribbon'. Furthermore, the vowel quality of /o/, /i/, /u/ in Persian is different from English, since there is no secondary articulation in Persian.

TASK: Persian phonemes

Task-type used: Picture-word elicitation

Picture-Word Elicitation is used in this test. Subjects see a picture and an English translation of the target word. They pronounce the target word and their responses and reaction times are recorded.

Design

Subjects are asked to produce words that contain Persian phonemes that do not exist in English. The first subtest evaluates the pronunciation of the rhotic /r/ and its three allophones, and the laryngeal /γ/ and its allophones, while the second subtest tests the pronunciation of /x/.

Items

The words for these tests are interspersed among the other tests in the phonology section.

TASK: Persian phonemes

Task-type used: Lexical decision

Lexical Decision is used in this test. Subject are told to listen to a list of words which consist of real words, phonologically possible non-words (or pseudo-words) and phonologically impossible non-words. They are asked to respond when they hear a "real" word in Persian and the reaction time is recorded.

Design

The subtests in this group contain words that are pronounced with English pronunciation. It is expected that reaction time will be longer in recognizing mispronounced real-words than in recognizing words with correct Persian phones. The prediction is that only advanced students will have internalized Persian sounds as distinct from American sounds. The following are subtests in this group:

1. Vowel quality with /i/, /u/ and /o/ vowels. English, unlike Persian, has a secondary articulation, and the vowels in the items in this test are pronounced as they are in Englisah, /i^{y}/ /u^{w}/ and /o^{u}/ .
2. Quality of liquid sounds /r/ and /l/. The rolled /r/ and the soft /l/ in Persian are unlike the English rhotic /r/ and the hard /l/. Furthermore, the Persian /r/ is devoiced in coda position. The test items are pronounced as in the English.

Items

The words for these tests are interspersed among the other tests in the phonology section. Each subtest in this group contains 20 phonetically legal pseudowords, 20 phonetically illegal non-words, and 40 real words as fillers. The real words range in frequency. The following are some possible examples:

Table P11: Sample items for Persian phonemes subtest 1 (with lexical decision)

Non-word	Vowel	Position
oban	o	onset
sofake	o	nucleus
peko	o	coda
usfan	u	onset
dækuž	u	nucleus
xaču	u	coda
ikmat	i	onset
zibar	i	nucleus
moqæši	i	coda

Table P12: Sample items for Persian phonemes subtest 2 (with lexical decision)

Non-word	Sound testing	Position
rabun	r	word-onset
mobqar	r	word-coda
lavedun	l	word-onset
pefsil	l	word-coda

TASK: Persian phonemes

Task-type used: Phoneme monitoring

In this Phoneme Monitoring test, the subjects are told to listen for a particular phoneme in a list of words. Every time the subjects hear that phoneme, they should press a button.

Design

There are four subtests in this group each evaluating the subjects' knowledge of a particular phoneme by contrasting them with sounds that have similar features and that English speakers have difficulties distinguishing. The following are the list of subtests in this group:

1. Recognition of /x/ as distinguished from /k/ and /h/ in onset position and from /k/ and /ɣ/ in coda position.
2. Recognition of /q/ and its allophones as distinguished from /g/.
3. Recognition of /h/ as distinguished from /x/ in onset position and ability to discern the sound /h/ in coda position.

Items

The items contain equal numbers of words that contain the phoneme being tested and the phoneme that it is being contrasted to. In each instance, the phonemes appear in various positions within the word. The detailed numbers are listed below for each subtest:

1. Subtest 1: Recognition of /x/

   There are 20 items containing /x/ in onset position and 20 items containing /x/ in coda position. The subtest also includes 10 items containing /k/ in onset and 10 containing /k/ in coda. In addition, there are 10 items containing /h/ in onset position and including /ɣ/ in coda position.

Table P13: Sample items for Persian phonemes subtest 1 (with phoneme monitoring)

Persian Prompt	IPA	Gloss	Phoneme Position	Expected Response
حشره	hæšære	insect	onset	no
کارمند	karmænd	white-collar worker	onset	no
خرچنگ	xærčæng	crab	onset	yes
شاخ	šax	horn	coda	yes
شیک	šik	chic	coda	no
شرق	šærq	east	coda	no

2. Subtest 2: Recognition of /q/
   (a) word-level onset: 20 /g/ and 20 /q/
   (b) word-level coda: 20 /g/ and 20 /q/. Includes 10 of each as final segment in consonant clusters
   (c) word-medial onset: 10 /g/ and 10 /q/
   (d) word-medial coda: 10 /g/ and 10 /q/

Table P14: Sample items for Persian phonemes subtest 2 (with phoneme monitoring)

Persian prompt	IPA	Gloss	Phoneme Position	Expected Response
گرم	gærm	warm	onset	no
قوم	qo^{w}m	tribe	onset	yes
غروب	ɣorub	sunset	onset	yes
میگو	meygu	shrimp	medial/onset	no
لاغر	laɣær	thin	medial/onset	yes
زرنگ	zeræng	shrewd	coda-cluster	no
سنجاق	sænjaq	pin / hairpin	coda	yes
عشق	ešq	love	coda-cluster	yes

3. Recognition of /h/
    (a) word-level onset: 10 /h/ and 10 /x/
    (b) word-level coda: 10 /h/ and 10 /x/. Includes consonant clusters
    (c) word-medial onset: 10 /h/ and 10 /x/
    (d) word-medial coda: 10 /h/ and 10 /x/

Table P15: Sample items for Persian phonemes subtest 3 (with phoneme monitoring)

Persian Prompt	IPA	Gloss	Phoneme Position	Expected Response
هیکل	heykæl	body	onset	yes
ده	deh	village	coda	yes
بیحال	bihal	tired / unergetic	medial/onset	yes
بهرام	bæhram	Bahram	medial/coda	yes
خوشمزه	xošmæze	tasty	onset	no
برزخ	bærzæx	purgatory	coda	no
مسخره	mæsxære	ridiculous	medial/onset	no
اخبار	æxbar	news	medial/coda	no

MORPHOLOGY

7 Plurals

Although all nouns in Persian can be made plural by adding the plural morpheme *ha*, there exist also a number of morphemes that can be used only on certain categories of nouns. There is a special plural morpheme for animate nouns (especially for humans). The plural morpheme for animates is /*–an/*. There are two allomorphs /*–yan/* (after vowels) and /*–gan/* (for words ending in the /e/ sound). Samples of these three forms are: *zæn* 'woman' which becomes *zæn-an* 'women', *danešju* 'university student' which becomes *danešju-yan* 'university students', and *nevisænde* 'writer' which becomes *nevisænde-gan* 'writers'. In addition to this Persian plural suffix, there are plural suffixal morphemes borrowed from Arabic such as *–un* as in *ruhani-un* 'clergy' and *–in* as in *mosafer-in* 'travelers'. Note that the former is also included in the phonology test for avoidance of the intervocalic high glide. Thus, whereas the response *ruhani-**yun*** would be considered correct for the morphology task, it would be counted as an erroneous response for the phonology test for glides.

Another method of forming plurals involves the use of the Arabic templatic plural form, which has been borrowed from Arabic into Persian. There are several template forms and each Arabic noun belongs to a separate template category. For example, *fekr* 'thought' belongs to the so-called **æf'al** template forming its plural as *æfkar* 'thoughts'; but *qærn* 'century' belongs to the **fo'ul** template becoming *qorun* 'centuries' in plural form. It is believed that these rules may not be part of the morphology of Persian and therefore only speakers who have higher level knowledge of Persian would be familiar with the various Arabic template plural forms.

TASK: Plurals

Task-type used: Paradigm elicitation

Paradigm Elicitation is used to test this feature.

Design

The goal of this test is to find out whether the subjects have knowledge of the contexts in which the different plural morphemes or templates appear. There are two subtests: The first subtest is used in conjunction with the phonology 'stress placement' test. The subject hears a series of Persian nouns and is asked to make the words plural without using the *-ha* morpheme. This will force the subject to use the more specific plural suffixes. The second subtest evaluates the

subjects; knowledge of the Arabic templatic forms. Once again, the subject is asked to provide the plural form without using the *–ha* suffix. Since the given nouns cannot be made plural by adding another suffix, this forces the subject to produce a plural form by using Arabic template morphology.

Items

All items used in the first subtest are animate nouns that can be made into plural form. The items include nouns of different categories which would require a specific plural suffix. Persian animate nouns, for instance, require the use of the *–an* morpheme and its allomorphs. Arabic loan words may take *–in* or *–un*, depending on the type of noun. Animate Arabic loans may also appear with the Persian animate plural *–an*. Crucially, the Arabic suffixes can only appear on Arabic loan words and not on Persian words. There are 32 items divided equally between nouns requiring the *–an* suffix, the *–gan* suffix, the *–yan* suffix, and the Arabic suffixes (i.e., 8 of each category).

Table P16: Sample items for subtest 1 of plurals

Pers. prompt	IPA	English gloss	Expected response	IPA of Persian Plural
کارگر	kargær	worker	کارگران	kargæ**ran**
نویسنده	nevisænde	writer	نویسندەگان	nevisænde**gan**
زلزله زده	zelzelezæde	earthquakevictims	زلزله زدەگان	zelzelezæde**gan**
زورگو	zurgu	bully	زورگویان	zurgu**yan**
روحانی	ruhani	cleric	روحانیان/ روحانیون	ruhani-**un**/ruhani-**an**
خائن	xa'en	traitor	خائنان/ خائنین	xa'e**nin**/xa'e**nan**

The items used in the second subtest are Arabic nouns that are being used in Persian in both singular and plural forms. Four most commonly used Arabic template forms are tested, containing 6 items from each templatic form. These nouns do not have a plural form with the Arabic *–in* or *–un* suffixes, and therefore the Arabic templatic form must be used.

Table P17: Sample items for subtest 2 of plurals

Persian Prompt	IPA	Gloss	Expected response	Category
فکر	fekr	thought	æfkar	Template 1 (æf'al)
قرن	qærn	century	qorun	Template 2 (fo'ul)
جامعه	jame'e	society	jævame'	Template 3

مدرسه	mædrese	school	mædares	Template 4

TASK: Plurals

Task-type used: Lexical decision

Lexical Decision is used to test this feature.

Design

The goal of this test is to find out whether the subjects have knowledge of the various noun classes in Persian in terms of animacy, word origin, and selectional restrictions in forming the plural. In this test, subject is presented with a pair of words consisting of a singular noun and a possible (grammatical or ungrammatical) plural form of that noun, and he or she must determine whether the plural form in the pair is acceptable for the noun.

Items

The items include both animate and inanimate Persian nouns, as well as kind nouns. In addition 4 Arabic plural forms are being tested. There are a total of 40 items, split equally between acceptable and unacceptable plural forms. The words used in this test have not been used in the other plural tests for Persian. The detailed breakdown of the conditions is as follows:

- 8 good Arabic broken plural items
- 6 good Arabic *–at* suffixal plurals
- 1 good Arabic *–in* suffixal plurals
- 3 good Persian animate plurals
- 2 good Persian kind plurals
- 6 ungrammatical Arabic broken plurals on Persian nouns
- 1 ungrammatical Arabic *–at* plural on Persian noun
- 1 ungrammatical Arabic *–in* plural on Persian noun
- 4 ungrammatical Persian animate plural on inanimate noun
- 3 ungrammatical Persian animate nouns that disallow animate plurals
- 4 ungrammatical Arabic plural forms
- 1 ungrammatical form of Persian kind plural

Table P18: Sample items for animate plural subtest of plurals (with lexical decision)

Sing. Item	Plural Item	Singular Persian	Plural Persian	Response	Description
ozv	æ'za	عضو	اعضاء	yes	Arabic broken pl
dælil	dælayel	دلیل	دلایل	yes	Arabic broken pl
kælæme	kælæmat	کلمه	کلمات	yes	Arabic suffix -at pl
kudæk	kudækan	کودک	کودکان	yes	Persian animate pl
karxane	karxanejat	کارخانه	کارخانه‌جات	yes	Persian kind pl
pænjære	pænajer	پنجره	پناجر *	no	Arabic broken pl on Persian
qali	qaliat	قالی	قالیات *	no	Arabic pl -at on Persian N
heyvan	heyvanin	حیوان	حیوانین *	no	Arabic pl -in on Persian N
kuze	kuzegan	کوزه	کوزه‌گان *	no	Persian animate pl on inanimate N
loɣæt	loɣæl	لغت	لغل *	no	wrong Arabic pl

8 Verbal negation

The negative verbal prefix *næ-* or *ne-* appears on the beginning of the conjugated verbal unit, hence in simple verbs, it attaches to the beginning of the whole verb form whereas in compound verbs, it appears on the beginning of the light verb and not on the preverbal element. For instance, negation attaches to the beginning of simple verbs, hence *xabidæm* 'I slept' becomes ***næ****xabidæm* 'I did not sleep', and *dævide budid* 'you had run' becomes ***næ****dævide budid* 'you had not run'. In the future tense, the negative suffix appears in the beginning also as in ***næ****xahæm ræft* 'I will not go'. In compound verbs, negation appears on the light verb element and not on the preverb as in *hærf mizædi* 'you were talking' becoming *hærf* ***ne****mizædi* 'you were not talking'.

TASK: Verbal negation

Task-type used: Paradigm elicitation

Paradigm Elicitation is used to test this feature.

Design

The goal of this test is to find out whether the subjects have knowledge of the position where the verbal negative morpheme appears. This test is used in conjunction with the phonology test **Error! Reference source not found.**. The subject hears a series of Persian verbs and is asked to provide the negative form.

Items

The items were controlled for tense (with and without auxiliary), person and agreement, and whether the conjugated verb is in simple or compound form.

Table P19: Sample items for verbal negation

Persian Prompt	IPA	Expected response	Persian	Verb	Tense	Person
خوابیدم	xabidæm	næxabidæm	نخوابیدم	Simple	Past	1s
تلفن کرده بودیم	telefon kærde budim	telefon nækærde budim	تلفن نکرده بودیم	LV	PastPerf	1p
می شنید	mišenid	nemišenid	نمی شنید	Simple	PastCont	3s
نامه می نویسید	name minevisid	name neminevisid	نامه نمی نویسید	LV	Pres	2p

9 Agentive nouns

In Persian, agentive nouns can be formed by adding suffixes to the present stem of the verb or another noun. There are in particular three suffixes that denote agenthood and these are *–či, -kar*, and *-gær.* For example, the suffix *–gær* can be added to the noun *kar* ‘work’ to form *kargær* ‘worker’.

TASK: Agentive nouns

Task-type used: Lexical decision

Lexical Decision is used to test this feature.

Design

The goal of this test is to find out whether the subjects have knowledge of the morphology for forming agentive nouns, targeting the three common suffixes *–či, -kar*, and *-gær*.

Items

Items are divided into an equal number of grammatical and ungrammatical agentive nouns using one of the three suffixes as the correct response. In other words, 20 items were developed for each suffix, comprising 10 acceptable and 10 unacceptable agentive nouns. Items range in frequency.

Table P20: Sample items for agentive noun

Persian prompt	**IPA**	**English gloss**	**Condition**	**Expected response**
پست‌چی	post-či	mailman	-či	acceptable
* درچی	dær-či	*Intent: *doorman*	-či	unacceptable
جنایت‌کار	jenayæt-kar	murderer	-kar	acceptable
* گزارش‌کار	gozareš-kar	*Intent: *reporter*	-kar	unacceptable
رُفت‌گر	roft-gær	street-sweeper	-gær	acceptable
* استخوان گر	ostexan-gær	*Intent:*bone-setter*	-gær	unacceptable

10 Negation

There are several prefixes for negating a noun or an adjective in Persian. Each prefix appears on a specific noun or adjective class and is therefore constrained in its use. The main prefixes for negating are the following:

- *ɣeyr*, used on Arabic adjectives and denoting ‘in/im-’ or ‘not’
- *zed*, used on nouns and meaning ‘anti’
- *na*, used with Persian and Arabic adjectives
- *bi*, used with Persian and Arabic adjectives and denoting ‘-less’
- *næ*, used on present stems of verbs meaning ‘un-’
- *bedun*, used before a noun and meaning ‘without’

TASK: Negation

Task-type used: Lexical decision

Lexical Decision is used in this test.

Design

The goal of this test is to find out whether subjects have knowledge of the morpho-lexical selection rules for negating nouns (including derived nouns) and adjectives.

Items

Six morphemes of negation were chosen and divided into an equal number of acceptable and unacceptable uses. Frequency was controlled for such that for each prefix, there are 3 acceptable and 3 unacceptable forms in the high frequency range, and 1 acceptable and 1 unacceptable form in the low/mid frequency range, totaling 48 items.

Table P21: Sample items for the negation morphemes test

Persian Prompt	IPA	English gloss	Lexeme Tested	Expected Response
بی‌مزه	bimæze	tasteless	*bi*	good
* بی‌شناس	* bišenas	*Intent: unknown*	*bi*	bad
ناراحت	narahæt	uncomfortable	*na*	good
* نانمک	* nanæmæk	*Intent: not salty; insipid*	*na*	bad
غیر قانونی	γeyreqanuni	illegal	*γeyr*	good
* غیرعرضه	* γeyreorze	*Intent: inept*	*γeyr*	bad
ضد دولتی	zededolæti	anti-governmental	*zed*	good
* ضد توان	* zedetævan	*Intent: debilitating*	*zed*	bad
زبان‌نفهم	zæbannæfæhm	animal (w/out language)	*næ*	good
* نشرف	* næšæræf	*Intent:dishonorable*	*næ*	bad
بدون شرح	bedunešærh	unxplained	*bedun*	good
* بدون ممکن	* bedunemomken	*Intent: impossible*	*bedun*	bad

11 Subject clitics

Persian colloquial speech, and in particular the Tehrani dialect, makes use of subject clitics. These are pronominal morphemes that refer back to the subject of the sentence as demonstrated below in (1).

(1) sara ræft-**eš**
Sara went-**CL.3SG**
'Sara left.' (Lit. 'Sara left-she')

The subject clitic may not be used on all verbs and its use is constrained by the type of verb and context. This feature is used only in speech and almost never in written form. It is expected that L2 speakers will be unfamiliar with the usage of the subject clitic.[3] It is predicted that even though the use of the subject clitic is for the most part limited to the dialect of Tehran, non-Tehranis will have judgments on this task. The prediction is that the results will demonstrate the divergence between non-subject-clitic dialects of native Persian speakers and non-native speakers.

TASK: Subject clitics
Task-type used: Lexical decision

Lexical Decision is used in this test.

Design
To test subjects' knowledge of the usage of the /-*eš*/ subject clitic, 60 items have been developed, split equally between grammatical and ungrammatical uses of the morpheme. Care was taken to control for the homophonous object clitic in transitive constructions. The subject clitic can only refer to a third person singular subject. Because this clitic cannot refer to plural objects, especially animate ones, all direct objects are plural. Thus, in the sentence 'the dog bit-***eš*** the boy', it is possible for the –***eš*** clitic on the verb to refer to the direct object, 'boy'. However, in the minimally different sentence 'the dog bit-***eš*** the boys', it is impossible for the clitic to have reference to the plural direct object. In this sentence, it must refer to the subject. The use of the

[3] This contrasts with heritage speakers who are expected to demonstrate better proficiency in using the subject clitic.

subject clitic is highly constrained and all items were checked for acceptability by native speakers.

Items

Verb tense, negation and verb class (such as unaccusative, unergative or transitive) have been controlled for. Other considerations in the design of the items include sentence length and whether the verb is simple or compound. Person and number on the subject have been considered, and direct objects of transitive verbs are plural animates which disallow the singular clitic.

Table P22: Sample items for the subject clitic test

Accept.	Persian prompt	English gloss
good	محمود کو؟ قبل از این که من بیام، رفته بودش	Where's Mamoud? He had left-Cl before I came
good	اون باید می‌گفت که امروز ممکنه نیادش	He should have said that he may not come-Cl today
good	اون‌یکی خوشگل نیستش	That one Neg-is-Cl [isn't-Cl] handsome
* Bad	دخترایی که اومده بودن، فوری رفتنش	The girls who had come, left-Cl quickly
* Bad	سهراب داره میافتدش	Sohrab was in the process of falling-Cl

12 Subject-verb agreement

In Persian, the subject and verb generally agree in number. However, if the subject is inanimate, agreement is optional. For example, a plural inanimate subject can appear with a singular agreement on the verb as shown:

(2) lebasaš kæsif miše
clothes-his dirty becomes
'His clothes are getting dirty.' (Lit. 'His clothes is getting dirty.')

English does not permit optional agreement, and animacy distinctions in the grammar are rare (with the exception of "it" which generally cannot refer to animate/human nominals). This

feature therefore involves two-part learning: optionality in verbal number agreement, and an animacy distinction.[4]

TASK: Subject-verb agreement

Task-type used: Grammaticality judgment

Grammaticality Judgment is used to test this feature.

Design

Subjects hear sentences with plural subjects and the verb with or without plural agreement and must determine whether the sentence is grammatical or not. The conditions are the following:

	Plural agreement	**Singular agreement**
Animate subject	6 items	6 items
Inanimate subject	6 items	6 items

Items

The items were controlled for frequency and include 2 each of high, medium, and low frequency words in each cell shown above. Each item includes sentences of 6 and 7 words. All cells are matched in sentence length and word frequency.

Table P23: Sample items for subject-verb agreement

Persian Prompt	**English gloss**	**Conditions**	**Freq**	**Expected Response**
لیوان‌های قرمز از بوداپست خریده شدن.	The red glasses were purchased in Budapest.	inanimate/ +Agr	high	acceptable
بخش‌های صنعتی در اواخر امسال افتتاح می‌شود.	The industrial departments will be inaugurated at the end of this year.	inanimate/ noAgr	low	acceptable
مدیران داخلی با منشی جدید ملاقات کرد. *	*The internal managers interviewed the new secretary	animate/ noAgr	mid	unacceptable
مسافرا بار رو توی مینی‌بوس گذاشتن.	The passengers put their baggage in the minibus.	animate/ +Agr	high	acceptable

[4] This is another area where better performance is expected of heritage speakers.

13 Causation

Causation can be expressed either morphologically using a causative morpheme on the present stem of the verb, or by using a causative light verb. The strategy used depends on the verb type and is strongly constrained. For instance, the verb *tærsidæn* 'to fear' can become a causative by adding the morpheme *–an* shown in bold in *tærs**an**dæn* 'to scare/to cause fear'. But for the verb 'to open, the choice of the light verb determines if the verb is intransitive or transitive/causative:

(3) dær baz **miše**
door open becomes
'The door is opening.'

(4) dæro baz **mikone**
door open makes
'(He/she) is opening the door.'

The recognition of causation has been known to be difficult for students. The hypothesis is that L2 and heritage speakers might have different rates and levels of success in acquiring semantic, syntacic, lexical, and morphological ways of expressing causation.

TASK: Causation
Task-type used: Grammaticality judgment

Grammaticality Judgment is used to test this feature.

Design
Subjects hear sentences with causative and non-causative verb forms and they need to determine whether the sentence is grammatical or not. Half of the sentences are constructed using simple verbs (which may take a causative morpheme), and the others are formed using light verbs. The same verb is used in the following conditions:

(a) causative and acceptable

(b) causative and not acceptable (e.g., used in intransitive sentence)

(c) non-causative and acceptable

(d) non-causative and not acceptable (e.g., used in transitive sentence)

In addition, 5 items with causative interpretation in simple verbs are being piloted. They are: *suxtæn* 'burn', *pušidæn* 'wear, cover', and *pičidæn* 'wrap' without the overt causative morpheme. These verbs can take the causative morpheme, but due to language change, they also seem to convey both a causative and a non-causative meaning without having to add the causative affix.

Items

The conditions are the following:

	Causative & Acceptable	**Causative & Unacceptable**	**Non-Causative & Acceptable**	**Non-Causative & Unacceptable**
Simple Verb	10 items	10 items	10 items	10 items
Light Verb construction	10 items	10 items	10 items	10 items

Table P24: Sample items for causation

Persian prompt	**English gloss**	**Condition**	**Grammatical?**
صدای هواپیما همیشه مریمو می ترسونه	The sound of a plane always frightens me.	causative simple verb	Yes
پسرتون از این جور فیلما می ترسونه؟ *	* Does your son frighten of such movies?	causative simple verb	No
بچه ها از این معلم می ترسن	The children are afraid of this teacher.	non-causative simple verb	Yes
فیلمای الفرد هیچکاک منو می ترسن *	* Alfred Hitchcock's movies are afraid me.	non-causative simple verb	No
این چند خط شعر منو به گریه انداخت	These few lines of poetry made me cry.	causative light verb construction	Yes
شادی همین که فیلم رو دید، به گریه انداخت *	* As soon as she saw this movie, Shadi made cry.	causative light verb construction	No
وقتی که شاگردا نمره هاشونو دیدن، به گریه افتادن	When the students saw their grades, they started to cry	non-causative light verb construction	Yes

* پسرت خواهرشو دوباره به گریه افتاد	* Your son started to cry her sister again	non-causative light verb construction	No

14 Arabic template morphology

Persian makes use of certain Arabic words constructed based on the template morphology of the language. Hence, a number of words can be formed based on the same root; these words share certain orthographic features and in most cases, they also have certain common meanings. For example, the root *še'r* 'poetry' also forms the words *ša'er* 'poet' and *moša'ere* 'poetical contest'. Native speakers can generally distinguish related words derived from the same Arabic root, which allows for a higher vocabulary knowledge and better spelling. It is, however, not clear whether heritage or L2 speakers would also be discern words related to the same root form.

TASK: Arabic template morphology
Task-type used: Multiple choice

Multiple Choice is used to test this feature.

Design
Subject is presented with 3 words from the same Arabic root, 2 of which are common and one is low-frequency, as well as one phonetically similar distracter. The subject must identify the unrelated form. This task is also a test of lexical knowledge as it is almost impossible to handle without a fairly advanced level of vocabulary.

Items
There are a total of 17 items with 4 words in each item. Three of the four words are derived from the same root, two of those being more common, one less so. The fourth choice is a distracter which is phonetically similar but unrelated to the Arabic root. The words in each item have been controlled for frequency. An example of an unrelated word is as follows: The Arabic root for "exit" is the sequence [x-r-j], which give us the three words ***exraj***, ***xoruj***, *mæ**xræj***. The unrelated word given for this sequence is *xejalæt* 'embarrassment' which does not contain the 3-vowel sequence of the previous three tokens.

Table P25: Sample items for the Arabic templatic morphology test

IPA 1	Gloss	Form 2	IPA 2	Distractor	IPA	Outsider	IPA
šaer	poet	مشاعره	moša'ere	شعور	šo'ur	معاشرت	mo'ašeræt
mæskæn	domicile	اسكان	eskan	مسكّن	mosækken	نسكافه	neskafe
exraj	excise	خروج	xoruj	مخرج	mæxræj	خجالت	xejalæt

15 Deverbal nouns

Each verb in Persian has a present stem and a past stem, which are used in forming different tenses. Certain nouns in Persian are formed on the past stem of the verb as in *goftar* 'saying' (formed on *goft*), while others are formed on the present stem as in *ranænde* 'driver' (formed on *ran*) or *jæheš* 'jumping' (formed on *jæh*). In addition to this strategy, certain affixal morphemes can also be used to denote thematic categories such as person, place, or thing. For example, the word *azmayešgah* 'laboratory' is actually formed on the present stem of the verb *azmudæn* 'experiment, try', followed by the nominalizing morpheme *–eš*, and followed by the suffix *–gah* which denotes place. Thus the word *azmayešgah* can really be analyzed as 'a place for doing experimentation'.

The deverbal noun elements targeted in this test are listed below and the relevant affix for each category is shown in bold letters:

(1) Agentive noun:	*ran* 'drive' + ***ænde***	*ran**ænde***	'driver'
(2) Place of:	*xab* 'sleep' + ***gah***	*xab**gah***	'dormitory'
(3) x-able:	*xordæn* 'eat' + ***i***	*xordæn**i***	'edible'
(4) Patient / participle:	*mord* 'die' + ***e***	*mord**e***	'dead'
(5) Action noun:	*jæh* 'jump' + ***eš***	*jæh**eš***	'(the act of) jumping'

TASK: Deverbal nouns

Task-type used: Multiple choice

Multiple Choice is used to test this feature.

Design

The goal of this test is to see whether subjects can guess the meaning of a non-word based on the structure of the word and the affixes within it. The subject is given a prompt such as "One who does _____" using the appropriate form of a non-word in the blank. He or she is then given 5 possible non-words and must choose the one with the correct form to match the role denoted in the prompt (in this case, agentive). For example, given the prompt "A place where one **kurupid**s" and the 5 choices of *kurupgah*, *kurupide, kurupidæni, kurupeš, kurupænde*, the subject must choose the first one *kurupgah*, since that is the only one that might refer to that meaning. Note that *kurupid* is a non-word and the meanings can only be guessed based on the form of the word and the affixes used.

Items

Each item includes one sentence prompt in Persian describing the meaning targeted, followed by 5 choices that the subject needs to select from. Each choice uses one of the forms described in the feature description above: **agentive noun**, **place of**, **x-able**, **participle**, and **action noun**.

Table P26: Sample items for deverbal nouns. The correct choice is shown in bold.

English Gloss of Persian prompt	Choice 1	Choice 2	Choice 3	Choice 4	Choice 5
Something that is "bulid-able" is …	bulešgah	bulide	buleš	***bulidæni***	bulænde
Someone who "mikamæd"-s is ...	kabešgah	kabide	kabeš	kabidæni	***kabænde***
The place where one "kurupid"-s is …	***kurupešgah***	kurupide	kurupeš	kurupidæni	kurupænde
The act of "mæršidæn" …	mæršešgah	mæršide	***mæršeš***	mæršidæni	mæršænde

SYNTAX

16 Argument structure

Argument structure of the verb plays an important role in Persian syntax. This property is clearly exhibitied in denotations of causation, which can be by means of a light verb, by means

of a lexical "cause" verb and by a causative infixival morpheme. Thus the first subtest of arguments structure targets learners' acquisition of the possible ways to denote causation in Persian. Items have been designed to assess subjects' knowledge of simple and light verb causative constructions, and related argument structures, as well as their knowledge of verb classes.

The second subtest targets knowledge of the uses of accusative suffix *–ra*. This test evaluates subjects' knowledge of uses of this as a specific direct object marker, and as a Topic marker. Animacy distinctions, especially plus human distinctions, the structure of the noun phrase, specificity and non-specificity, topicality and telicity all play a role in the appropriate use of this morpheme.

Subtest 3 focuses on the subcategorization of prepositions. Verbs in Persian subcategorize for different prepositions. These must be individually memorized. Many are different from the English and indeed defy logical explanation. Sometimes, the English equivalent contains no preposition and the verb only takes a direct object. It is expected therefore to see transfer from English, and errors due to inadequate learning. Even heritage speakers whose Persian is advanced display transfer effects in the most commonly used expressions. In order to use preposational phrases correctly, speakers must be aware of verbal selectional rules.

Subtest 4 tests verbal aspect. It is predicted that learners will have difficulty in understanding the nuances of tense. This includes verbal aspect which is often only learned at very advanced levels. This test will define the levels at which students of Persian acquire verbal aspect.

TASK: Argument structure

Task-type used: Grammaticality judgment

Grammaticality Judgment is used in this test.

Design

The three ways of denoting causation have been selected. A fourth condition is a light verb construction with no causative meaning. In addition, there is a newly developing phenomenon in the language, non-causative verbs which now denote causation. An example is the verb *suxtæn* 'to burn' in the sense of 'something is burning'. In order to denote 'I burned the food', a causative morpheme is used. This verb is undergoing a change, however, and can now be used

in its simple form to mean “burning something”. The test includes a number of these items to compare with native judgments.

Items

There are four types of items with 10 acceptable and 10 unacceptable samples of each and a fifth type of the newly developing uses of unaccusative verbs. These item types are listed below. Sample items appear in Table P27.

Item Types in the Test for Causative Constructions

1. causative simple verb N=20
2. non-causative simple verb N=20
3. causative light verb construction N=20
4. non-causative light verb construction N=20
5. test of (newly developing) causation on non-causative verb N=5

Table P27: Sample items for the causation subtest of argument structure

Persian prompt	English gloss	Condition	Grammatical?
صدای عصبانی بابابزرگ بچه ها رو از ترس لرزوند	The sound of grandpa's angry voice made the children tremble with fear.	causative simple verb	Yes
دیدم که شونه هاش رو می لرزوندن *	* I saw that his shoulders were making tremble.	causative simple verb	No
پیرزن تو برف مونده بود و از سرما داش می لرزید	The old woman was stuck in the snow and was trembling from the cold.	non-causative simple verb	Yes
زلزله های اخیر تموم استان مرکزی رو لرزیدن *	* The latest earthquakes shook all of the central province.	non-causative simple verb	No
خیلی تاریک بود. من رفتم چراغو روشن کردم	It was very dark. I went and turned the light on.	causative light verb construction	Yes
چراغای خیابونا سر ساعت هشت روشن می کنن *	* The streetlights turn on at exactly 8 o'clock	causative light verb construction	No
هر چی می کردم، سیگارم روشن نمی شد	Whatever I did, my cigarette wouldn't light up	non-causative light verb construction	Yes
عصر که رسیدیم خونه، تلوزیون رو روشن شدیم *	* When we arrived home in the evening, we turned on the TV	non-causative light verb construction	No

TASK: Argument structure

Task-type used: Grammaticality judgment

Grammaticality Judgment is used for this test.

Design

The use of the accusative morpheme is complex. As a direct object marker, it is sensitive to specificity, animacy, and the noun phrase structure of the object. For example, pronouns, referential expressions, and possessor-possessee constructions as direct objects must be marked with /-ra/. In addition, certain Topic expression can be marked with this morpheme, but only if they are telic. These conditions have been controlled for.

Items

The test is comprised of acceptable and unacceptable samples with and without the /–ra/ morpheme in the following contexts:

1. on specific direct object
2. on clausal direct object
3. as topic marker on time adverbial
4. as topic marker on telic location adverbial
5. ungrammatical on non-specific direct object
6. ungrammatical on indirect object
7. ungrammatical on PP

Sample items appear in Table P28.

Table P28: Sample items for 'ra' affix subtest of argument structure

Persian prompt	English Gloss	G/J	/ra/?
هر چی می گردم، کیفم رو پیدا نمی کنم.	I keep searching, but cannot find my bag.	good	w/-ra
می تونم قرض کنم اون کتاب؟ *	Can I borrow that book?	bad	no -ra
تعطیلات رو رفتیم اروپا.	For vacation, we went to Europe.	good	w/-ra
همیشه از همین درو بیرون می رم. *	I always go out from this door.	bad	w/-ra
من با بچه ها رفتم مدرسه.	I went to school with the kids.	good	no -ra

TASK: Argument structure

Task-type used: Fill in the beep

Fill in the Beep is used to test this feature. Subjects hear a sentence with a sound (a beep, for example) denoting the position of a missing preposition. Subjects must supply the missing preposition.

Design

Many verbs select for a preposition. The verb "answer" in English selects for a direct object, whereas its synonym "respond" selects for a prepositional phrase "to someone/something". Often the preposition selected by the verb is arbitrary. Some verbs require the same preposition in English and Persian, while other verbs select for no or a different preposition in one language vs. the other.

Items

Two items have been designed for each for the following verbs with their corresponding Persian prepositional equivalent. It was noted whether the Persian form was similar to English or different. This gives us a subset-superset scenario. The goal is to determine if learners have more difficulty in Persian items that appear as a superset to English or vice-versa. The format of the test items appears below, in Table P29.

Verb/Predicate Phrase	Persian preposition required
• enjoy	from
• hosted	from
• borrow from	from
• to be deprived of	from
• take advantage of	from
• say good bye to	with
• compare with/to	with
• answer/ respond to	to
• belong to	to
• interested in	to
• ajudicated/ found guilty	to
• at the request of	to
• in Persian	to

• participate in	in
• about a subject	about
• on the phone	at
• on/as schedule	on
• on/as schedule	on
• on their own feet	on
• to work	head/top
• on time	head/top
• in a good mood	in

Table P29: Sample items for the preposition subcategorization subtest of argument structure

IPA of Persian prompt	English gloss	Persian Preposition	Condition to English
dustaye æfγanemun če xub [BEEP] ma pæziraii kærdæn	How well our Afghan friends hosted [**x**] us!	from/of	Diff
enšatuno [BEEP] ingilisi benevisin	Write your essay [**in**] English.	to	Diff
to dari [BEEP] mohæbate mæn su'estefade mikoni	You are taking advantage [**of**] my affection	from/of	Same
kešværhaye orupaii æm [BEEP] in næmayešga šerkæt mikonæn	The European nations will also be participating [**in**] this exhibition	in	Same

TASK: Argument structure

Task-type used: Multiple choice

Multiple Choice is used to test this feature. Subjects hear a sentence and must choose the English translation from a choice of three English sentences.

Design

Verbal aspect in Persian denotes telicity. Without an understanding of verbal aspect comprehension is distorted. Persian makes use of simple and compound verbs to denote aspect and these have been controlled for in this test. Verbs using an unrelated structure were also included in this test to factor out a chance probability, and to identify those subjects that do not have ability in identifying the aspect denoted by the verb phrase.

Items

An equal number of items were chosen in the following conditions:

- 6 items with Verbal aspect denoting telicity
- 6 items with Verbal aspect denoting atelicity
- 6 items that are unrelated but use similar structure

Table P30: Sample items for the verb aspect subtest of argument structure. The correct choice is shown in bold.

Persian prompt	English translation (A)	English translation (B)	English translation (C)
دست هوشنگ درد گرفت.	**Hooshang's hand started hurting.**	Hooshang was hurting in the hand.	Hooshang took something in his hands.
نرگس به سختی نفس می کشید.	Narges took a deep breath.	**Narges was breathing with difficulty.**	Narges sucked it out with difficulty.
دختر برادرم افتاد و گریه کرد.	My brother's daughter started crying.	My brother's daughter is crying.	**My brother's daughter fell down and cried.**

17 Complex sentences

Knowledge of a language must include the ability to parse embedded structures. These structures are not only longer in length but also create syntactic complexity by making tense requirements on the embedded verb. In addition, verbs select for different subordinating elements which are often arbitrary.

The first subtest of complex sentences target tense sequences.
Feature tested: Tense dependencies in multi-clausal sentences

The second subtest targets subjects ability to choose the correct subordination conjunct.
Feature tested: Learner's knowledge of appropriate conjunct to introduce subordinate clause in bi-clausal sentences. Multiple choice test translating the English equivalent.

Subtest 3 is a general test to guage subjects proficiency in simple vs. complex Sentences.
Feature tested: Level of acquisition based on sentence length and sentence complexity. Paired with Korean test to evaluate possible similarities in responses.

TASK: Complex sentences

Task-type used: Grammaticality judgment

Grammaticality Judgment is used to test this feature.

Design

Conditions:

Tense sequences that are parallel to English – grammatical items

Tense sequences that are parallel to English – ungrammatical items

Tense sequences that are unlike English – grammatical items

Tense sequences that are unlike English – ungrammatical items

Items

Total number of items: 125

Table P31: Sample items for sequence of tenses subtest. The * marks the sentences that are ungrammatical in Persian.

Persian sentence	Cond	English gloss
می‌گه زندگیش سخت می‌گذره.	1	S/he says her life is hard
اگه می‌ری، اینم بده به رضا!	1	If you are going, give this to Reza
گزارش دادن خرسی که دیروز به مزرعه اومده بود، رفته.	1	They reported that the bear that
وقتی به دانشگا بریم، لباس رسمی پوشیدیم. *	2	*When we went to university, we wore formal (official) dress
تو رستوران منتظر شدیم تا مامان می‌رسید. *	2	*We waited in the restaurant until mom arrived
هر وقت علی رو دیدین، به ما بگین!	3	Whenever you saw Ali, tell me.
تا غذاتو تموم نمی‌کنی، از سر میز بلند نشو! *	4	*Until you don't finish your food, don't get up from the table.

TASK: Complex sentences

Task-type used: Multiple choice

Multiple Choice is used to test this feature. You will see an English sentence where one word is underlined. You will then hear 4 Persian prompts. Choose the Persian item that fits best as the translation of the underlined word in English. Press 1, 2, 3, or 4 to denote the number of the response you choose.

Design

Conditions: subset/superset items of the following Persian conjuncts

- ægær
- ægær če
- æz anja ke
- aya N=2
- ba in ke N=2
- bæd æz in ke
- bælke
- bæraye in ke
- be mæhze in ke
- čenan ke
- dær suræti ke
- dær zemn in ke
- engar ke
- hæman tori ke
- hæmin ke
- ke N=2
- mægær in ke
- mesle in ke
- qæb æz in ke
- ta N=5
- vægærna
- væli
- væqti ke

Equated with the following English conjuncts:

- after
- although
- as
- as if
- as long as
- as soon as
- because
- before
- but

- by the time
- even though
- if
- in order to
- in the event that
- not only but also
- or (else)
- since
- so that
- that
- unless
- when
- whether
- while

Items

Total number of items: 25

Table P32: Sample items for the conjunction choice subtest

English prompt	Expected Response	Distracter 1	Distracter 2	Distracter 3
Since I was tired, I decided to stay home.	æz anja ke	ba vojude in ke	kaš	ta
As soon as he saw me, he hid in the room.	ta	ba an ke	hær væx	čon
Although they were poor, they prepared a sumptuous meal for us.	ba in ke	særænjam	gozæšte æz an ke	čon ke
As long as you're in love with this man, you won't consider anyone else.	ta	čon ke	væqti ke	aya
Come early **or** you won't make it to dinner.	vægærna	mægær in ke	ta	æmma

TASK: Complex sentences

Task-type used: Elicited imitation

Elicited Imitation is used to test this feature.

Design

Conditions - Equal number of simple and complex sentences controlled for sentence length:

- 5 words
- 8 words

- 11 words

Special feature: Sentences are matched as closely as possible (including meaning) with Korean sentences to probe for possible similarities or patterns in outcome.

Items

Total number of items: 30

Table P33: Sample items for simple vs. complex sentences subtest

Persian prompt	English gloss	No. words	Condition
من شنبه واسه شام کباب خوردم.	I ate kabob for dinner Saturday.	5	Simple
اون جوونا تا ساعت دو بعد از ظهر تو اون پارک قلیون می کشیدن.	Those guys were smoking a hookah in the park until 2 o'clock in the afternoon.	8	Simple
یکشنبه ی گذشته، روز مادر، برادرم برا مامان به جای گل یه ماهی خرید.	Last Sunday on Mother's Day my brother bought my mom a fish instead of flowers.	11	Simple
بچه وقتی تو بغل مامانش بود خوابش برد.	The baby fell asleep while in his mom's arms.	5	Complex
پسرم قبل از اینکه تکلیفاشو انجام بده، وقتی داشت کارتون تماشا می کرد، خوابش برد.	My son fell asleep while watching cartoons on TV without finishing his homework.	8	Complex
به اونایی که می خوان خاطره ی خوبی از تهران داشته باشن، توصیه می کنیم که یه سری به کافه شوکا بزنن.	For those who wish to have a good memory of Tehran, we recommend that they visit Café Shouka.	11	Complex

18 Classifiers

Syntax 3.3 – Classifiers

Feature tested: knowledge of classifiers in Persian

TASK: Classifiers / Elicited imitation

Task-type used: Elicited imitation

Elicited Imitation is used.

Design

Items

Total number of items: 26

Table P 34: Sample items for the classifiers test

Persian prompt	English Translation	Classifier	Used for
از نمایشگاه کتاب امسال، ده جلد کتاب داستان خریدم.	I bought 10 books from this year's book expo.	jeld	books
چوپان ده، دو رأس گوسفند خود را از دست داد.	The village shepherd lost two of his sheep.	ræ's	cattle
شما و همسرتون چند تا بچه ی مشترک دارین؟	How many common children do you and your spouse have?	ta	count nouns
تو این پادگان هزار نفر سرباز خدمت می کنن.	One thousand soldiers serve in this base.	næfær	humans

19 Relative Clauses

Subtest1 – Relative clauses

Feature tested: Learner's ability to identify the focus of a relative clause based on the Keenan-Comrie (1977) accessibility heirarchy

Subtest 2 – Relative clauses and Binding

Feature tested: Level of acquisition based on complex structure; test of binding in relative clauses.

TASK: Relative clauses

Task-type used: Picture selection

Picture Selection is used. Each page contains a series of three pictures. As you look at the picture, you will hear a tape-recorded voice describing a person or animal in one of the three pictures. Your job is simply to click on the picture showing the person or animal described.

Design

Based on Keenan's hierarchy.

Items

Crosslinguistic test, with pictures for each sentence.
There are three practice sentences. The other items belong to the following four categories:

1. Subject gap sentences
2. Object gap sentences
3. Indirect object gap sentences
4. Gap in object of preposition (Case in other languages)

Table P35: Sample items for relative clauses

Persian Test Item	English Translation	Condition
معلمی که شاگردو نگا می‌کنه	The teacher who is looking at the student	Subject gap
شاگردیو که معلم داره نگا می‌کنه	The student that the teacher is looking at	Object gap
پاسبونی که آشپز ازش سیگارا رو می‌گیره	The policeman that the chef is getting the cigarettes from	Gap in object of prep.
زنی که مردی براش گل میخره	The woman that a man is buying flowers for	Indirect object gap

TASK: Relative clauses

Task-type used: Elicited imitation

Design

[Still to be developed]

Items

LEXIS and COLLOCATIONS

20 Light verbs

The Persian verb phrase is complex. In addition to the familiar simplex verbs both transitive and intransitive, many Persian verbs are a composite, as in light verb constructions, or a combination of verb plus preposition. In assessing a learner's ability, knowledge of collocations, verb + preposition phrase, and light verb constructions must be targeted. Tests were designed to measure level of proficiency as well as to evaluate the extent of transfer effects from English.

TASK: Light verbs / Grammaticality judgment
Task-type used: Grammaticality judgment

Grammaticality Judgment is used to test subjects' knowledge of this feature.

Design

Learners of Persian will exhibit transfer from English. This task was based on Spadarro (1996) "core over-extension" task with a number of changes with a focus on light verb constructions (e.g. "take a bath"). Another change was to make it an acceptability judgment task instead of a multiple choice, because: 1) light verb constructions would produce strong enough judgments that preclude contrast choices; and 2) multiple-choice, particularly with 4 choices, is more difficult in a purely auditory mode. Diverging from Spadarro, it was decided it was insufficient to test whether subjects could tell apart "good" target language sentences from "bad" ones, and instead opted for the more interesting issue of whether English shows transfer effects. The goal is to determine whether native English learners of the target languages will have more or less problems in identifying correct/incorrect light verb constructions, when the light verb matches or does not match the English construction.

Instructions: You will hear a series of Persian sentences. Decide if the sentence is acceptable in Persian or not and press XX if "YES" and XX if "NO".

Items

This test is comprised of a total of 80 sentences, half acceptable and half unacceptable. Whether they match or diverge from analogous English sentences was controled for in the following way:

- 20 good items where the translation matches the English usage
- 20 bad items where the correct form is a match of English, that is changed in Persian to a potentially meaningful item that has no meaning.
- 20 good items where the Persian words do not match the translation of the English words
- 20 bad items where the Persian phrase is a mismatch to English, but where the word(s) have been changed to match the English equivalent

Sample test items from each of these categories appears in Table P36. Note that the asterisk (*) on the Persian sentence denotes unacceptability.

Table P36: Sample items for the light verbs test. Light verb elements appear in bold.

Persian prompt	**English translation**	**Correct Persian verb**	**Condition**
تو جمعه بازار فروشنده ها حسابی بهمون **تخفیف دادن**.	The sellers at the Friday bazaar **gave** us a really good **discount**	give discount	matches English
توی جشن تولد فرزانه، تا تونستیم **عکس گرفتیم**.	At Farzaneh's birthday party, we **took** as many **pictures** as we could.	take pictures	matches English
در آخر هر فصل فروشگاه ها **تخفیف های** خوبی **می برن** *	[Bad sentence – Literally: At the end of the year, the stores **take** good **discounts]**	give discount	matches English
بهترین دوستم قراره تابستون **ازدواج کنه**.	My best friend is **getting married** in the summer.	do marriage	differs from English
امیدوارم تا سال دیگه هر کی آرزوشو داره، **ازدواج بگیره** *	[Bad sentence – Literally: I hope by the end of the year, whoever wishes can **take marriage]**	do marriage	differs from English

21 Ezafe

Noun compounds in Persian require a linking morpheme referred to as an *ezafe*. In fact, all noun modification requires an *ezafe* suffixal morpheme –(*y)e* on the noun. Some very high-frequency noun-noun compounds, however, are becoming lexicalized so that they do not require/permit the *ezafe*. For example, *ab(-e) portoqal*, 'orange juice' and *ab(-e) sib* 'apple juice' are both

possible without the *ezafe*, *ab*(-e) gorje farhangi* ‘tomato juice’ is not acceptable without an *ezafe* morpheme between the two nouns. Whereas native speakers have clear judgments as to which compounds no longer need to be in ezafe constructions, it is expected that only very advanced learners will be able to correctly omit the *ezafe*. Lower level speakers, on the other hand, will incorrectly delete the *ezafe*, and the expectation is that they will over-generalize and will erroneously accept many of the bad non-ezafe examples.

TASK: Ezafe

Task-type used: Lexical decision

Lexical Decision is used for this test to assess learners’ knowledge of noun compounds where the ezafe compound marker can/must be deleted.

Design

This task was based on Spadarro (1996) “word discrimination” task whereby subjects are supposed to identify morphologically plausible but non-existent words, like the English *kindwill*, *nightbreak*. It was agreed that for the purposes of our project that the teams would not limit themselves to compounding-type word formation, since the availability of that kind of morphology varied a great deal across languages. The original design was to take examples of non-existant words, mixed with an equal number of real words. It was important that the the real words share properties with the non-words. Taking the English as an example, if most of the non-words are compounds, most of the real words should be similar kinds of compounds. This task can/should be easily mixed in with other lexical decision items.

All examples have been piloted with native speakers too ensure accuracy in acceptability.

Items

All the test items appear without ezafe, half of these acceptable and half unacceptable. A sample appears in Table P37.

Table P37: Sample items for the lexicalized noun-noun compounds (ezafe) test

IPA of prompt	Persian prompt	Gloss	Acceptability
pompbenzin	پمپ بنزین	gas station	no ezafe
æsbabbazi	اسباب بازی	toy	no ezafe
*abgojefærængi	* آب گوجه‌فرنگی	tomato juice	needs ezafe
*simtelefon	* سیم تلفن	telephone wire	needs ezafe

22 Proverbs

The use of proverbs in Persian is a natural and frequent part of normal discourse. Even individuals who have little or no formal education and who may live in rural areas are familiar with and employ proverbs as analogies. A learner's use (understanding and communicating) of Persian will be impaired if s/he cannot understand the most frequently used proverbs and sayings. It is expected that learners in the U.S. will show a deficiency in their knowledge of this important communication tool, and that even the most advanced learner will have problems in this area. Acquisition f Persian cannot be complete without the ability to understand and use Persian proverbs. Thus, a test was designed to asses learners' familiarity with high-frequency Persian proverbs.

TASK: Proverbs

Task-type used: Error correction

Error Correction is used for this test.

Design

Based loosely on Spadarro (1996) "multi-word unit correction" task, subjects are told that each sentence contains a Persian proverb, but that one of the words is incorrect. This is a two-part test: first, the subject must identify the problem word, and second, give the correct word. It was decided that getting both responses would be more informative, and without both responses, this task wouldn't give us much than other tasks. In this way, learners' familiarity with often-used proverbs can be ascertained by the fact they notice an error, even if they cannot correct that error.

Items

In designing the prompts, it was important not to lose the meaning of the proverb. It was observed, for example, that changing the verb often resulted in a complete loss of meaning, such that the proverb was not recoverable. An English example is the following:

Actual proverb: A rolling stone gathers no moss.

Test Prompt: A rolling stone gathers no dust.

Expected response: 1) "dust" is wrong; 2) replace with "moss"

Bad test prompt: A dry stone gathers no moss.

In the bad test prompt, subjects frequently have trouble retrieving the proverb or even ascertaining whether the prompt was a proverb to begin with. Sample items appear in Table P38.

Table P38: Sample items for the proverbs test

IPA	Correct Word	IPA	English gloss
æz divar dær umædæn, be čah oftadæn	چاله	čale	to come out of the wall, to fall into the *well/ditch
emruz o emšæb kærdæn	فردا	færda	to make today *tonight/tomorrow
pa tu gelime kæsi kærdæn	کفش	kæfš	to put one's foot in another's *gelim/shoe

23 Idioms

Whereas most native speakers have knowledge of common idioms and collocations, L2 learners frequently fall short in this area. To be deemed a near-native or very advanced speaker, one must make use of and understand of common idiomatic phrases. This test measures the learner's familiarity with common idiomatic expressions. A double-duty test, it also measures lexis, as many of the phrases in the test items include vocabulary which may be unfamiliar.

TASK: Idioms

Task-type used: Phrase completion and translation

Phrase Completion is used for this test. Whether or not subjects were able to complete the phrase, they are asked to provide the English meaning.

Design

The goal is to assess the learner's familiarity with high-frequency idioms and collocations in Persian. This test has two parts. The first is designed to measure learners' recognition of and ability to produce collocational/idiomatic phrases. The second part is designed to evaluate vocabulary. Subjects' ability to translate the first part of these short phrases will indicate level of vocabulary, even when they may not be able to complete the idiom.

Two of the Spadaro (1996) tasks were combined for this task: In the "multi-word unit completion task, subjects hear the beginning of a multi-word unit, a collocation or idiom, and must complete it by saying the final word. (In Spadaro, subjects had to complete three-word units, where the first two words were provided and the final word was missing.) In the "multi-word unit supply" task, Spadarro's subject were given one word, and had to provide the entire phrase in which that word usually occurred, e.g. (gift of the) gab, spick (and span). Controlling idioms in one of these ways in Persian was not possible, so short multi-word units that could be consistently completed by native speakers were chosen, with a maximum of 3 or 4 lexical words (including the target word).

A number of items were tested on native speakers and only those that received immediate and accurate responses were chosen for this test. Many of the phrases in Persian are oft-used because they rhyme, for example *zærq o bærq* "thunder and lightning" and *gol o bolbol* "flowers and nightingales", or because they have a reduplicative component, as in, *bede bestun* "give [and] receive", and *duri o dusti* "distance and [preserves] friendship". Note that care was given to exclude opposites because these are paired for reasons other than their lyrical or collocational properties. So, whereas *jæng o solh* "war and peace" was excluded, *jæng o dæva* "war/fight and quarrel" (to mean "fighting and quarrerling") was included. A special note in instructions to the subject: Subject must be told that the correct response is not an opposite, such that the prompt: *jæng* 'war'does not produce the response " __ *o solh"* 'and peace'. Test items

have been piloted and the findings show that given these instructions, subjects perform fine. Sample items appear in Table P39.

An innovation was the addition of the translation portion of this test. As a follow-up to completing the phrase, subjects are asked to give a loose translation of what the expression means. This second part of the test will evaluate subjects' lexis in two ways:

1. Do subjects who successfully complete the phrase know what it means?
2. Do subjects who cannot complete the phrase know the meaning of the words in the prompt?

Table P39: Sample items for the idioms test

Persian Idiom	IPA	Prompt	Expected response
تر و تازه	tær o taze	tær o . . . تر و	تازه = taze
درب و داغون	dærb o daɣun	dærb o. . . درب و	داغون = daɣun
قاطی پاطی	qati pati	qati . . . قاطی	پاطی = pati

HOLISTIC

24 Sentence comprehension

Second language speakers are at a particular disadvantage in noisy environments. However, many real life situations, from the cocktail party to the airport, require comprehension in noise. Advanced proficiency in a language must include the ability to communicate and understand accurately in real-life environments. It has been shown that performance is affected by length of utterance, word frequency, contextual information, and noise level. These factors were controlled for in this test to assess overall competence.

Persian causative constructions, in which the verb contains an infixal morpheme, are also included in this test. This is a frequently used construction and could result in massive misunderstanding if not interpreted correctly. A less than ideal hearing environment is ideal to determine subjects' ability to pick out this morpheme.

TASK: Sentence comprehension

Task-type used: Comprehension in noise

Comprehension in Noise is used in this test. Subjects hear a sentence with some level of white noise superimposed over it and are asked to repeat the final word.

Design

Persian is a verb final language that contains many light verb constructions as well as auxiliary tense verbs. A potential problem would have been to ask subjects to repeat the final word without controlling for what the final word would be. For example, the light verb 'to fall' consists of two lexical items, *zemin* 'ground' and *xordæn* 'to eat'. Asking subjects to repeat the final word of the sentence 'the boy fell' where the sequence in Persian would be 'boy ground ate', is confusing. Should subjects respond with the final word 'ate'? Or with the verb phrase 'ground ate' which means 'fell down'? To complicate matters, auxiliaries follow the verbal complex, so that 'the boy had fallen' has the Persian sequence 'boy ground eat was'? What response would subjects be expected to provide? Only the three word verbal phrase 'ground eat was' has meaning. Any of the words individiually become non-sensical within the context of the whole sentence. For this reason, only simplex verbs and tenses with no auxiliaries were selected.

Special Design: Because both Persian and Korean are head-final languages that have simplex as well as light verbs, the items have been designed to be as nearly identical as possible in both the Persian and Korean versions of this test with the final word being a simple verb. Not only are the sentences the same in both languages (with minor exceptions due to cultural specific items like foods), but sentence length and word frequency have also been matched in each sentence. In addition, agreement, tense, and number of morphemes on the verb have been controlled for. There are a total of four main conditions. Each condition contains an equal number contextual and non-contextual verb phrases. In contextual sentences, the verb can be predicted from the rest of the sentence, as in "Hamid a question from his mother asked". In non-contextual sentences, the verb cannot be predicted from the preceding words, as in "Parvin the dishes in the box counted". It is more probable to expect that the verb will be "asked" in a sentence containing the sequence "someone, a question, from his mother ___". But one does not know what the verb might be in a sequence "someone, the dishes in the box ___". It might be "put", "placed", "bought", "liked", etc. Each of these contextual units contains 8 sentences with mid to high frequency words and two sentences containing low frequency words.

Items

Table P40: Sample items for sentence comprehension. The final word to be repeated is shown in bold.

Cond	Contxt	Freq.	English translation	Persian Sentences
1	C	mid/high	Hamid **asked** this question from his mother.	حمید این سئوالو از مادرش **پرسید**
1	C	low	The maid **hung** the socks and pants on the clothesline.	پرستار بچه جوراب‌ها و شلوارها را روی طناب **آویخت**
1	NC	mid/high	Parvin **counted** the dishes in the box.	پروین بشقابای توی جعبه رو **شمرد**
2	C	mid/high	The soup **is boiling** on the stove.	آش روی اجاق **می‌جوشه**
2	C	low	The boy fell down and **is** now **moaning**.	پسره افتاد زمین و الان داره **می‌ناله**
3	C	mid/high	our uncles **are shaving** their beard in the bathroom.	عموای من دارن ریشاشونو تو دستشویی **می‌تراشن**
4	NC	mid/high	Last night, we **gave** our friends **a ride**.	دیشب ما رفقامونو **رسوندیم**

25 Accent detection

Native speakers can tell a lot about a person's background indirectly through his speech patterns. Not only do individuals have an idiolect, the specific dialect that they speak, but they also have several registers which they use depending on the context and the audience. A person from New York City has a New York accent, and will use a different registers to speak to an elderly professor and to a teen-age gas station attendant. Just as a Bronx accent is distinctive to most Americans, native speakers averywhere can immediately identify the non-standard dialects of well-known geographic regions. Thus, Iranians can tell almost instantly if the person he is speaking to is Iranian, Afghan, or Tajiki. A native Persian speaker can also easily discern several regional accents within Iran. The ability to discern easily recognizable accents and registers is necessary to be considered an advanced or near-native speaker.

TASK: Accent detection

Task-type used: Multiple choice

Multiple Choice is used in this test. Subjects hear a passage of speech (about 10-15 seconds) are asked to respond to questions that ask something specific about the speaker's background. They respond by pressing a button that corresponds to a choice they see in English on the screen. A passage may have multiple questions, depending on how fine-grained the responses need to be, i.e. what country?, what region within that country?

Design

The choice of what factors to consider was based on how easy it would be for a native speaker to respond accurately. Thus, country of dialect was an immediate choice. Regional dialects that every Iranian is familiar with were included. In addition, Persian speakers are sensitive to ethnicity, age, level of education and strength of religious adherence because these are discernible in a person's speech. Socio-economic judgments are made within a context, and are obviously not all-or-none types of considerations. Test questions have been designed to provide two opposites and subjects are asked to decide about the speakers relative to one another. Note that the speakers chosen for the question regarding the age of the speakers both have very young voices, so it is not the voice that will determine the response but rather linguistic factors such as vocabulary and sentence structure.

Items:

Which speaker do you think has had more formal education?

Which speaker do you think is younger?

Which speaker do you think is more religious?

Note that this test is to be expanded to cover the requirements for the heritage language case study.

Items

<u>Conditions:</u>

- dialect by country
- dialect by regions with Iran
- dialect of youth vs. mature idiolect
- high education level vs. lower educational level
- degree of religious affiliation

26 Narration

The level of vocabulary, complexity of sentence structure, and sentence length in free narration is a broad indicator of linguistic competence. The more advanced a speaker is, the higher the ratio of low frequency vocabulary and sentence embeddings. Sentences tend to be longer, verb forms and tense sequences more complex. The test battery includes a test to assess general competence in free narration.

TASK: Narration

Task-type used: Picture narrative

Picture Narrative is used for this test. The test is designed to include an opportunity for both free and directed narration.

Design

Subjects are presented with a cartoon sequence and asked to relate an extended story based on the cartoons. Following the story narrative, subjects are asked to expound on the following questions, one eliciting a supported opinion and another hypothesizing:

1. Since prehistoric times, many animals have become extinct and many more are currently on the list of endangered species. Is this an inevitable process? Is it important to preserve the disappearing animal and plant life? Why or why not?

2. Do you think the girl in the story should be considered a heroine for standing up for her beliefs, or she should feel ashamed because of the ridicule and her disagreement with popular belief, and why?