Variable hiatus in Persian

Koorosh Ariyaee and Peter Jurgec (University of Toronto)

Introduction. A great deal of recent phonological research has focused on variable sound patterns. For instance, Becker et al. (2011) demonstrate that laryngeal alternations in Turkish depend on word length and consonant place. In this paper, we contribute to this literature by looking at variable hiatus resolution in Persian. Like many languages, Persian generally disprefers vowel sequences, but it does so variably. Moreover, while many languages prefer elision of the first vowel (V_1), Persian typically elides V_2 (Casali 1997, 1998). This creates a pressure for suffixes that consist of a single vowel to be deleted entirely. We conducted a production and perception experiment which confirm that with monosegmental suffixes hiatus is in fact tolerated.

Background. In Spoken Persian, when vowel-final roots are followed by vowel-initial suffixes, the realizations vary: V_2 can be deleted, [ʔ] can be epethesized or hiatus is preserved, as shown in (1). However, not all possibilities are equally attested: if the suffix consists of a single vowel, elision is not clearly grammatical ($^{???/*}$baba-∅), and V_1 elision is always ungrammatical.

(1)

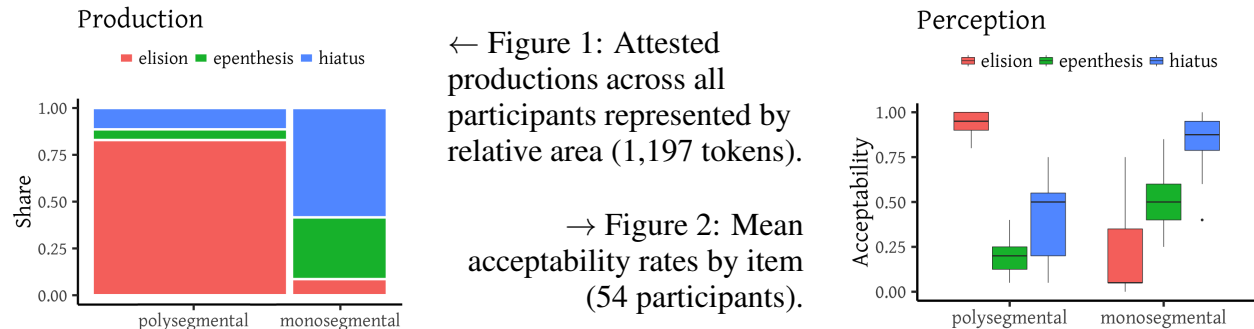|	'our'	'his/her'	'my'	'the'		
dæftær	dæftær-emun	dæftær-eʃ	dæftær-æm	dæftær-e	'office'	
baba	baba-mun	baba-ʃ	baba-m	$^{???/*}$baba	'dad'	(V_2 elision)
	baba-ʔemun	baba-ʔeʃ	baba-ʔæm	baba-ʔe		(epenthesis)
	baba-emun	baba-eʃ	baba-æm	baba-e		(hiatus)
	*bab-emun	*bab-eʃ	*bab-æm	*bab-e		(V_1 elision)

The Persian hiatus pattern is cross-linguistically rare. First, there are only two other known languages that have V_2 elision while never having V_1 elision (Casali 1997). Second, very few reported languages exhibit variation in which hiatus is allowed, but also variably resolved in two different ways (Garrido 2013). The variability is mirrored in the existing literature on Persian hiatus: while Sadeghi (1986), Shaghaghi (2000), Dehghan and Kord (2012) focus on epenthesis, Jam (2015) studies elision, and Estaji et al. (2010), Yazarlou (2014) suggest hiatus is retained.

Production. To explore the variation and its restrictions, we conducted a word-formation production experiment. Seven participants were provided with a consonant-final root and its derived form with a vowel-initial suffix, and they were then asked to derive vowel-final roots with the same suffix. The stimuli consisted of 171 words (V-final roots followed by V-initial suffixes), which were selected depending on the following variables: V_1, V_2, and suffix length. Since there is no corpus of Spoken Persian, the aim of this experiment was to gauge the extent of variation rather than to tightly control for all factors. We were also restricted by lexical gaps: for instance, there are no productive *u*-initial suffixes or polysegmental *a*-initial suffixes.

The results, confirmed by an acoustic examination, reveal that V_1 elision is never attested, and V_2 elision is by far the most common. The suffix length matters: V_2 elision (henceforth, simply elision) is extremely rare with monosegmental suffixes, as shown in Figure 1. The vowel combination matters, but only in the choice of the epenthetic segment, which is dependent on the first vowel: variable *j*-epenthesis is attested with [i] while ʔ-epenthesis is found elsewhere. The participants also produced variant realizations which are in line with the data in (1). For example, hiatus was common with monosegmental suffixes, but also attested with longer suffixes.

Perception. We next ask whether the generalizations observed in real words are extended to nonce words in a perception experiment. Our auditory stimuli were perfectly balanced: we chose 30 nonce roots that were derived with 6 suffixes (3 polysegmental and 3 monosegmental). Every one of 54 native speakers heard 30 of these words, each under three conditions (elision, epenthesis, hiatus; randomized). At each trial, the participants were asked to judge the paradigm (root + derived

word) as acceptable or not. The results indicate that V_2 elision is the most acceptable variant with polysegmental suffixes, whereas hiatus is the most acceptable variant with monosegmental suffixes (Figure 2). We fit the acceptability in a fully-crossed mixed-effects logistic regression model with Condition (Helmert coded: elision vs. other, hiatus vs. epenthesis) and SuffixLength (simple coded) as fixed effects, and Item and Participant as random effects. Elision had significantly higher acceptability than the other two conditions ($z = 2.95, p = .003$); at the same time, elision had significantly lower acceptability rates with monosegmental suffixes ($z = -18.24, p < .001$).

Production

← Figure 1: Attested productions across all participants represented by relative area (1,197 tokens).

→ Figure 2: Mean acceptability rates by item (54 participants).

Perception

Analysis. We fed the results of the perception experiment to a Maximum Entropy learner (Goldwater and Johnson 2003; Hayes and Wilson 2008), with four key constraints shown in (2). REALIZEMORPHEME ($\equiv$ Morphemes must have output realizations; Kurisu 2001) was assigned the highest weight, followed by DEP and *HIATUS. With polysegmental suffixes, the elision candidate (2-a) is preferred over the other two, while with monosegmental suffixes, hiatus (3-c) is the most common. Regardless of the suffix length, the probability of hiatus (c) is estimated at rates 1.8-times higher than the epenthesis (b). This follows directly from the violation profile of these candidates which are identical in (2) and (3). This prediction closely matches the perception and production data, suggesting that the proposed constraints are adequate.

(2)

/hutʃɑ-emun/	REALIZEMORPHEME $w = 2.2$	DEP $w = 1.5$	*HIATUS $w = 0.9$	MAX $w = 0.0$	$\mathcal{H}$	p
a. hutʃɑmun				-1	-0.0	.61
b. hutʃɑʔemun		-1			-1.5	.14
c. hutʃɑemun			-1		-0.9	.25

(3)

/hutʃɑ-e/	REALIZEMORPHEME $w = 2.2$	DEP $w = 1.5$	*HIATUS $w = 0.9$	MAX $w = 0.0$	$\mathcal{H}$	p
a. hutʃɑ	-1			-1	-2.2	.14
b. hutʃɑʔe		-1			-1.5	.31
c. hutʃɑe			-1		-0.9	.55

Discussion and conclusions. We found that elision is the most common resolution with polysegmental suffixes, but rare with monosegmental suffixes. Hiatus and epenthesis are both possible, but the former is more frequent than the latter, regardless of suffix length. Even though V_2 elision is cross-linguistically rare, Casali (1997, 1998) nevertheless predicts such a pattern in classic OT, and our results confirm its generality and productivity. In sum, this is the first experimental study showing that hiatus resolution can be a variable process, where a range of phenomena typically found across languages (elision, epenthesis, and hiatus) are observed in a single language.