

# ON PERSIAN EPENTHESIS IN SUFFIXATION: AN EXPERIMENTAL STUDY\*

*Elham Rohany Rahbar  
University of Toronto*

The goals of this work are: (i) to examine epenthesis in suffixation in Persian, a process which appears to be conditioned by vowel properties and syllable structure; and (ii) to show that it is not an active synchronic process in Persian and is limited to some frozen cases, and therefore is not dependent on vowel properties. This is important in particular because there is not agreement in the literature on whether the Persian vowel system is quantitative or qualitative (e.g., Samareh 1977, Windfuhr 1979, Najafi 2001 -see Rohany Rahbar (in progress)). And it is shown here that a process which may seem to provide an argument for quantity cannot be considered as evidence for vowel structure in Persian.

## 1. Introduction

Persian has an epenthesis process that may occur when a consonant cluster is created at a stem-suffix boundary. An epenthetic vowel (the vowel –e, and in a few cases –o or –a) may be inserted in order to break up the created consonant clusters. This process occurs only in a limited number of cases and with some stem structures. Considering *a*, *e*, *o* to be lax vowels and *a*, *i*, *u* to be tense vowels in Persian, with stems having/ending in the shape CV<sub>lax</sub>C no epenthesis occurs; while with stems having/ending in the shape CV<sub>lax</sub>CC, CV<sub>tense</sub>C, CV<sub>tense</sub>CC, epenthesis occurs in some cases and not in others.

- (1) CV<sub>lax</sub>C (no epenthesis)
  - a. gam ‘sadness’ + gin → gam.gin ‘sad’ \*ga.me.gin
  - b. ſen ‘sand’ + zar → ſen.zar ‘sandy terrain’ \*ſe.ne.zar
  - c. gol ‘flower’ + dan → gol.dan ‘vase’ \*go.le.dan
- (2) CV<sub>lax</sub>CC (epenthesis may occur)
  - a. arɕ ‘value’ + mand → ar.ɕo.mand ~ arɕ.mand ‘valued’
  - b. roft ‘past stem of *roftan* ‘to sweep’ + gar → rof.te.gar ~ roft.gar ‘street sweeper’
  - c. mehr ‘kindness’ + ban → meh.ra.ban ~ mehr.ban ‘kind’
- (3) CV<sub>lax</sub>CC (no epenthesis)
  - a. xaſm ‘anger’ + nak → xaſm.nak ‘angry’ \*xaſ.me.nak
  - b. keſt ‘farming’ + zar → keſt.zar ‘farmland’ \*keſ.te.zar
  - c. gadr ‘value’ + dan → gadr.dan ‘appreciative’ \*gad.re.dan

---

\* This research was partially supported by SSHRC grant #410-2008-2645 to Elan Dresher and Keren Rice. I am grateful to Keren Rice, my supervisor, for her ongoing help and guidance. I would also like to thank Ron Smyth for his invaluable advice on the experiment and statistics. Thank you also to Elan Dresher and Yoonjung Kang for their comments. Any errors are, of course, my own.

- (4) CV<sub>tense</sub>C (epenthesis may occur)
- kar ‘work’ + gar → ka.re.gar ~ kar.gar ‘worker’
  - pas ‘watch, guard duty’ + ban → pa.se.ban ~ pas.ban ‘policeman’
  - ruz ‘day’ + gar → ru.ze.gar ~ ruz.gar ‘times, days’
- (5) CV<sub>tense</sub>C (no epenthesis)
- kar ‘work’ + gah → kar.gah ‘workshop, atelier’ \*ka.re.gah
  - pas ‘guard duty’ + gah → pas.gah ‘police station’ \*pa.se.gah
  - rig ‘pebble’ + zar → rig.zar ‘desert’ \*ri.ge.zar
- (6) CV<sub>tense</sub>CC (epenthesis may occur)
- mand ‘past stem of *mandan* ‘to last’ + gar → man.de.gar ‘lasting’
  - saxt ‘past stem of *saxtan* ‘to build’ + man → sax.te.man ‘building’
  - xast ‘past stem of *xastan* ‘to want’ + gar → xast.gar ~ xas.te.gar ‘suitor’
- (7) CV<sub>tense</sub>CC (no epenthesis)
- ist ‘stop’ + gah → ist.gah ‘station’ \*is.te.gah
  - xast ‘past stem of *xastan* ‘to rise’ + gah → xast.gah ‘origin’ \*xas.te.gah
  - rast ‘right, truth’ + gu → rast.gu ‘truthful’ \*ras.te.gu

The observation that epenthesis may occur in suffixation with particular stem structures lead some to consider the environment for epenthesis to be conditioned by properties of the vowel (e.g., Samareh 1977, Lazard 1992). Samareh relates the occurrence of the epenthesis to the length of the vowel of the root in cases such as *amuzgar* ~ *amuzegar* ‘teacher’. Presenting examples such as *pasban* ~ *paseban* ‘policeman’, Lazard considers the epenthesis to occur with stable vowels, *a, i, u* -he calls *a, i, u* stable because they have a relatively constant duration and are not subject to change in quality.

## 2. A possible account

While details of prosodic structure in Persian remain to be developed, the following is the outline of an account.

Epenthesis is often argued to result from syllabification demands (e.g., Selkirk 1981, Itô 1989). Assuming a moraic representation of vowels (e.g., Hyman 1985, Hayes 1989), the vowels which I call lax, *a, e, o*, occupy a single mora, while those which I call tense, *a, i, u*, occupy two morae. Assuming that a syllable can accommodate two morae, the postvocalic consonant following the monomoraic vowel can receive a mora (e.g., *gam* → *gamqin* ‘sad’) and is thus prosodically licensed in this way. But following the long vowel the consonant is not moraic (e.g., *kar* → *ka.re.gar* ~ *kar.gar* ‘worker’). While it can be licensed by associating to the prosodic word when it is final, when a suffix is present, the consonant syllabifies as an onset with epenthesis.

One can also consider the final consonant after long vowels to receive a mora, which means that a syllable can accommodate more than two morae. The latter analysis of the syllable structure does not change the possibility of the occurrence of epenthesis as it creates a superheavy syllable, with which the language allows epenthesis in some cases as shown in unsuffixed words such as *aḡna* ~ *aḡena* ‘familiar’, *aḡkar* ~ *aḡekar* ‘apparent’. When a monomoraic vowel is followed by a consonant cluster in the root, a vowel may be epenthesized when a consonant-initial suffix is present. The final coda consonant of the root acts as the onset of the syllable with an epenthetic vowel. When a suffix is added to a root which contains a long vowel followed by a cluster of two consonants, to avoid having four morae in the root after suffixation, the epenthesis occurs. I have to emphasize that this is not an account of the moraic structure of Persian and it is just to show that the heaviness of some structures (heavy due to the bimoraic vowel or a monomoraic vowel followed by a consonant cluster) can lead one to consider epenthesis to be conditioned by vowel properties.

It should be noted that the process, even given the above explanation, can still be accounted for based on quality, as follows. In order to account for the epenthesis based on the tense/lax distinction (that is a qualitative distinction), one can follow the direction of an analysis according to which features play a role in projecting syllable structure (see van Oostendorp 1995 for discussion). Quantity is not underlying, but vowels with the feature [tense] project two morae, unlike lax vowels, which project a single mora. A syllable is bimoraic only if the vowel has the feature [tense]. That is, [tense] is underlying and tense vowels are redundantly bimoraic. Thus the account given above based on mora is possible without the implication that quantity is underlying.

Regardless of the analysis one may provide for this process, given that the number of suffixed forms which do not show epenthesis is far more than those which show it, the main question is: if the occurrence of epenthesis is motivated by the properties of vowels and the difference in the syllable structure due to these properties, why does epenthesis not occur with all cases which include roots with those vowels/syllable structures? This is the question which will be addressed below.

### 3. On stems and suffixes

In order to determine whether the conditions under which epenthesis occurs are systematic I examine a variety of factors (i.e., clusters types, productivity of suffixes, and frequency), in 3.1-3.3, and found that they do not explain the variations in epenthesis. But before discussing these factors, I need to make a few points about stems and suffixes, as follows.

One cannot consider some particular roots to always force epenthesis.

- (8)  $ba_G$  ‘garden’  
 $ba_G + ban \rightarrow ba.Ge.ban \sim ba_G.ban$  ‘gardener’  
 $ba_G + tʃe \rightarrow ba_G.tʃe$  \* $ba.Ge.tʃe$  ‘small garden’

One cannot consider some particular suffixes to always show epenthesis.

- (9) -ban  
 pas + ban → pas.ban ~ pa.se.ban ‘police officer’ (pas ‘guard duty,  
 watch’)  
 baG + ban → baG.ban ~ ba.ge.ban ‘gardener’ (baG ‘garden’)  
 miz + ban → miz.ban \*mi.ze.ban ‘host/hostess’ (miz ‘table’)

Having said that, there are some suffixes which show more tendency towards epenthesis. These are -man, and -gar. There are some suffixes with less or no tendency towards epenthesis, such as: -tʃe, -dan, -dis, -zar, -sar, -san, -kade, -gun, -gin, -nak, -vaj. The suffixes -gar and -ban are in between the first two groups.

### 3.1 Cluster types

There is no particular tendency in terms of the clusters created in suffixation in cases where epenthesis is found, as shown in (10).

- (10) With m-initial suffix, words with final: z, t, r, j, d, ʒ  
 With b-initial suffix, words with final: s, G, r  
 With g-initial suffix, words with final: d, z, t, r, j

### 3.2 Productivity of suffixes

There is no particular correlation between the productivity of the suffixes and their tendency to show the vowel in suffixation process. The suffixes -gar, -ban, -gar, with which epenthesis is observed, are productive (Kalbasi 1992). There are, however, several other productive suffixes, such as -kade, -dan, etc. (again according to Kalbasi 1992), which do not show epenthesis.

### 3.3 Frequency

There are sound changes which affect words of low frequency and those of high frequency differently (e.g., Phillips 1984, Antilla 2006, Kang 2003, 2005, 2007). The question is: are the suffixed forms which show epenthesis more frequent compared to their root? In this case the dominant pattern, the non-epenthesis pattern, would not affect them as they have their own pattern due to their being more frequent than their root. In order to examine this possibility, I used Persian googling and Penglish googling (Penglish: Persian written in English script).

Persian googling turns out to be unreliable in this regard for various reasons, such as: in the Perso-Arabic script, *a*, *e*, *o* are not written. Therefore, the epenthetic vowel is not shown in written forms, or, for example, the result of *mehr* ‘kindness’ to be compared to *mehrban* ~ *mehraban* ‘kind’ includes the Persian words *mohr* ‘stamp’ and *mahr* ‘a sum of money that the bridegroom undertakes to pay the bride’.

Penglish googling is not completely reliable either. The advantage of Penglish is that the epenthetic vowel is inserted in written forms so one can tell the difference between, for example, *shademan* /ʃademan/ ~ *shadman* /ʃadman/, which is not possible to do through Persian google search. There are, however,

problems with Penglish, such as: a word may be written in more than one way (e.g., *u* could be written as *oo* or *ou*; *a* could be written as *aa* or *a*).

A general problem with both types of googling is that it is not clear how many of the words in the google search are based on Standard Persian spoken in Iran, which is the focus in my study, and how many are based on other related languages and dialects. Even loan words could change the result of googling. For example, when one googles a word like *pas* to be compared to *pasban* ~ *paseban* ‘police officer’ to see which one is more frequent, in addition to the Persian word, the loan word *pas* ‘pass’ which has entered into the language to mean ‘pass a ball (in soccer, etc)’ and ‘pass a course’ also appears.

Nonetheless, I googled several words and their suffixed forms and no pattern was observed in terms of frequency between the stems and their suffixed forms. An example of the google searching is presented in (11):

|      |   |                               |  |
|------|---|-------------------------------|--|
| (11) | baG   | ‘garden’                      | 2,470,000                                      |
|      | <i>Google search (in Persian script)</i>            |                               |  |
|      | a.  | baGban/ bageban ‘gardener’    | 125,000 (epenthesis possible)                  |
|      | b.  | baGdar ‘sb who owns a garden’ | 24,500 (no epenthesis)                         |
|      | c.  | baGʃe ‘small garden’          | 818,000 (no epenthesis)                        |
|      | <i>Penglish (Persian written in English script)</i> |                               |  |
|      | a.  | baGban/ bageban ‘gardener’    | baghban/bagheban<br>144,000/3,810              |
|      | b.  | baGdar ‘sb who owns a garden’ | baghdar<br>9,050                               |
|      | c.  | baGʃe ‘small garden’          | baghche/baghcheh <sup>1</sup><br>48,200/29,200 |

Note that the Persian corpora (e.g., Hamshahri corpus (2008), Bijankhan corpus (2007)) cannot be considered as a source for my search because they are based on a limited data and also on written Persian, which is not the focus in my work, which is more on speech, and in addition, the written Persian does not show the epenthetic *e*.

To sum up, types of consonant clusters, productivity of suffixes, and frequency do not account for the epenthesis process.

#### 4. The epenthesis in suffixation: historically

Given that there are present consonant-initial suffixes which were historically vowel-initial, it is worth to investigate whether the suffixes with which epenthesis may occur today had –e in their initial position in their historical forms, and then lost their initial –e and became consonant-initial. Under this account, the –e which is observed today is a consequence of the historical form of these suffixes. Consulting Farahvashi’s Middle Persian to Modern Persian

<sup>1</sup> In Persian, the words which end in –e are written with a silent –h at the end. In Penglish, both with-silent *h* and without-*h* versions are observed so I searched both versions for the Penglish of the word *baGʃe*.

dictionary (1967), Farahvashi's Modern Persian to Middle Persian dictionary (1974), and Kalbasi (1992), I show that this is not the case. There is no correlation between the occurrence of the vowel before the suffix in the present time and its occurrence in the past, as the following examples show.

The present suffix -nak was formerly -ēnak but it does not show the initial vowel in suffixation today.

|      |                       |           |  |
|------|-----------------------|-----------|--|
| (12) | <i>Middle Persian</i> |           | <i>Modern Persian</i>                    |
|      | a.                    | tars-ēnak | tars-nak 'scary' (tars 'fear') *tarsenak |
|      | b.                    | bīm-ēnak  | bim-nak 'fearful' (bim 'fear') *bimenak  |

The present suffix -ban had the form -bān/-pān in Middle Persian but it may show an epenthetic vowel in suffixation today.

|      |                       |         |  |
|------|-----------------------|---------|--|
| (13) | <i>Middle Persian</i> |         | <i>Modern Persian</i>                              |
|      | a.                    | bāy-pān | baG-ban ~ baG-e-ban 'gardener' (baG 'garden')      |
|      | b.                    | pās-pān | pas-ban ~ pas-e-ban 'policeman' (pas 'guard duty') |

There are cases in which the vowel occurs with -ban neither in its earlier version nor in its current one.

|      |                       |          |   |
|------|-----------------------|----------|---|
| (14) | <i>Middle Persian</i> |          | <i>Modern Persian</i>                   |
|      |                       | marz-pān | marz-ban 'border guard' (marz 'border') |

Looking at some specific words, there are cases whose *e* at the suffix boundary seems to be a historical residue. Consider Modern Persian *arḡ* 'value' + -mand.

|      |                       |          |  |
|------|-----------------------|----------|--|
| (15) | <i>Middle Persian</i> |          | <i>Modern Persian</i>                      |
|      |                       | arḡōmand | arḡomand (~ arḡmand) 'valued' (arḡ-o-mand) |

In Middle Persian, this suffix was in fact -ōmand, which is -mand today.

|      |                       |             |                                       |
|------|-----------------------|-------------|---------------------------------------|
| (16) | <i>Middle Persian</i> |             | <i>Modern Persian</i>                 |
|      | a.                    | xrat-ōmand  | xerad-mand 'wise' (xerad 'wisdom')    |
|      | b.                    | dard-ōmand  | dard-mand 'suffered' (dard 'pain')    |
|      | c.                    | hunar-ōmand | honar-mand 'artist' (honar 'art')     |
|      | d.                    | sūt-ōmand   | sud-mand 'beneficial' (sud 'benefit') |
|      | e.                    | kār-ōmand   | kar-mand 'employee' (kar 'work')      |

The vowel at the suffix boundary in *arḡomand* seems to be a historical residue.

However, there are cases with the vowel at the suffix boundary in the present time which cannot be historically explained, as follows:

(i) There are words which did not have the vowel in Middle Persian and now they may show it.

- (17) *Middle Persian*                      *Modern Persian*
- a.    rōḡ-kār                      ruz-gar ~ ruz-e-gar                      ‘times’ (ruz ‘day’)
- b.    āmōz-kār                      amuz-gar ~ amuz-e-gar                      ‘teacher’  
       (amuz ‘present stem of *amuxtān* ‘to learn sth, to teach sth to sb’)
- c.    ajāt-kār                      jad-gar ~ jad-e-gar                      ‘memento’ (jad ‘memory’)
- d.    bāy-pān                      baG-ban ~ baG-e-ban                      ‘gardener’ (baG ‘garden’)

(ii) There are words which may show the vowel at the suffix boundary today but are not found in Middle Persian dictionaries.

- (18) a.    šad-man ~ šad-e-man                      ‘happy’                      (šad ‘happy’)
- b.    bad-ban ~ bad-e-ban                      ‘sail’                      (bad ‘wind’)

To sum up, the historical investigation cannot fully explain the epenthesis.

## 5. Experiment on epenthesis in suffixation

In order to determine whether epenthesis is productive in the language, I conducted an experiment with 10 native speakers of Persian.<sup>2</sup> The experiment included both production and perception:

- (i) Task 1: *Production (Reading)*                      real words
- (ii) Task 2: *Production (Question and answer)*                      real and made-up words
- (iii) Task 3: *Production (Wug test)*                      made-up words
- (iv) Task 4: *Perception (Acceptability rating)*                      real and made-up words

### 5.1 Task 1: Production (Reading)

There were two texts in Persian. Both texts had several suffixed words, including those words which can take epenthesis and those which cannot. One of the texts was more formal and the other more informal. Some words were repeated more than once. The advantages of having a reading task were as follows: (i) It showed the basic pattern for each speaker; (ii) It showed intra-speaker variation; (iii) It showed in general how much tendency towards epenthesis exists in the language, that is, in cases where there is optional epenthesis how much the epenthesis occurs. The results of this task are as follows:

- (19) *The result of epenthesis-possible cases* (“E” stands for epenthesis; without-epenthesis versions are shown in “No E” row and with-epenthesis versions in “With E” row; Misc stands for miscellaneous).

<sup>2</sup> Many thanks to my consultants (in the order of participating in the experiment): Ali Esmaili, Farideh Tajik, Shadi Farshadfar, Vahid Danaee, Shery Shahabi, an anonymous consultant, Rana Mohammad Esmail, Poopak Haghi, Nikoo ParvinNejad, Lili ParvinNejad. Their help and willingness to participate in my experiment are much appreciated.

|        | # of Tokens | Total |      | %     |
|--------|-------------|-------|------|-------|
| No E   | 390         | 187   |      | 47.95 |
| With E | 390         | 203   |      | 52.05 |
|        |             |       | Misc | 0     |

There is not much difference between the epenthesis-including version and the non-epenthesis version, as (19) shows.

(20) *The result of cases with which epenthesis is not expected*

|        | # of Tokens | Total |      | %     |
|--------|-------------|-------|------|-------|
| No E   | 410         | 409   |      | 99.76 |
| With E | 410         | 1     |      | 0.24  |
|        |             |       | Misc | 0     |

With one exception, the words were pronounced without epenthesis by all speakers.

To sum up: (i) There is a set of words which can get epenthesis and there are words which cannot get epenthesis although they share the same root structure; (ii) Among those words which can get epenthesis, there is no particular tendency towards using or not using epenthesis, considering the overall result 47.95% vs. 52.05%.

## 5.2 Task 2: Production (Question and answer)

This production task includes both real and made-up (randomized) words. Participants were seated in front of a computer. They heard a question (“What did Ali say?” (in Persian)) and saw a slide on the screen which consists of ‘a root + a suffix’ (e.g., “*kar* ‘work’ + *-gar*” (written in Persian script)). They were asked to put together the root and the suffix, and put the word they made in a blank space in a frame sentence (“I think he said....” (participant’s response)).

The words which were examined are categorized into three groups:

- (i) Real words with which epenthesis is possible
- (ii) Real words with which it is not expected to see epenthesis
- (iii) Made-up words

The results of this task are as follows. I begin with real words.



(21) *The result of real words with which epenthesis is possible*

|        | # of Tokens | Total |      | %     |
|--------|-------------|-------|------|-------|
| No E   | 280         | 159   |      | 56.79 |
| With E | 280         | 121   |      | 43.21 |
|        |             |       | Misc | 0     |

No significant difference is observed between the epenthesis-including versions and those without.

(22) *The result of real words with which epenthesis is not expected*

|        | # of Tokens | Total |      | %   |
|--------|-------------|-------|------|-----|
| No E   | 120         | 120   |      | 100 |
| With E | 120         | 0     |      | 0   |
|        |             |       | Misc | 0   |

No epenthesis is found with the words with which it is not expected to get epenthesis.

Now let us look at made-up words. There are 24 made-up words. For each of the four root structures,  $CV_{\text{lax}}C$ ,  $CV_{\text{tense}}C$ ,  $CV_{\text{lax}}CC$ ,  $CV_{\text{tense}}CC$ , there are 6 words, 3 words with suffixes with which there are more cases of epenthesis, and 3 words with suffixes with which there are no or rare cases of epenthesis in real language. In terms of vowels, the lax vowels in  $CV_{\text{lax}}C$  include one example of each lax vowel. The same holds for the lax vowels in  $CV_{\text{lax}}CC$ . The tense vowels in  $CV_{\text{tense}}C$  also include one example of each tense vowel. The same holds for the tense vowels in  $CV_{\text{tense}}CC$ . I did not separate made-up words into two groups as epenthesis-expected and non-epenthesis-expected since the non-epenthesis is by far the dominant pattern regardless of the structures of the roots.

(23) *The result of made-up words*

|        | # of Tokens | Total |      | %     |
|--------|-------------|-------|------|-------|
| No E   | 240         | 235   |      | 97.92 |
| With E | 240         | 5     |      | 2.08  |
|        |             |       | Misc | 0     |

Out of 240 tokens of made-up words, only 5 cases were pronounced with epenthesis (note: those 5 cases had one of the structures with which epenthesis may occur in the language). This strongly confirms that non-epenthesis is the dominant pattern.

To sum up: (i) *Real words*: Epenthesis is limited to a set of words; (ii) *Made-up words*: No epenthesis is the dominant pattern.

### 5.3 Task 3: Production (Wug test)

This task includes only made-up words. The participants were given a list of Persian sentences and a list of suffixes. They were asked to make a suffixed form out of a word in a sentence using one of the suffixes and put it in the following sentence. For example, consider *mes* ‘copper’, from which, by adding the suffix -gar, there is *mesgar* ‘coppersmith’, which can be put in a sentence such as ‘Ali was working/dealing with *mes*. Ali was *mesgar*’. For a made-up word such as *haf*, it is possible to make *hafgar* to be put in the sentence. Note that there were different sentences for different suffixes suitable to their meaning/function (e.g., agentive, locative, etc.).

I chose six suffixes: two suffixes which show more cases of epenthesis (i.e., -gar, -man), two suffixes which show no epenthesis or rarely show it (i.e., -nak, -qah), and two suffixes which are in between (i.e., -gar, -ban).

The list of words of the made-up roots included the structures under study. For each suffix, 12 words were created, 3 words for each root structure. The 3 words for lax vowels with CV<sub>lax</sub>C structure include one example with each lax vowel. The same is the case for lax vowels with CV<sub>lax</sub>CC structure. The 3 words for tense vowels with CV<sub>tense</sub>C structure include one example with each tense vowel. The same is the case for tense vowels with CV<sub>tense</sub>CC structure.

The result of this task is as follows.

#### (24) The result of made-up words (wug test)

|        | # of Tokens | Total |      | %   |
|--------|-------------|-------|------|-----|
| No E   | 720         | 720   |      | 100 |
| With E | 720         | 0     |      | 0   |
|        |             |       | Misc | 0   |

The result is categorical and clearly shows the non-epenthesis pattern.

### 5.4 Task 4: Perception (Acceptability rating)

A list of real and made-up suffixed words was recorded. For each word, there were two versions, one with epenthesis and the other without. The words were randomized, so the two versions were not in a row. The participants were asked to rate each word on the following scale: √ (good, acceptable, possible); ? (so-so); X (bad, unacceptable, impossible). The list of the roots was written on paper

and the speakers were asked to put one of the three signs (√, ?, or X) in front of each word after they heard the suffixed forms.

The results are as follows. I begin with the real words.

(25) *The result of real words where epenthesis is possible*

|              | Total | %     |
|--------------|-------|-------|
| No E (√)     | 200   | 89.29 |
| With E (√)   | 205   | 91.52 |
| No E (?)     | 6     | 2.68  |
| With E (?)   | 2     | 0.89  |
| No E (X)     | 18    | 8.04  |
| With E (X)   | 17    | 7.59  |
|              |       |       |
| Total No E   | 224   |       |
| Total With E | 224   |       |

Both versions are acceptable without much difference in terms of percentage (89.29% and 91.52%).

(26) *The result of real words where epenthesis does not occur although the word has one of the structures with which one may get epenthesis in the language*

|              | Total | %    |
|--------------|-------|------|
| No E (√)     | 24    | 100  |
| With E (√)   | 2     | 8.33 |
| No E (?)     | 0     | 0    |
| With E (?)   | 1     | 4.17 |
| No E (X)     | 0     | 0    |
| With E (X)   | 21    | 87.5 |
|              |       |      |
| Total No E   | 24    |      |
| Total With E | 24    |      |

With real words with which one does not get epenthesis although based on their structure should be able to, the acceptability of non-epenthesis version is far more than the acceptability of with-epenthesis version.

(27) *The result of real words with which epenthesis is not expected (the CV<sub>lax</sub>C structure)*

|              | Total | %     |
|--------------|-------|-------|
| No E (√)     | 32    | 100   |
| With E (√)   | 0     | 0     |
| No E (?)     | 0     | 0     |
| With E (?)   | 1     | 3.13  |
| No E (X)     | 0     | 0     |
| With E (X)   | 31    | 96.88 |
|              |       |       |
| Total No E   | 32    |       |
| Total With E | 32    |       |

Real CV<sub>lax</sub>C roots strongly show the non-epenthesis pattern.  
Now let us look at made-up words.

(28) *The result of made-up words where epenthesis can be possible (the structures other than CV<sub>lax</sub>C)*

|              | Total | %     |
|--------------|-------|-------|
| No E (√)     | 134   | 93.06 |
| With E (√)   | 41    | 28.47 |
| No E (?)     | 4     | 2.78  |
| With E (?)   | 7     | 4.86  |
| No E (X)     | 6     | 4.17  |
| With E (X)   | 96    | 66.67 |
|              |       |       |
| Total No E   | 144   |       |
| Total With E | 144   |       |

No epenthesis is far more acceptable.

(29) *The result of made-up words with which epenthesis is not expected (the CV<sub>lax</sub>C structure)*

|              | Total | %     |
|--------------|-------|-------|
| No E (√)     | 45    | 93.75 |
| With E (√)   | 2     | 4.17  |
| No E (?)     | 2     | 4.17  |
| With E (?)   | 4     | 8.33  |
| No E (X)     | 1     | 2.08  |
| With E (X)   | 42    | 87.5  |
|              |       |       |
| Total No E   | 48    |       |
| Total With E | 48    |       |

The without-epenthesis forms of these words received a very high percentage of acceptability and the with-epenthesis forms a very high percentage of unacceptability.

### 5.5 Summary of the experiment

I ran an experiment including made-up and real words testing Persian native speakers for both production and perception through different tasks in order to study the synchronic status of epenthesis in Persian and to find out the underlying generalizations in the speakers' minds regarding this suffixation process and epenthesis. The result is as follows:

*Production:* (i) For *made-up* words, the non-epenthesis pattern is the general pattern. (ii) For *real* words, for those words with which one can take epenthesis both versions are fine. For real words which do not take epenthesis, the version without epenthesis is by far more acceptable, and the version with epenthesis is highly unacceptable.

*Perception:* (i) For *made-up* words, the non-epenthesis is the acceptable pattern in general both for words with CV<sub>lax</sub>C structures and for words with other structures. If the speakers accept epenthesis-including versions, it is more with words with structures other than CV<sub>lax</sub>C. (ii) For *real* words, for the words with which one can take epenthesis both versions are fine. For those real words with which it is not expected to see epenthesis, the non-epenthesis versions are far more acceptable, and the epenthesis-including versions are highly unacceptable.

## 6. Conclusion

Given the limited number of words which have both epenthesis-including and non-epenthesis versions and the results of the experiment, I argue that epenthesis is not a synchronically productive process in Persian and is limited to some frozen cases, and therefore cannot be considered as evidence for vowel structure (see Rohany Rahbar (in progress) for a detailed discussion).

## References

- Antilla, A. 2006. Variation and opacity. *Natural Language and Linguistic Theory*. 24(4), 893-944.
- Farahvashi, Bahram. 1967. Farhang-e Pahlavi. Entesharat-e Bonyad-e Farhang-e Iran.
- Farahvashi, Bahram. 1974. Farhang-e Farsi be Pahlavi. Tehran: Anjoman-e Asar-e Melli.
- Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253-306.
- Hyman Larry M. 1985. *A theory of phonological weight*. Dordrecht: Doris.
- Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217-260.
- Kalbasi, Iran. 1992. *Saxt-e Eshteghaghi-ye Vazhe in Modern Persian*. Tehran: Moassese-ye Motaleat and Tahghighat-e Farhangi.
- Kang, Yoonjung. 2003. Sound changes affecting noun-final coronal obstruents in Korean. In McClure, W. ed. *Japanese/Korean Linguistics* 12.
- Kang, Yoonjung. 2005. The emergence of the unmarked in an analogical change. Handout from Seoul Forum in Phonology, Seoul National University.

- Kang, Yoonjung. 2007. Frequency effect and regularization in Korean nouns. Handout from the Workshop on Variation, Gradience and Frequency in Phonology. Stanford, July 6-8.
- Lazard, Gilbert. 1992. *A Grammar of contemporary Persian*. Mazda, Costa Mesa, California. Translated from French by Shirley Lyons.
- Najafi, Abol Hassan. 2001. *Mabani-y-e Zabanshenasi*. Tehran: Niloufar Press.
- Phillips, Betty. 1984. Word frequency and the actuation of sound change. *Language* 60, 320-42.
- Rohany Rahbar, Elham. In progress. Aspects of Persian phonology and morphophonology. PhD thesis. University of Toronto.
- Samareh, Yadollah. 1977. *The arrangement of segmental phonemes in Farsi*. University of Tehran Press.
- Selkirk, E. 1981. Epenthesis and Degenerate Syllables in Cairene Arabic, *MIT Working Papers in Linguistics* 3: 209-32.
- Van Oostendorp, Marc. 1995. Vowel quality and syllable projection. Tiburg dissertation in Language Studies.
- Windfuhr, Gernot L. 1979. *Persian Grammar: History and State of its Study*. Trends in Linguistics 12. Mouton Publishers.