

Analyzing Tri-syllabic Geminated Words in Persian According to Optimality Theory (OT)

Aliye Kord Za'feranlu Kambuziya, Farnoosh Taherloo, Ferdows Aghagolzade, Arsalan Golfam

Linguistics Department, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran

Email address:

Akord@modares.ac.ir (A. K. Z. Kambuziya), ftaherloo@yahoo.com (F. Taherloo), Ferdows@yaho.com (F. Aghagolzade),

Golfam@modares.ac.ir (A. Golfam)

To cite this article:

Aliye Kord Za'feranlu Kambuziya, Farnoosh Taherloo, Ferdows Aghagolzade, Arsalan Golfam. Analyzing Tri-syllabic Geminated Words in Persian According to Optimality Theory (OT). *International Journal of Language and Linguistics*. Vol. 4, No. 6, 2016, pp. 220-229.

doi: 10.11648/j.ijll.20160406.14

Received: November 16, 2016; **Accepted:** December 2, 2016; **Published:** December 20, 2016

Abstract: This study is aimed to analyze gemination in tri-syllabic words in Persian. Data for the study is collected from Dehkhoda Medium Dictionary. The collected data categorized in different tables in Excel. To classify the data, various factors such as structure of syllable, number of morphemes, combination of morphemes (being simple word, derived, compound or derivational-compound), grammatical categories, and the language of origin are considered. The data is analyzed in three stages. First, the whole data including Persian and borrowed geminated words is studied. Then, Persian geminated words are selected and studied separately. Finally, the data is analyzed to answer the main question of current research: What types of morphemes in Persian can satisfy Obligatory Contour Principle (OCP). This question leads the researchers to hypothesize that derivational morphemes can satisfy OCP. To examine this hypothesis, Optimality Theory (OT) employed to figure out the order of limitations in optimum syllabic combination.

Keywords: Phonology, Optimality Theory (OT), Obligatory Contour Principle (OCP), Syllabic Structure, Gemination

1. Introduction

Gemination and its usage in Persian is rarely studied, since many researchers may consider gemination as Arabic phonetical pattern that is borrowed and used in Persian. However, a few geminated words that trace back to Old Persian such as /parre/ are still in use in Modern Persian¹. A few researches have been carried out on gemination in Persian from various perspectives like gemination in written Persian², in syllabic structure³, and its changes in Persian⁴. This research, however, attempts to study the frequency of tri-syllabic geminated words in Persian, their syllabic structure, their origin, frequency of geminated consonants in tri-syllabic words, frequency of simple, derived and compound words, and probable constraints on occurrence of gemination in word or syllabic ending according to OT. The

main objective of the study is presenting statistical results of the tri-syllabic geminated words in Persian to find out what type of morphemes can satisfy OPC. It is hypothesized that derivational morphemes play the most important role.

2. Definition of Gemination

Gemination has been defined in several ways which will be mentioned in the review of literature. Here, only Crystal's (2008) is represented. According to him⁵:

"Gemination is a term used in phonetics and phonology for a sequence of identical adjacent segments of a sound in a single morpheme, e.g. Italian *notte* /notte/ ('night'). Because of the syllable division, a geminate sequence cannot be regarded as simply a 'long' consonant, and transcriptional differences usually indicate this, e.g. [-ff-] is geminate, [-f:] is long. The special behavior of geminates has been a particular focus in some approaches to non-linear phonology, as a part of the discussion of the way in which quantitative

1 VahidianKamyar, 1992: 26

2 Vahidian Kamyar, 1992

3 Bijankhan, 2005

4 Kord Za'feranlu Kambuziya & Taj Abadi, 2012

5 Crystal, 2008: 206

phenomena should be represented. Those long segments which cannot be separated by epenthetic vowels ('true' geminates, represented with multiple associations) are said to display geminate 'inseparability' or 'integrity'. Those which fail to undergo rules because only one part of the structure satisfies the structural description are said to display geminate 'inalterability'. True geminates are contrasted with 'fake' or 'apparent' geminates, where identical segments have been made adjacent through morphological concatenation."

3. Research Method

The current corpus-based research has been conducted in three stages. First, 1535 geminated words has been selected among 60000 entries of Dehkhoda medium Dictionary. Then, all the words has been divided into groups based on several linguistic factors including word structure, number of morphemes, type of words (based on the combination of morphemes), grammatical categories, and the language of origin. The groups are registered separately in Excel and statistical analysis is conducted on these groups of words. Since the majority of geminated words are borrowed from Arabic (75.58 percent of the whole data), it was required to run the statistical analysis on two separate groups of words. Therefore, 371 Persian geminated words were studied and analyzed separately. The statistical analysis of these words lead to extraction of 101 geminated Persian words. The place of gemination in these words confirmed the hypothesis of the present research: no geminated consonant exists in word final position in Persian.

4. Review of Literature

Gemination in Persian has been studied from different points of views: gemination in written Persian, types of gemination in Persian, geminating and degeminating, and variation and changes in geminated sounds. The most important studies are mentioned in this part.

In his descriptive study, Vahidian Kamyar (1992), defines gemination in this way: "when two same consonants become adjacent in a way that the first one has no vowel while the second one has a vowel, the former is omitted and the latter becomes geminated." His study mostly deals with writing and diacritics.

Mir'Emadi (1994) has concentrated on students' problems in reading geminated words and discusses whether or not we need to show gemination while we are writing. He analyzes this phenomenon phonetically, and attempts to find the context in which gemination may occur.

Samare⁶ (1985) who studied gemination phonetically, explains this phenomenon in Persian as an incomplete articulation. He states that when two same consonants or two distinguished consonants with the same place of articulation become adjacent, completion of one sound and preparation of

the adjacent sound come together which results in fusion and longer duration. So, to produce these two consonants, lips are closed just once and there is no preparation for the second consonant.

In another study, Kord Za'feranlu Kambuziya (2006) presents two groups of words. Comparing these two groups shows that their difference is the occurrence of two same consonants in one group. She names this phenomenon as gemination. She studies differences in true and fake gemination, stylistic variation in using gemination, and the occurrence of consonant repetition, using gemination for rhyming in poetry, and degemination.

Najafi⁷ (2011) defines gemination as a set of same phonemes. He believes that Persian vowels cannot be geminated; so, this phenomenon can only be found in consonants. As a result, it is better to call gemination as "repetitious consonant". According to Najafi, gemination belongs to borrowed Arabic words and it can be omitted in Persian words except when omission can cause ambiguity.

For Zabih Nia (2012), showing or not showing the gemination in Persian written form is the question of a study. He believes that Persian speakers are aware of the existence of gemination in the words they produce, but he doubts if showing written form of gemination can make reading easier for Persian speakers. He states that gemination in Persian is formed by a phonological rule and it is used by Persian speakers unconsciously.

The process of gemination and degemination is studied by Kord Za'feranlu Kambuziya and Taj Abadi (2012). They study the phonological process that leads to the production of geminated consonants in some languages like Persian. They also declare that gemination and degemination are related to other phonological processes such as amplification, reduction, assimilation and compensatory lengthening. The main achievements of the study are the fact that degemination opposes gemination, and compensatory lengthening and consonant insertion are the most frequent phonological process after degemination.

Finally, Zirk and Skaer (2013) attempt to give a new interpretation of phonetical opposition of geminated sounds, and the differences between the articulating process of lexical and post-lexical geminations. Their findings show that gemination within a word has the same duration as gemination between two words. However, articulation of gemination between two words lack [+tense] feature so, it has considerable impact on the vowels before geminated consonant.

5. Theoretical Framework

McCarthy's (2008) optimality theory is the basis of the current study's framework. In this approach, an input representation is related to a set of chosen output representation. Various types of filters operate on input to select a well-formed, optimum choice. The selection is

⁶ Samare, 1985: 45-46

⁷ Najafi, 2011: 79

governed by the linguistic word-formation rules and limitations that are ordered according to the rules priority in each language⁸.

In this research gemination is analyzed based on optimality framework. As Bijankhan⁹ suggests, phonological patterns of gemination in the boundary of two morphemes in Persian declare dissimilation. He considers three possible hypotheses for word final consonants: a) Final consonant cannot be geminated (there should be a phonological rule that makes a consonant geminated before vowels); b) final consonant can become geminated (because it makes articulation more natural and omission of a phonological element is more common than inserting one); c) final consonant can be considered as a changeable consonant (so that it can have a kind of realization when there is a vowel after that). Bijankhan explains the occurrence of gemination in word/syllable final position using optimality theory and indicates that there is a constructing limitation in Persian that does not allow word or syllable final gemination. Therefore, it is possible to consider words' final position as obligatory contour boundary for gemination. He introduces three limitations for geminated final consonant in Persian: DEP-IO, MAX-IO, and OCP. OCP forbids word final gemination, MAX-IO does not let omission of phonological element in input and DEP-IO constraints phonetical insertion in output. He defines six permutations among which the following is the optimum:

(A) OCP >> MAX-IO, DEP-IO

Here the order of DEP-IO, MAX-IO limitations is not significant. Tables 1 & 2 show the process of optimum selection. We can notice that either C_1VC_2 or $C_1VC_2C_2$ as input has the same output of C_1VC_2 .

Table 1. C_1VC_2 is the input.

C_1VC_2	OCP	MAX-IO	DEP-IO
⇒ C_1VC_2			
$C_1VC_2C_2$	*!		*

Table 2. $C_1VC_2C_2$ is the input.

$C_1VC_2C_2$	OCP	MAX-IO	DEP-IO
⇒ C_1VC_2		*	
$C_1VC_2C_2$	*!		

Bijankhan believes that in geminated words such as /haGGi/ (A right) the final consonant of the first syllable is geminated but the structure of the word leads to a separation in which the first member of the geminated consonant comes as a final consonant for the first syllable and the second geminated consonant becomes initial part of the second syllable. Therefore, this separation does not violate OCP limitation. This combination cannot choose the optimum selection, so requires new limitation. As noted here,

gemination and structure of syllable related to each other directly, so a new limitation is required. Bijankhan introduces alignment (ALIGN) limitation as follows:

ALIGN (Stem, R, Syll, R):

"There is a syllable for each root in a way that the right border of root coincide the right border of a syllable as if syllable and root has one border at right side."¹⁰

He explains that OCP constraint must govern ALIGN constraint because OCP prevent the occurrence of geminated consonant in syllable final position and can oppose ALIGN constraint which prevent the existence of two geminated consonants in syllable borders. Following tables show the function of these two constraints. Table 3 shows geminated words before affixes and table 4 shows geminated word combined with other words in which the second member of consonant cluster does not appear:

Table 3. Geminated words before affixes.

$/C_1VC_2C_2+e/i/$	OCP	ALIGN	MAX-IO
⇒ $C_1VC_2C_2e/i$		*	
$C_1VC_2C_2.e/i$	*!		
$C_1V.C_2e/i$		*	*

Table 4. Geminated word combined with other words.

$/C_1VC_2C_2+CV.../$	OCP	ALIGN	MAX-IO
⇒ $C_1VC_2.CV$		*	*
$C_1VC_2C_2.CV$	*!		

In following section of this research these constraints and their function on tri-syllabic geminated words are analyzed.

6. Data Analysis

A few important facts should necessarily be noted first before starting data analysis: a) in data collecting, some entries were found that were written with a geminated consonant at word final position. As mentioned before, Persian word formation rules do not allow this; so, these entries were excluded from the current research; B) True gemination does not take place in the border of two morphemes of a compound word. For instance, in a word like /hammahal/ (same place) which consists of two morphemes /m/ is not a geminated phoneme; c) the words that are geminated due to stylistic differences, such as /sadde – i / versus /sade – i/ (easy) are excluded from this study; d) Onomatopoeic and expressive words that are geminated to show emphasize or showing excitement such as /d3ezzevzezz/ (used for showing anger) are included in current research.

6.1. Analysis of All Tri-syllabic Geminated Words

In the first step of data analysis, a general overview of all geminated words is studied. This statistical analysis includes

⁸ Crystal, 2008: 343

⁹ Bijankhan, 2005: 198

¹⁰ Bijankhan, 2005: 203

all Persian and borrowed words. To find out the frequency of geminated words, all the depicted words were divided to several groups according to their syllabic structure. The result of the statistical analysis can be found in figure 1, shows that CV.CVC.CVC structure is the most frequent one with geminated consonant. In this research, 858 words with CV.CVC.CVC structure are found which stand far higher

than second group. It should be mentioned that the majority of these words are simple words borrowed from Arabic; Persian words of this group are mostly derived ones. Among all other structures, CVCC.CVC.CV and CV.CVC.CVCC have the least number of words; only four geminated words with this structure were found.

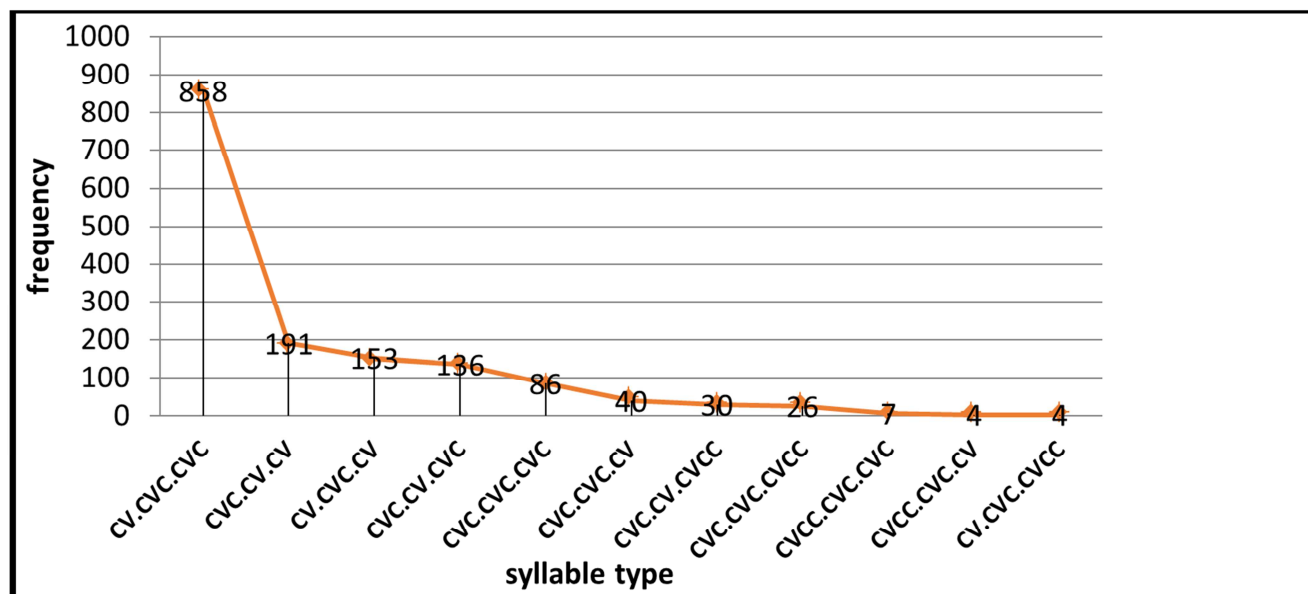


Figure 1. Syllabic Structure Type and Frequency of Tri-syllabic Geminated Words.

Moreover, table 5 gives an example of each structure to clarify the differences of studied groups. The whole elicited data was 1535 geminated words. However, 23 words had double gemination like /haGGonnas/ (people's right) /haGGollah/ (God's right) /ʃaGGoraGG/ (stiffly); Therefore, 1558 tokens were analyzed.

Table 5. Examples of Different Syllabic Types of whole data.

No	syllabic structure	frequency	example	meaning	Word type
1	CVC.CVC ₁ .C ₁ VC ₁₁	86	/porgovvat/	powerful	compound
2	CVC ₁ .C ₁ VC.CVCC	26	/sevvomʃaʒs/	3 rd person	derivational-compound
3	CV.CVC ₁ .C ₁ V	153	/bihessi/	numb	derived
4	CV.CVC ₁ .C ₁ VC	858	/tahammol/	tolerance	simple
5	CVC ₁ .C ₁ V.CV	191	/dʒarrahi/	surgery	derived
6	CVC ₁ .C ₁ V.CVC	136	/hoggebaz/	slicker	compound
7	CVC.CVC ₁ .C ₁ V	40	/jaczarra/	a little	compound
8	CVC.CV.CVCC	30	/haddohasr/	limitation	compound
9	CV.CVC ₁ .C ₁ VCC	4	/mofarranj/	foreign	simple
10	CVCC.CVC ₁ .C ₁ VC	7	/ʒordhemmat/	aimless	compound
11	CVCC.CVC ₁ .C ₁ V	4	/dastʔarre/	small saw	compound
-----	-----	1535	-----	-----	-----

Figure 2 shows the origin of the words. According to this figure, among 1535 selected words, 1160 are Arabic, 371 words are Persian, 2 words are Russian, and there exist 1 Latin and 1 French word. In terms of Percentage, 75.58 percent of words were Arabic, 24.17 were Persian, 0.13 were Russian, 0.06 was Latin and 0.06 was French.

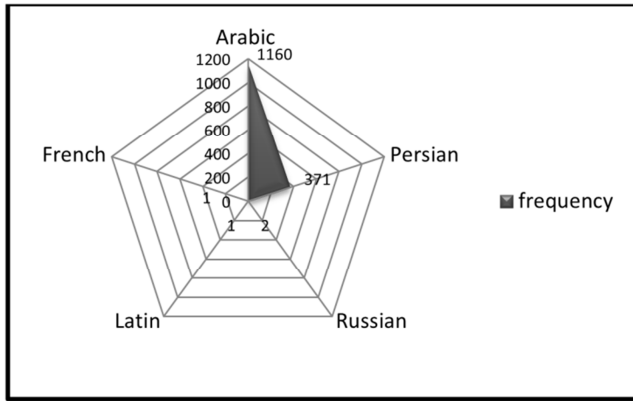


Figure 2. Frequency of Tri-syllabic Geminated words according to language.

Table 6 shows borrowed geminated words from Russian, French and Latin.

Table 6. Non-Arabic Geminated Borrowed Words.

Language of origin	Phonological form	meaning
Latin	/belladon/	Belladonna
French	/selloloz/	cellulose
Russian	/šammate/	alarm
Russian	/šammata/	alarm

The collected data is also analyzed according to the number of morphemes. Figure 3 shows that 1116 words of whole data, 72.70 percent, consist of one morpheme; 346 words, 22.54 percent, have two morphemes and only 73 words, 4.76 percent, are made up of three morphemes.

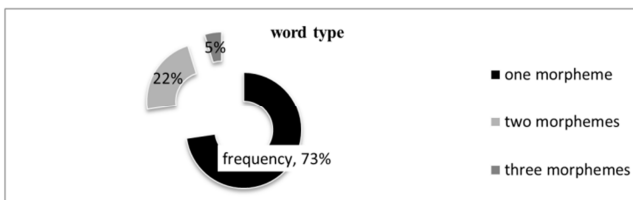


Figure 3. Tri-syllabic Geminated Words based on Number of Morphemes.

Figure 4 shows different types of words based on morphemes combination. As it is shown, the majority of words, 1143 words which means 72.7 percent, are simple ones; 246 words, 13.87 Percent, are compound and 180 words, 11.37 percent, are derived words. The rest, 26 words, are derivational- compound words.

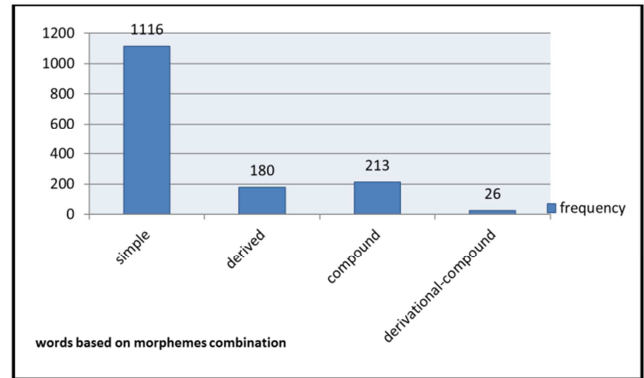


Figure 4. Frequency of Geminated Words based on Type of Words.

Figure 5 divided the words based on their grammatical category including: noun, adjective, adverb, verb, sentence, infinitive, short infinitive, derived nominal, denominal verbs, minor clause, relative pronoun, conjunctives, onomatopoeic words, reduplicated phrase and conjunctive phrase. Geminated adjectives are the most frequent in Persian: 777, 50.61 percent of tri-syllabic geminated words are adjectives. Then infinitives and nouns are the most frequent with correspondingly 359 and 201 words. The least frequent grammatical categories are short infinitive and verb with only one word for each of them.

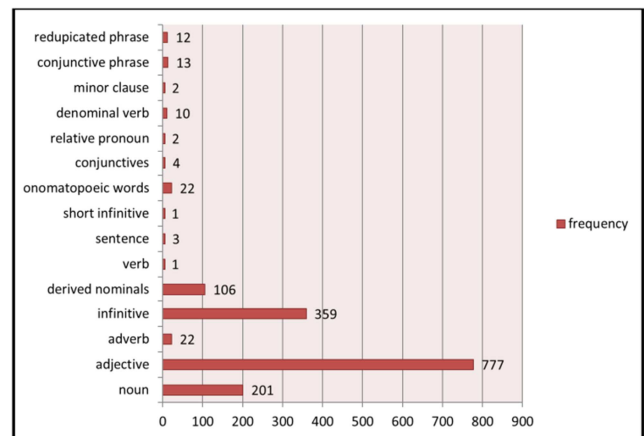


Figure 5. Frequency of tri-syllabic geminated words based on grammatical category.

Figure 6 distinguishes the words with the geminated phonemes in the border of first and second syllables from the ones with the geminated phoneme in the border of second and third syllables. In general, there were 1547 words with geminated phoneme in the border (12 was double geminated). A glance at the figure 6 indicates that the number of the words with geminated phonemes in the border of second and third syllable is more frequent with the total number of 1113, 71.95 percent. Moreover, the most frequent geminated phoneme in the words with geminated phonemes in the border of first and second syllable is /r/ with 81 words and /l/ with 57 words while /ʒ, tʃ, ʃ, p/ never occurred. On the

other hand, /l/ was the most frequent geminated phoneme in the words with geminated phonemes in the border of second

and third syllable, occurring 145 times while /r/ was the second one occurring 141 times. /ʃ, ʒ/ had no frequency.

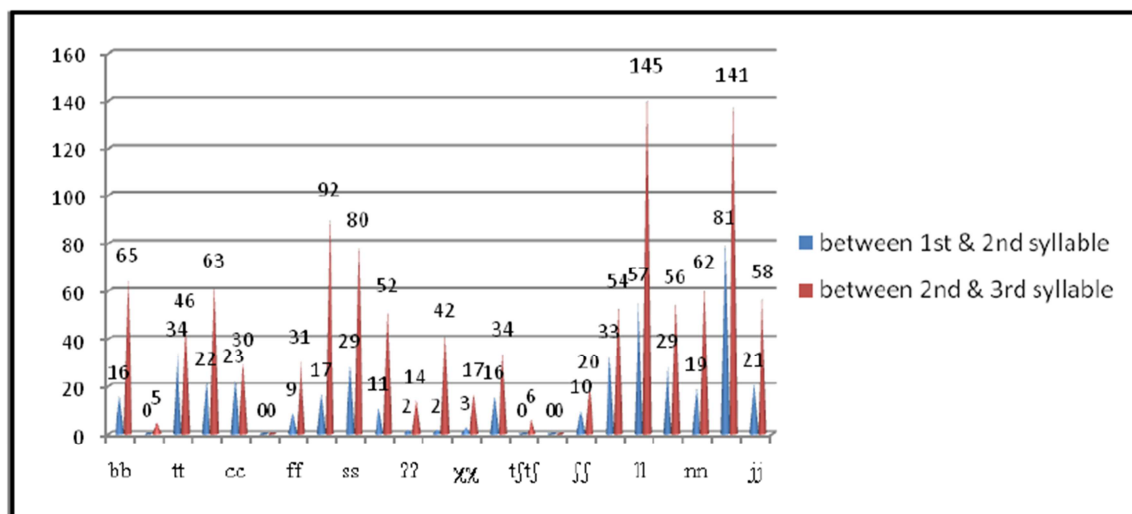


Figure 6. Frequency of geminated phonemes in the border of syllables.

Table 7 is allocated to the 11 words in which geminated phoneme placed in the third syllable. It is necessary to say that the words with geminated words in one syllable are mainly used for emphasis or demonstration of excitement¹¹. In table 7, C₁ shows syllable border and C₂ shows geminated phoneme in one syllable.

Table 7. Frequency of words with geminated phoneme in one syllable.

phoneme	frequency	Phonological form	Syllabic structure
dd	1	/ʃaddomadd/ (power)	CVC ₁ .C ₁ V.CVC ₂ C ₂
zz	3	/dʒezzovezz/ (onomatopoeic word for frying)	
		/ʔezzotʃezz/ (crying)	
		/ʔezzodʒezz/ (crying)	
XX	1	/ʃayXXoraXX/ (stiffly)	
ff	1	/Gallogaff/ (hatred)	
GG	3	/ʃagGoraGG/ (stiffly)	
		/laGGodaGG/ (savannah)	
mm	1	/temmoremm/ (soil & water)	
nn	1	/hennohenn/ (gasping)	

Table 8 shows the sequence of the vowels in syllabic structure of geminated words. Total number of possible sequence of long and short vowels is 216. However, 123 sequences are not found in geminated words that were analyzed in this research. It seems that short vowels usually come before geminated consonant while long vowels mostly come after¹². For instance, in /maddijat/ (materials), long vowel /i/ is recognizable. Actually, it is /e/ in /maddede/ (material) that is changed to /i/ because it assimilates to its' following consonant. According to table 8, the most frequent sequences are /oaa/, and /aao/ which take place 339 and 289 times correspondingly. The least frequent sequences belong to sequence of long vowels.

Table 8. Frequency of vowels' sequence in geminated words.

aaa	37	aaa	4	iaa	4	uaa	
aae	6	aae	2	iae	2	uae	
aaO	298	aaO	1	iao		uao	
aaα	34	aaα	5	iaα	1	uaα	
aaI	55	aaI		iai		uai	
aaU	4	aaU		iaU		uaU	
aea	11	aea	1	iea	6	uea	
aeE	3	aeE		ieE		ueE	
aaa	10	aaa	339	aaa	4	iaa	4
aae	2	aae	192	aae	2	iae	2
aaO		aaO		aaO	1	iao	
aaα	10	aaα	62	aaα	5	iaα	1
aaI	5	aaI	27	aaI		iai	
aaU	1	aaU	5	aaU		iaU	
aea	4	aea	6	aea	1	iea	6
aeE	2	aeE		aeE		ieE	

11 Kord Za'feranlu Kambuziya, 2006: 340

12 Kord Za'feranlu Kambuziya, 2006: 350

aeo		eeo		oeo	1	deo		ieo		ueo	
aea	8	eea	8	oea	8	dea		iea		uea	
aei	2	eei	3	oei		dei	1	iei	1	uei	
aeu	1	eeu		oeu		deu		ieu		ueu	
aoa	33	eoā	3	ooa	15	doa	3	ioa		uoa	1
aoe	2	eoē	13	ooē		doē		ioē	1	uoē	
aoō	5	eoō	1	ooō	4	doō	1	ioō		uoō	
aoā	21	eoā	4	ooā	1	doā	2	ioā	1	uoā	
aoi	6	eoī	4	ooī	2	doi		ioī		uoi	
aoū	3	eoū	1	ooū		dou	2	iou		uou	
ada	41	eda		oda	13	daa		ida		uaa	
ade	16	ede		ode	1	dae	1	ide		uae	
ado	1	edo	1	odo		dao		ido		uao	
adā	9	edā		odā	4	daā		idaā		uaā	
adi	95	edi	1	odi	8	dai		idi		uai	
adu		edu		odu	1	daū		idu		uaū	
aia	8	eia	7	oia	7	dia	2	ida		uia	
aie	1	eie		oie	2	die		ide		uie	
aio		eio		oio		dio		ido		uio	
aiā	2	eiā	1	oiā	1	dia	1	ida		uia	
aii		eii	3	oii		dii	1	idi		uii	
aiu		eiu	2	oiu		dau	1	idu		uiu	
aua		eua		oua	1	dua		ida		uua	
aue	1	eue		oue		dau		ide		uue	
auo		euo		ouo		duo		ido		uuo	
auā		euā		ouā		dua		ida		uua	
aii	1	eui		oui		dai		idi	1	uii	
auu		euu		ouu		duu		idu		uiu	

6.2. Analysis of Persian Tri-syllabic Geminated Words

The main goal of the research in this part is analyzing 371 Persian geminated words to find out the most frequent syllabic structure, grammatical categories, and word structure. Figure 7 illustrates the Persian geminated words based on their syllabic structure. The most frequent structure according to figure 7 is CVC.CV.CV with 133 samples, and the least frequent one is CVCC.CVC.CV with 4 samples. It should be noted that 125 words with CVC.CV.CV syllabic structure are derived words.

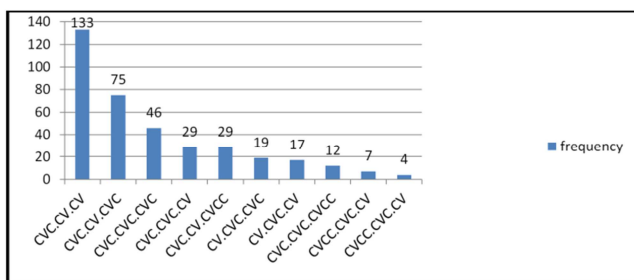


Figure 7. Frequency of syllabic structure of Persian geminated words.

Table 9 indicates a statistical analysis of combination of morphemes in Persian geminated words. As mentioned before, the majority of tri-syllabic geminated Persian words are derived. The next level is occupied by compound words with 155 words and the least frequent structure belongs to simple words. Indeed, 279 words consist of two morphemes,

70 consist of three morphemes and 22 are simple words.

Table 9. Frequency of Persian geminated words based on combination of morphemes.

Type of words	Frequency
Derived	168
Compound	155
Derivational Compound	26
Simple	22

Table 10 shows the grammatical category of Persian geminated words. According to this analysis, the most geminated Persian category is adjective, while short infinitive and denominal verb are the least frequent with only one occurrence.

Table 10. Frequency of tri-syllabic geminated Persian words based on grammatical category.

Category	Frequency	Category	Frequency
adjective	166	Infinitive	6
derived nominal	95	adverb	6
Noun	83	Short infinitive	1
onomatopoeic words	13	denominal verb	1

6.3. Analysis of Data based on Optimality Theory

In this part the collected data is analyzed based on optimality theory. As mentioned before, OCP constraints the occurrence of gemination in words' final position. Moreover, the limitations introduced as OCP, MAX-IO, DEP-IO cannot

lead to the optimum selection. Thus, another limitation defined as ALIGN is required. In this part, Persian tri-syllabic words are analyzed according to ALIGN constraint. Focusing on this constraint 101 Persian tri-syllabic

geminated words can be studied among 371 selected words. Figure 8 shows the statistical distribution of these words. Since ALIGN constraint should be studied on the border of two morphemes, the simple words are excluded.

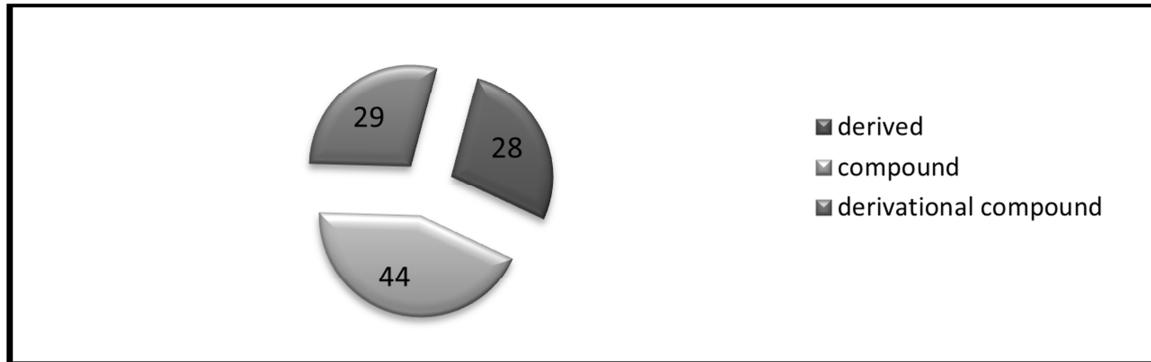


Figure 8. Frequency of words based on combination of morphemes (Final Data).

Table 11 shows the frequency of data showed in figure 8 according to grammatical category where adjectives and nouns are located in the highest rank.

Table 11. Frequency of words based on grammatical category.

category	Frequency	Category	Frequency
Adjective	43	Infinitive	5
Noun	42	Adverb	2
Derived nominal	9		

Various derivational suffixes are added to the simple words which result in gemination process. Table 12 shows these affixes with a Persian instance for each.

Table 12. Derivational suffixes in tri-syllabic geminated Persian words.

No	word	meaning	suffix	phonological form	final meaning	Final syllabic structure
1	/laɡ/	loose	/-andan/	/laɡɡandan/	loosen	CVC ₁ .C ₁ VC.CVC
2	/χas/	special	/-ana/	/χassana/	specially	CVC ₁ .C ₁ V.CV
3	/zar/	gold	/-ina/	/zarrina/	golden	CVC ₁ .C ₁ V.CV
4	/laɡ/	loose	/-idan/	/laɡɡidan/	to loose	CVC ₁ .C ₁ V.CVC
5	/taraɡ/	crack	/-e/ (noun maker)	/taraɡge/	fire cracker	CV.CVC ₁ .C ₁ V
6	/doʃaɡ/	two parts	/-e (adjective maker)	/doʃaɡga/	being in two parts	CV.CVC ₁ .C ₁ V
7	/ɡor/	roar	/-ida/	/ɡorrida/	roaring	CVC ₁ .C ₁ V.CV
8	/ɡor/	roar	/-anda/	/ɡorrandā/	roaring	CVC ₁ .C ₁ VC.CV
9	/χas/	particular	/-aʃan/	/χassajan/	particular ones	CVC ₁ .C ₁ V.CVC
10	/se/	three	/-omin/	/sevvomin/	the third	CVC ₁ .C ₁ V.CVC
11	/se/	three	/-om/	/sevvomi/	third one	CVC ₁ .C ₁ V.CV
12	/zar/	gold	/-in/	/zarrinbal/	golden wing	CVC ₁ .C ₁ VC.CVC
13	/dal/	full	/-a/	/dalladal/	replete	CVC ₁ .C ₁ V.CVC
14	/bor/	cut	/-a/ (noun maker)	/borraji/	sharpness	CVC ₁ .C ₁ V.CV
15	/ʔoho/	oh	/-va/	/ʔohovva/	oh (with more emphasis)	CV.CVC ₁ .C ₁ V
16	/becol/	total	/-i/ (adjective maker)	/becolli/	totally	CV.CVC ₁ .C ₁ V
17	/bihes/	numb	/-i/	/bihessi/	numbness	CV.CVC ₁ .C ₁ V
18	/ʃerep/	Sound of pouring water	/-i/	/ʃereppi/	Adverb of this sound	CV.CVC ₁ .C ₁ V
19	/mahal/	location	/-i/ (adjective maker)	/mahalli/	local	CV.CVC ₁ .C ₁ V

In the collected data, 44 compound words exist in which the second member of geminated consonant cluster appears due to combination of two morphemes. Most of these compound words are reduplicated or conjunctive phrases. It seems that occurrence of /va/ in Persian (“and” in English), which pronounces as /o/ leads to appearance of second consonant and gemination. For instance, in a word like /Gaʃiʃ₂ozaʔf/ (faint) the geminated consonant, /ʃ₂/, appears

because of /o/ in the second syllable while in its infinitive form as /Gaʃcardan/, /ʃ/ is pronounced just once since there is no vowel after that.

In this part, some derived tri-syllabic geminated Persian samples are analyzed through OCP, MAX-IO, and ALIGN constraints. Table 13 shows three possible situations of gemination’s occurrence. By analyzing these three constraints, we can figure out which structure is suitable for

derived words. In table 13, a Persian word is analyzed which is pronounced as /laG/ (loose) in normal condition, but by adding a causative suffix, /an.dan/ a vowel comes after /G/ and provides the opportunity of the second consonant's appearance. If we consider /laGG+an.dan/ (loosen) as deep structure, then three syllabic structures are possible: /laG.Gan.dan/, /laGG.ʔan.dan/ and /la.Gan.dan/. According to the introduced constraints, /la.Gan.dan/ cannot be the optimum choice since it violates all constraints except OCP. /laGG.ʔan.dan/ also violates OCP which is the most important constraint here. Therefore, the optimum choice is /laG.Gan.dan/ which violates only ALIGN limitation.

Table 13. Three possible situations of gemination's occurrence in derived word /laGG+an.dan/.

/laGG+an.dan/	OCP	ALIGN	MAX-IO
⇒ laG.Gan.dan			
laGG.an.dan	*!		
la.Gan.dan		*	*

Table 14 analyzes another instance of derived word in which the second member of geminated consonant cluster appears when the suffix is added to the root. The analyzed word consists of /be.col/ (total) with adverb making /-i/ suffix. Here, the only possible optimum choice is /be.col.li/ (totally), since it only violates ALIGN constraint; while /be.coll.i/ violates OCP and /be.co.li/ violates two constraints.

Table 14. Three possible situations of gemination's occurrence in derived word /be.coll+i/.

/be.coll+i/	OCP	ALIGN	MAX-IO
⇒ be.col.li		*	
be.coll.i	*!		
be.co.li		*	*

All derived words of current research's data with the suffixes in table 12 follow the same rule mentioned above.

Table 15 shows the order of constraints in compound words. The analyzed word is consists of two simple words, /χas/ (special) and /ʔam/ (general), that are connected to each other by /o/. Here again, the optimum construction is /χas.so.ʔam/, since it shows least violation.

Table 15. The order of constraints in compound word /χass+o+ʔam/.

/χass+o+ʔam/	OCP	ALIGN	MAX-IO
⇒ χas.so.ʔam		*	
χass.o.ʔam	*!		
χa.so.ʔam		*	*

7. Results

The main achievements of current study at the first stage are as follows:

- (1) The most frequent syllabic structure of tri-syllabic geminated words is CV.CVC.CVC with 858 words of

1535 analyzed words in this study.

- (2) The majority of tri-syllabic geminated words in Persian, almost 75 percent of analyzed data of this study, are borrowed from Arabic and only four words are borrowed from other languages.
- (3) It was presupposed that most of the geminated words should contain more than one morpheme; however, the analysis showed that 72.7 percent of geminated words were simple, which are borrowed mostly from Arabic. Other words are compound, derived and derivational-compound respectively.
- (4) The most frequent tri-syllabic geminated words were adjectives (50.61 percent of whole data). Infinitives stand at the second rank.
- (5) Regarding to the analysis of geminated phonemes in the border of syllables, the most frequent phonemes were found. In the border of first and second syllables, /r/ occurred more than others s and /p, ʒ, tʃ, ʃ/ could not be found. In the border of second and third syllables, the most frequent phonemes were /l/ and /r/ while /ʒ, ʃ/ could not be found.
- (6) In Persian, geminated phonemes do not come at the end of one syllable when a word with potential gemination used separately and the geminated phoneme distributed in two syllables when a vowel comes after, except when a speaker wants to show excitement or emphasize. In the analyzed data, 11 words were found in which both parts of geminated phonemes occurred in one syllable. The geminated phonemes in this situation were /m, n, ɣ, ʃ, χ, z, d/.
- (7) Possible vowel sequences in Persian geminated words are 216 combinations, but 123 sequences could not be found in this analysis. The most frequent vowel sequence was /oaa/ with 339 founded case. In this analyzed data, only 4 samples were found with the combination of three long vowels. It can be interpreted that mostly short vowels come before geminated consonants.

In the second part only Persian geminated words were studied. The main achievements of the second part are as follows:

- (1) The most frequent syllabic structure is CVC.CV.CV that was found 133 times. This finding is in accordance with the whole data. It can be inferred that the occurrence of geminated consonants is more probable in open syllables.
- (2) Among Persian geminated words, derived ones were the most frequent. Unlike the analysis of whole data in which the simple words stood in first rank, here the simple words were the least frequent words. It can be inferred that gemination is not a common phonological process in Persian and it usually takes place as a result of combination of morphemes. Moreover, the words formed by two morphemes are the most frequent ones with 279 instances. This can confirm that gemination is usually happens in the border of two morphemes in Persian.
- (3) Analyzing various grammatical categories showed that

adjectives are the most frequent tri-syllabic geminated Persian words with 166 instances. This result is in accordance with the analysis of the whole data.

The last part of study is related to the analysis of the collected data according to optimality theory. The results are as follows:

- (1) The most frequent grammatical categories in this analysis were adjectives and nouns.
- (2) In the analyzed data, compound words had the highest frequency with 44 instances. The simple words were excluded from this part of study, since appearance of second member of geminated cluster is due to combination of two morphemes. Among whole data, 28 instances were derived words that usually came with suffixes such as /-andan /, /-ande/, /-i/, /-ine/.
- (3) In this part of data analysis, the main purpose is to study three constraints: OCP, ALIGN, and MAX-IO to find out the optimum choice. Where the deep structure of a derived word was [C₁VC₂C₂+derivational morpheme], while all suffixes start with a vowel, the optimum choice was /C₁VC₂.C₂V.../. The same syllabic structure was acceptable when two words connected by /o/, violating only ALIGN constraint.

8. Conclusion

Majority of tri-syllabic geminated words in Persian (75%) are borrowed from Arabic. In Persian there is no geminated consonant at the end of syllable except when a speaker wants to show excitement or emphasize. In geminated words, most frequent vowel sequences are combinations of short vowels. Open syllables were preferred for geminated syllabic structures and CVC.CV.CV is the popular one. Moreover, derived words have a lot of occurrence which shows that most of Persian geminated words consist of an Arabic word plus a derivational suffix. Most of these derivational suffixes are adjective makers; therefore adjectives are most frequent among other grammatical categories.

References

- [1] Bijankhan, Mahmud (2005), *Phonology of Optimality Theory*, Tehran: Samt.
- [2] Crystal, D., (2003). *A Dictionary of Linguistics & Phonetics*, (3rd ed.), Oxford: Blackwell.
- [3] Dehkhoda, Ali Akbar (2003), *Medium Dehkhoda Dictionary*, Tehran: Tehran University Publication.
- [4] Kord Za'feranlu Kambuziya , Aliye (2006), *Phonology: Rule Based Approach*, Tehran: Samt.
- [5] Kord Za'feranlu Kambuziya , Aliye& Taj Ababdi, Farzane (2012), Comparative Study of Gemination and Degemination in Persian and Iranian Dialects, *Journal of Language and Comparative Linguistics*: 3: 9, P: 127-152.
- [6] Mc. Carthy, John, J. (2004), *Optimality Theory in Phonology*, USA: Blackwell Publishing.
- [7] Mc. Carthy, John, J. (2008), *Doing Optimality Theory, Applying Theory to Data*, USA: Blackwell Publishing.
- [8] Miremadi, Ali (1994), The Role of Gemination in Persian, *Tehran: AlameTabatabayi University, Second Linguistic Conference*.
- [9] Najafi, Abolhasan (2011), *Basic Linguistics and its Usage in Persian*, Tehran: Nilufar.
- [10] Sadeghi, Ali Ashraf (2013), Gemination in Persian, *Lexicography*: 5 & 6: P 43-3.
- [11] Samare, Yadolah (1985), *Persian Language Phonetics* (2nd ed.), Tehran: Nashr-e Daneshgahi.
- [12] Vahidian Kamyar, Taqi (1992), Analyzing Gemination to Solve a Dictation Problem, *Journal of Persian Language Advancement*, 7: 30, P 24-27.
- [13] Zabih Nia, Omran (2012), Gemination: Necessary to be Written or Not in Persian, *Rhetorical Language Studies*, 2: 4.
- [14] Zirak, M & Peter, M. Skaer (2013), Evidence of Gemination in Persian (Phonetic and Phonological Study of Lexical and Post Lexical Geminates), *retrieved from* <http://www.hiroshima-u.ac.jp/upload/85/others/kiyou/2013/ningen2013-2.pdf> on 10/4/2016.