ABSTRACT

This paper addresses a question concerning phonological operations involved standard Farsi vowel Compensatory Lengthening (CL) due to glottal consonant deletion in syllabic coda. Specifically, I question two current views related to CL process. One view suggests that glottal allophonic weakening is compensated by vowel lengthening. Another view suggests that glottal consonant is deleted, but vacated mora is filled by preceding vowel. Based on an experimental approach, I claim that Farsi CL is a gradient process in which different magnitude of glottal gesture variation is accompanied by F0 decreasing of vowel onset and offset and lengthening of the vowel. Evidence for this claim comes from the results of two experiments. For experiment 1, ten sentences having ten CaC words were made. /a/ duration was doubled while each synthetic CaaC word supposed to be a phonetic counterpart of another CaCC lexical word. Ten other sentences were generated by such CaaC words. Ten subjects listened to these twenty sentences. Subjects' lexical ambiguity to understand CaC and CaCC words was 37.8%. For experiment 2, subjects colloquially uttered those twenty sentences with CaC and CaCC words. I found that, in addition to significant vowel lengthening for colloquial CaCC, /a/ onset and offset F0 was significantly decreased, while different magnitude of glottal gesture was realized in speech signal from a weak through complete deletion of glottals. This finding can be used to improve Farsi ASR and TTS systems.

1. INTRODUCTION

Phonologists consider CL as a widespread phonological process in which the deletion of one segment triggers the lengthening of another. Various manifestations of CL in different languages are described using timing unit theories such as moraic, CV and X theories [1]. Farsi (Persian language), native language of Iranians, has a kind of CL which involves the deletion of a coda glottal consonant accompanied with the lengthening of the adjacent nucleus vowel, according to a general agreement among Farsi phonologists. The phonological question I deal with is to what extent this deletion and lengthening is significant as two intermediate operations in the CL process. This paper is to demonstrate that instead of the complete deletion of glottal consonant, a kind of significant F0 decreasing occurs in parallel with vowel lengthening. Additionally, Farsi listeners are to resolve a type of Sapir’s phonetic illusion as a result of CL process using some acoustic cues. To support this hypothesis, two experiments are given. This finding can be used to enhance the quality of Farsi synthesizers and ASR systems.

2. CL PHONETICS

To explain my view, a background for CL phonetics is needed. Farsi contains six vowels : /i/ , /e/ , /a/ , /u/ , /o/ , /A/ , and two glottal consonants / ^ and / H / plus twenty-one other consonants. / l / , / u / and / A / are long vowels , insensitive to CL process, and / e / , / o / and / a / are short vowels , sensitive to CL process. The syllabic structure is formulated as CV(C)(C), although some phonologists argue the optionality of syllabic onset due to predictability of glottal stop in onset position [2]. Main stress falls on the final syllable of all words except for some verbs with specific inflectional prefixes. Word order is relatively free in colloquial Farsi , but conservative SOV in formal Farsi.

<table>
<thead>
<tr>
<th>CVCC</th>
<th>CV: C / CVCC</th>
<th>CVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ba:d</td>
<td>&quot;after&quot;</td>
<td>bad : &quot;bad&quot;</td>
</tr>
<tr>
<td>rob?</td>
<td>&quot;horror&quot;</td>
<td>rob : rob^</td>
</tr>
<tr>
<td>sa:d</td>
<td>&quot;prosperity&quot;</td>
<td>sad / ssa:d</td>
</tr>
<tr>
<td>tab?</td>
<td>&quot;nature&quot;</td>
<td>tab / tab^</td>
</tr>
<tr>
<td>man?</td>
<td>&quot;prevention&quot;</td>
<td>man / man^</td>
</tr>
<tr>
<td>rob?</td>
<td>&quot;quarter&quot;</td>
<td>rob / rob^</td>
</tr>
<tr>
<td>bahs</td>
<td>&quot;argument&quot;</td>
<td>bahs / bahs</td>
</tr>
<tr>
<td>sehr</td>
<td>&quot;magic&quot;</td>
<td>sehr / sehrH</td>
</tr>
<tr>
<td>zahr</td>
<td>&quot;poison&quot;</td>
<td>zahr / zahrH</td>
</tr>
<tr>
<td>sahm</td>
<td>&quot;share&quot;</td>
<td>sahm / sahm</td>
</tr>
<tr>
<td>tarh</td>
<td>&quot;project&quot;</td>
<td>tar / tarH</td>
</tr>
<tr>
<td>sharh</td>
<td>&quot;explain&quot;</td>
<td>shar / sharH</td>
</tr>
<tr>
<td>sar?</td>
<td>&quot;epilepsy&quot;</td>
<td>sar / sar^</td>
</tr>
</tbody>
</table>

Table1: Farsi CL data [3],[4]. [^] and [H] are used to denote weak allophones of /i/, as glottal stop, and /h/, as glottal fricative. sh and : are used to denote unvoiced alveo-palatal fricative and length.
word with no glottal consonant after vowel in Farsi lexicon as shown in table1. No long vowel occurs in V position of CL data. All CL phonetic data in Farsi are originally Arabic. The words in each row of CVCC and CV column in table1 contrast gesturally in the sense of [5] that glottal gesture is present in CVCC words, but absent in CVC words. The words in middle column are two versions for phonetic realization of CVCC column words due to vowel CL process.

The second group contains polysyllabic words, which are not of my interest here.

3. CURRENT VIEWS

Two views are in debate among Persian phonologists for CL process. The first view [3] maintains that vowel CL occurs as a result of glottal consonant weakening. Therefore, /h / and /?/ in coda position changes to weak allophones of [H] and [?], as represented below:

```
\begin{array}{c}
\text{?} \\
\hline
\text{H} \\
\text{C}
\end{array}
```

The second view [4] maintains both moraic and CV phonological analysis following Hock [4], to explain vowel CL. process. According to this view, glottal consonant deletes from the coda, but remaining mora is reassigned to the preceding vowel, as shown below:

```
\begin{array}{c}
\text{X} \\
\text{V} \\
\text{C} \\
\text{V}
\end{array}
```

X stands for any consonant, z for a non-glottal consonant, y for vowels and G for glottal consonants. Syllabic position of z and G can be interchangeable.

Briefly, regarding glottal gesture, the first view description is based on weakening, but not deletion, of glottal consonant which ends to vowel CL. Samareh [3] asserts that audition of the weak glottal allophones for Farsi listeners is almost impossible. This view, in CL analysis, does not care about magnitude reduction of glottal gesture, but pays attention to satisfy crossing association line prohibition constraint of generative nonlinear phonological theory. Two versions of phonetic representations are shown in CV-C / CVCC column of table1. I claim that what involves with CL is better interpreted as a quantitative variation of glottal gesture in coda position. To test this hypothesis I conducted two experiments.

3. EXPERIMENT 1

3.1 Design

Ten monosyllabic CaC words, so called type1 words indicated in table1, were selected. The words were embedded in ten sentences such that the main stress of the sentences normally placed to the embedded words. One educated native speaker with no linguistics knowledge was asked to read aloud the sentences. lal duration was doubled by copying each cycle of lal speech signal to the end of the cycle itself using a DSP software. In this way, ten type2 words of the form CaaC were made each of which was assumed to be an acoustic realization of one CaCC Farsi lexical word, (as predicted by CL theory), indicated in CVCC column of table1. For example, [ zaar ] , which can be an acoustic realization of the lexical word / zahr / “poison”, was made from [ zar ] “gold”. Hereby, ten other sentences were made by replacing words type1 with word type2. For instance, the sentence “ be ?u zaar dAd” , which supposed to be understood by listeners as “ he / she poisoned him / her”, was made from “ be ?u zar dAd” , which means “ he / she gave him / her gold “. By replacing [ zar ] with [ zaar ]. ( A is used to denote Farsi long low back vowel ). Then, ten Farsi educated subjects were asked to listen to random playing of such twenty sentences in off-line mode, and to write down meaning of each sentence. They were also asked to write down the most probable meaning in cases where they confused to identify the correct sentence between two related minimal sentences.

3.2 Results

Subjects were able to understand sentences with 84% correction rate, while they confused to make correct decision in 45% of cases. Therefore, subjects had no lexical ambiguity to identify CaCC words from corresponding CaC words in 84% * 55% = 46.2% of cases. I concluded that CL involved other phonetic parameters. The question arised was how phonologically this 37.8%, i.e. (84% * 45%), lexical ambiguity can be explained. Author, as a native speaker, observed that the crucial phenomenon relevant to the question which is neglected by phonologists comes from vowel F0 decreasing in phonetic realization of CaCC words. To test this hypothesis, experiment2 was conducted.

4. EXPERIMENT 2

4.1 Design

Ten subjects of experiment1 were asked to utter colloquially twenty sentences of which ten sentences were made by CaCC words and ten others by CaC words of table1, with main stress placed on the words in question. Vowel F0 contour was measured using pitch extraction software written based on cepstrum algorithm. Values of phonetic parameters were computed for /a/ vowel of both CaCC and CaC words by visual examination for all subjects using software facilities: (1) vowel onset F0 : F0(1); (2) vowel offset F0 : F0(2); (3) vowel duration : (L). Difference of vowel offset F0 from vowel onset F0, i.e. df12, for all words and vowel onset F0 difference for each related CaC and CaCC, i.e. df11, were computed. Figure 1 shows F0 contours for two embedded words [man] and [man?]. Table2 shows a part of computations for just three speakers. Three null hypotheses were defined:

(1) Means of the vowel onset F0 for CVC and CVCC words are the same.
(2) Means of the difference of the vowel offset F0 from the vowel onset F0 for CVC and CVCC words are the same. (3) Means of the vowel duration for CVC and CVCC words are the same.

Figure 1: F0 contour for two embedded words [man] and [man?]. More vowel duration and glottal stop deletion is seen for [man?]. Signal and F0 contour alignment shows phonetic parameter values for two words. Three hypotheses were tested according to statistics drawn from samples.

4.2 Results

A t-test is used to determine if the means in three hypotheses are significantly different. All three hypotheses were rejected with p<.05 such that for hypothesis 1 : t(99) = 6.96, hypothesis 2 : t(99) = 6.42, and hypothesis 3 : t(198) = 10.1. I concluded that glottal weakening or deletion undeniably changes overall shape of vowel F0 contour in a way that hypothesis testing showed. This is what CL current views neglected to account for. The question is how to theorize this finding within phonological framework.

5. DISCUSSION AND CONCLUSION

In this paper my effort was to address the question of phonological operations which involve vowel CL process. To answer question, two views were presented: allophonic and moraic. Allophonic theory predicts that glottal weakening causes vowel CL and rejects complete deletion of glottals. Moraic theory predicts that glottal consonant deletion causes vowel CL such that phonological length of the words in question is preserved, i.e. structural length is independent from segmental glottals [6]. However, results of the phonetic experiments revealed that there is one acoustic cue other than vowel lengthening that is preserved significantly in speech signal, i.e. F0 decreasing of vowel onset and offset. In general, phonetically typical glottal gestures occur just in the words containing glottal stop or fricative gemination, as in Arabic [7]. Since Farsi phonotactic constraints in syllabic coda does not allow glottal consonant gemination, therefore target value for glottal constriction degree would not lead to a complete closure for glottal stop and a critical value for glottal fricative. To account for this phenomenon, I examined the acoustic structure of subjects’ coda glottals in speech signal which observed in expanded waveform and spectrogram of experiment2 data:

<table>
<thead>
<tr>
<th>word</th>
<th>SPEAKER1</th>
<th>SPEAKER2</th>
<th>SPEAKER3</th>
</tr>
</thead>
<tbody>
<tr>
<td>man</td>
<td>125</td>
<td>123</td>
<td>126</td>
</tr>
<tr>
<td>man2</td>
<td>125</td>
<td>123</td>
<td>126</td>
</tr>
<tr>
<td>F0</td>
<td>125</td>
<td>123</td>
<td>126</td>
</tr>
<tr>
<td>df11</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>df12</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

(1) Glottal stop: either the different degree of glottal pulse irregularity for constricted voicing as a typical cue of creaky voice or complete glottal deletion was seen on the end of vocalic portion [8]. Although some subjects with Arabic background pronounced glottal stop with voiced pharyngeal fricative articulation, i.e. a pharyngealized vowel was produced.

(2) Glottal fricative: either the different degree of breathy voice or complete glottal deletion was seen on the end of vocalic portion.

Therefore, phonetic evidences provide counter-argument for both allophonic and moraic theories. As I argued above, a significant finding for CL data acoustic structure was nucleus vowel onset and offset lower F0 in comparison with that of non-CL data. To construe the results of phonetic experiments within phonological theory, I suggest that in Farsi CL it is more appropriate to regard glottal consonants as segments that have only a laryngeal gesture specification that orally are unmarked and dependent to preceding vocalic gesture. To any degree that
glottal gesture is not satisfied to reach the intended target, i.e. stop or frication, through CL process, and therefore a laxer laryngeal setting is occurred, glottal consonants come close to vocalic gesture which ultimately causes more length for preceding vowel. Therefore, contrary to moraic phonology, my suggestion predicts that vowel structural length is dependent to glottal segments in coda position. This finding strongly confirms the discussion of phonetics and phonology integration [9],[10]. Another message of this paper is of course for Farsi words, both vowel length and vowel lower F0 for CL words are ASR community, that to identify CL words from non-CL words, both vowel length and vowel lower F0 for CL words are phonetically significant acoustic cues. For future work, experiment 1 will be conducted with vowel lower F0 synthesized into consideration alongside of vowel lengthening procedure.

6. REFERENCES