CiV Lengthening: Productivity and the Emergence of the Unmarked

Adeline Tan

1. Introduction

In English CiV Lengthening, non-high lax vowels are reportedly not allowed in the _CiV environment when a morpheme boundary intervenes (Chomsky & Halle 1968, Baković 2013). For example, *Caˈn[æ]d-ian with the lax [æ] is ill-formed while Caˈn[eɪ]d-ian with the tense [eɪ] is grammatical. This restriction against lax vowels does not apply to non-derived words. For instance, monomorphemic ˈc[æ]meo is grammatical because no morpheme boundary occurs within the [æ]CiV sequence.

At first blush, CiV Lengthening appears to be a derived environment effect because it is blocked from applying in non-derived environments like *cameo*. However, CiV Lengthening is subject to an additional blocking effect. Burzio (2005) observes that CiV Lengthening is similarly blocked in derived environments when the target vowel is not newly stressed. Consider the stem *Orwell* (ˈOrˌw[ɛ]ll), which retains the lax [ɛ] upon *-ian* affixation: ˌOrˈw[ɛ]ll-ian, *ˌOrˈw[i]ll-ian. This contrasts with the stem *Canada* (ˈCan[ə]da), which disallows the lax [æ] in the very same derived environment: Caˈn[eɪ]d-ian, *Caˈn[æ]d-ian.

The two reported generalizations regarding the applicability of CiV Lengthening in derived environments are summarized in (1):

(1) a. CiV Lengthening applies to a newly-stressed vowel upon affixation with a CiV Lengthening suffix.
   b. CiV Lengthening is blocked from applying to a non-newly stressed vowel upon affixation with a CiV Lengthening suffix.

It should be noted that *-ian* is not the only CiV Lengthening suffix. Other CiV Lengthening suffixes include *-ious* (*e.g.* mel[ə]dy: mel[oʊ]d-ious, *mel[ɑ]d-ious) and *-ial* (*e.g.* cust[ə]dy: cust[oʊ]d-ial, *cust[ɑ]d-ial). The relevant tense-lax pairs for CiV Lengthening are the very same ones observed for Trisyllabic Shortening (TSS). Take for instance the TSS pair [oʊ]~[ɑ][1]. CiV Lengthening has applied to mel[oʊ]d-ious. Yet, under different phonological conditions, the same stem shows up with the lax [ɑ]: mel[ɑ]dic.

The rest of the paper is organized as follows. In §2, I report experimental results that show that the two major generalizations regarding CiV Lengthening in (1) are productive. In doing so, I refine the generalization in (1b) from one of blocking to a retention effect. In the case of CiV Lengthening, the retention effect creates a challenge to the traditional derived environment effect story. Since the retention effect is synchronically active, I propose an alternative analysis of CiV Lengthening, in which the application of CiV Lengthening in (1a) is treated not as a case of a derived environment effect, but rather as a case of the emergence of an unmarked preference for stressed vowels to be heavy. This analysis is presented in §3. The paper closes with a general discussion in §4.

2. Experiment

The two generalizations to be tested are shown in (1). In order to generate the relevant experimental conditions, it is necessary to take a closer look at each of these generalizations.

* Adeline Tan, University of California, Los Angeles, aderaetan@ucla.edu. Thank you to Kie Zuraw as well as audiences at the UCLA Phonology Seminar and the West Coast Conference on Formal Linguistics 41 for helpful feedback and comments.

[1] *e.g.* prov[oʊ]k, prov[ɑ]cative.

 Proceedings of the 41st West Coast Conference on Formal Linguistics, ed. Nikolas Webster et al., 574-583. Somerville, MA: Cascadilla Proceedings Project.

Let us first turn our attention to the generalization in (1a), which states that "CiV Lengthening applies to a newly-stressed vowel upon affixation with a CiV Lengthening suffix". In order for the target vowel to become newly-stressed in the affixed form, it must have been unstressed in the unaffixed form. In English, this means that the vowel must have been a reduced vowel in the unaffixed form. Since CiV Lengthening has been reported be restricted to the **non-high** vowels, I will restrict the set of reduced vowels to [ə]. In other words, all unaffixed forms relevant to the generalization in (1a) can only bear [ə] as their target vowel. Examples of such unaffixed test stems include [ˈkudəb] and [ˈziləv]. This thus forms the first of the experimental conditions: the ə-stems. In order for the generalization in (1a) to be found productive, the ə-stems must show a preference for a tense vowel upon affixation with a CiV Lengthening suffix.

The generalization in (1b) states that "CiV Lengthening is blocked from applying to a non-newly stressed vowel upon affixation with a CiV Lengthening suffix". In this case, the target vowel must have borne some level of stress (primary or secondary) in the unaffixed form. In other words, the target vowel must have been a "full" vowel (*i.e.* not a reduced vowel) even in the unaffixed form. We have already seen that full lax vowels retain their laxness upon affixation (*e.g.* ˈOrˌw[ɛ]ll: ˌOrˈw[ɛ]ll-ian, *ˌOrˈw[i]ll-ian). While generally not discussed for CiV Lengthening, it can be seen that full tense vowels retain their tenseness upon affixation too (*e.g.* ˈMonˌg[oʊ]l: ˌMonˈg[oʊ]l-ian, *ˌMonˈg[ɑ]l-ian). Thus, an alternative way to characterize the blocking effect is to regard it as a retention effect. That is to say, there is pressure for the tenseness value of full vowels in the unaffixed form to be retained upon affixation with a CiV Lengthening suffix. To test for a retention effect, two additional conditions are necessary: the lax-stem condition & the tense-stem condition. Examples of unaffixed lax-stems include [səˈdæl] and [ˈtɪˌdɑn]. Examples of unaffixed tense-stems include [pəˈboʊk] and [ˈzuˌseɪm]. For the retention effect to be productive, the tense-stems should exhibit a greater preference for a tense vowel upon affixation than their lax-stem counterparts do.

The generalizations in (1) have thus been operationalized to the following hypotheses (2):

(2) a. ə-stems prefer a tense vowel upon affixation with a CiV Lengthening suffix.
  b. Upon affixation with a CiV Lengthening suffix, tense-stems prefer a tense vowel at a higher rate than their lax-stem counterparts.

2.1. Stimuli

The three experimental conditions and the general shape of the test stems for each of these three conditions are introduced in §2.1.1. Specific detail regarding how the individual test items were generated is presented in §2.1.2.

2.1.1. Conditions

Stems were of the shape $C_1V_1C_2V_2C_3$. The target vowel was V_2, V_2 being the vowel whose tenseness might change upon affixation. All experimental conditions were reflected only on V_2. All other segments were irrelevant to experimental manipulation.

For test stems, the target vowel, V_2, was subject to the following manipulations. There were three tenseness conditions: ə-stems, tense-stems, and lax-stems. Tenseness was crossed with backness such that within each of the tenseness conditions, half of the stems had a back target vowel and half had a front target vowel. For ə-stems, the backness condition was visible on the affixed options (but not on the unaffixed stem). When unaffixed, the V_2's of ə-stems were necessarily unstressed. In contrast, the V_2's of tense-stems and lax-stems had to bear stress. There were two stress-levels such that half of the V_2's of tense-stems had primary stress and the other half had secondary stress. Stress-level was similarly crossed with the lax-stem condition. This resulted in a total of 96 test stems. Table 1 summarizes the distribution of stems for each experimental condition.

Tenseness	Stress	Backness	
		Front	Back
ə-stem	None	16 [ˈkudəb]	16 [ˈziləv]
Tense-stem	Primary	8 [jəˈteɪk]	8 [pəˈboʊk]
	Secondary	8 [ˈzuˌseɪm]	8 [ˈtæˌsoʊp]
Lax-stem	Primary	8 [səˈdæl]	8 [ɹəˈlɑn]
	Secondary	8 [ˈnɪˌmæb]	8 [ˈtɪˌdɑn]

Table 1: Distribution of test stems across conditions.

2.1.2. Materials

CiV Lengthening was reported to be restricted to the non-high vowels. Hence, I decided to restrict the vowel pairs to the non-high vowels. The vowel pairs for the front and back conditions were [eɪ]~[æ] and [oʊ]~[ɑ] respectively.

I wrote a script to randomly generate the rest of the test stems (*i.e.* C_1, V_1, C_2, and C_3). The set of onsets (C_1 & C_2) were [p, t, k, b, d, ɡ, m, n, f, θ, s, ʃ, h, v, ð, z, ʒ, t͡ʃ, d͡ʒ, ɹ, l, j, w]. The set of codas (C_3) were [p, k, b, ɡ, m, n, f, θ, ʃ, v, ð, ʒ, l][2]. The set of vowels were [i, ɪ, eɪ, ɛ, æ, ɑ, ʌ, oʊ, ʊ, u, aɪ, aʊ, ɔɪ]. The frequency at which the script picked a sound depended on the sound's unigram frequency in the CMU Pronouncing Dictionary. Stems and affixed options that sounded like real words were rejected.

The 2AFC affixed options were generated by modifying the test stems in the following ways (3):

(3) a. The *-ian* [iən] suffix was added.
    i. All affixed options carried primary stress on the target vowel.
  b. Target vowels were modified such that for each 2AFC pair, one option would carry a tense target vowel while the alternative option would carry its lax counterpart.
    i. For ə-stems, half of the stems showed the front pair in the affixed forms while the other half showed the back pair in their affixed forms.

For example, the ə-stem [ˈkudəb] would have these two affixed options: [ˌkuˈdeɪbiən], [ˌkuˈdæbiən]. The lax-stem [ˈtɪˌdɑn] would have the two affixed choices: [ˌtɪˈdoʊniən], [ˌtɪˈdɑniən]. The tense-stem [jəˈteɪk] would have the affixed options: [jəˈteɪkiən], [jəˈtækiən].

There were 48 filler stems. The target vowels for the unaffixed filler stems were restricted to [ə] as well as the [aɪ]~[ɪ] and [aʊ]~[ʌ] pairs. The distribution of filler target vowels was identical to that of the test stems (albeit halved).

The stimuli were recorded by a phonetically-trained male native speaker of American English who was naïve to the research questions. Word-final stops were released. Otherwise, the pronunciation was as in American English.

[2] The sounds [t, d, s, z, t͡ʃ, d͡ʒ] were removed from the set of codas for test stems. This was done to sidestep potential confounding factors that might arise from *yod*-coalescence. The sounds [h, j, w] were removed too because they couldn't serve as codas. [ɹ] was excluded because tenseness is neutralized before an [ɹ] coda. Since C_3 becomes an onset upon *-ian* affixation, [ŋ] was excluded because it could not serve as an onset.

2.2. Procedure

The Experigen software (Becker & Levine 2014) was used to run the experiment online. Participants were instructed to find a quiet space and to put on headphones before beginning the experiment. They were then presented with the audio of two practice stems (*Portugal* and *Vietnam*), and given feedback on whether they correctly identified these words. Participants could play these practice stems multiple times to help adjust their headsets to a comfortable and clear volume.

The experiment began with two practice trials. In each trial, they saw two sentences, which each had a missing word. The first sentence had a missing noun while the second sentence had a missing adjective. Participants were then presented with a pair of audio files. Each audio file contained a noun~adjective sequence. For example, one audio file would play *Portugal~Portuguese* while the other would play *Portugal~Portulese*. Participants then picked the audio option that they preferred. An example trial screen is shown in Figure 1.

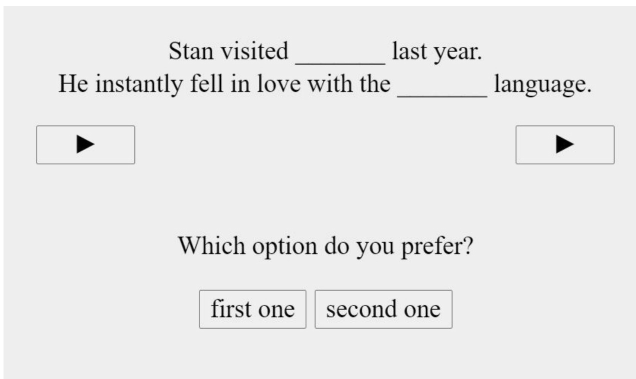

Figure 1: Trial screen.

The two practice trials were followed by the test and filler trials. Participants were informed that they might now encounter words that they were not familiar with, and that repeating the words out loud might aid them in the task. Half the trials had the tense affixed option played first while the other half played the lax affixed option first. The trials were shuffled, so the order of presentation of audio files (tense first or lax first), the order of stems, and the sentence-frame-to-audio-file pairings were different for each participant.

2.3. Participants

Participants were recruited from the UCLA Psychology Subject Pool. I excluded participants who were not native English speakers, as well as those who had taken more than three linguistics classes. After exclusions, a total of 24 participants remained.

2.4. Results

The proportion of tense responses was .71 for the ə-stems, .75 for the tense-stems, and .56 for the lax-stems (Figure 2). The error bars in all figures in this paper represent the 95% confidence interval.

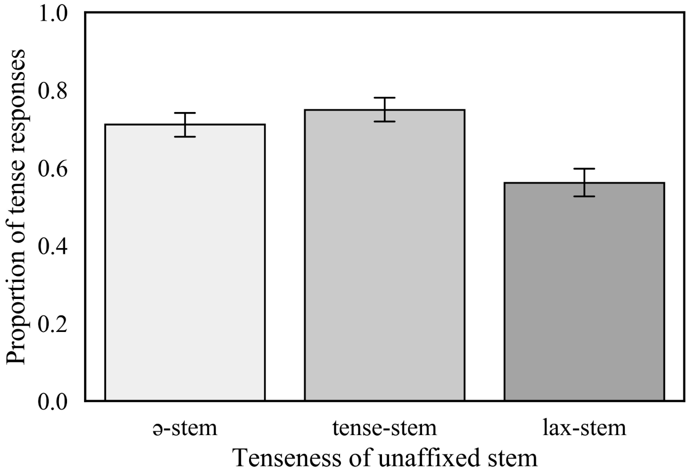

Figure 2: Proportion of tense responses for the three Tenseness conditions.

The proportion of tense responses was submitted to a mixed-effects logistic regression[3,4] with ə-stem set as the reference level. I included random intercepts for subjects and items. This model's parameters are presented in Table 2. For the ə-stem condition, the proportion of tense responses is indeed higher than chance ($\beta_0 = .97,\ p < .001$). This result suggests that a reduced vowel like [ə] prefers to become tense in the __CiV context.

Fixed effects	***Estimate***	***Std. error***	***p***
Intercept	.9702	.1429	<.001***
Tenseness = lax (*vs.* ə)	−.7002	.1350	<.001***
Tenseness = tense (*vs.* ə)	.2052	.1406	.144

Note: *p <.05, **p <.01, ***p <.001

Table 2: Fixed effects (Reference level = ə-stem).

I reformatted the mixed-effects logistic regression, this time setting tense-stem as the reference level. As before, random intercepts for subjects and items were included. The parameters of this model are presented in Table 3. For the tense-stems, the proportion of tense responses is higher than chance ($\beta_0 = 1.18,\ p < .001$). Compared to the tense-stems, lax-stems do in fact show a reduced porportion of tense responses ($\beta = -.91,\ p < .001$). This result suggests that for non-reduced vowels, the tenseness value in the unaffixed form has an effect on the tenseness value in the affixed form, with tense-stems preferring the tense vowel at a higher rate than their lax-stem counterparts.

Fixed effects	***Estimate***	***Std. error***	***p***
Intercept	1.1754	.1453	<.001***
Tenseness = ə (*vs.* tense)	−.2052	.1406	.144
Tenseness = lax (*vs.* tense)	−.9054	.1374	<.001***

Note: *p <.05, **p <.01, ***p <.001

Table 3: Fixed effects (Reference level = tense-stem).

[3] I used the `glmer()` function from the `lme4` package (Bates et al. 2015) in *R* (R Core Team 2018).

[4] These mixed effects logistic regression models only had one independent variable – unaffixed stem tenseness.

2.5. Discussion

Both of the hypotheses in (2) were supported, thus indicating that the generalizations in (1) were productive. ə-stems were found to prefer a tense target vowel upon affixation. This indicated that there was indeed a tense preference in the __CiV context when the original vowel was reduced. Tense-stems were found to have a greater tense preference than their lax-stems counterparts. This pointed towards the presence of a retention effect for non-reduced vowels, wherein a stem's original vowel quality was retained.

The proportion of tense responses for lax-stems was 0.56. This suggested that by themselves, lax-stems appeared to have a tense preference. This was, in fact, marginally so, though failing to meet the criterion for significance when α was set at .05[5]. This might seem surprising because if full vowels truly retained their original quality, we'd expect lax-stems to show a lax preference (all else being equal).

Nevertheless, there might be a confounding factor that skewed responses for all conditions towards the tensed affixed form. The experiment was designed to test participants' knowledge about the relation between the affixed form and it unaffixed counterpart. Stimuli were designed with this in mind; in order to nudge participants towards performing an affixation task (*i.e.* a wug task), the relevant unaffixed form was played immediately before the presentation of an affixed form. Despite this, a phonotactic task was also available in the background. To perform a phonotactic task, one need only compare the grammaticality of the two affixed forms without paying heed to the relation between the affixed form and its unaffixed counterpart. In a phonotactic task, the proportion of tense responses would reflect the goodness of tense vowels in the __CiV context (agnostic of morpheme boundary).

If participants had performed a mixture of the affixation and phonotactic tasks, we'd expect the vowel retention effects to be moderated by the phonotactic task. Consulting the Oxford English Dictionary, I found that the proportion of the tense vowels [eɪ, oʊ] out of the set of vowels [eɪ, æ, ɑ (& ɔ[6]), oʊ][7] in the __CiV context (agnostic of morpheme boundary) was .81.This indicated that any potential effect arising from an interfering phonotactic task should push responses towards the tense affixed form (and thus away from the lax affixed form). Given the potential interaction between the phonotactic and the affixation tasks, the trend towards a tense preference for lax-stems isn't unexpected, and should not constitute evidence against the retention of lax vowels in the __CiV context by lax-stems.

2.6. Interactions with backness & stress level

We have seen that the proportion of tense responses is dependent on the tenseness of the unaffixed stem. In this section, we will see that backness and stress level each moderate the relationship between a stem's unaffixed tenseness and the proportion of tense responses.

Back vowels amplify the retention effect of full-voweled stems (Figure 3). In a mixed effects logistic model[8] with random intercepts for item and subject, and with reference values set to "tense-stem, back, and secondary stress", the intercept was found to be significant ($\beta_0 = 1.26, p < .001$). Laxness ($\beta = -.92, p < .001$) and frontness ($\beta = -.61, p = .02$) each decreased the proportion of tense responses. However, being both lax and front led to a smaller decrease than expected ($\beta = .80, p = .02$). This indicated that front vowels had a smaller retention effect that their back-voweled counterparts did. This model found no other main effects or interactions at $\alpha = .05$. Model parameters which turned out to be significant at $\alpha = .05$ are shown in Table 4.

---

[5] For the mixed effects logistic regression model with lax-stem set as the reference level, there was a trend for the tense response to be higher than chance ($\beta_0 = .2700, p = 0.0518$).

[6] I restricted the set of vowels to [eɪ, æ, ɑ (& ɔ), oʊ] since these were the only target vowels used in the experiment.

[7] The /ɔ/ vowel does not exist in most varieties of American English. Since the /ɔ/ and /ɑ/ vowels have merged in General American English, I collapsed the /ɔ/ and /ɑ/ vowels into the /ɑ/ class. Regarding the experimental stimuli, the speaker produced all instances of /ɑ/ as [ɑ]. There were no stimuli with /ɔ/.

[8] The mixed effects models in this section include all three independent variables (unaffixed stem tenseness, backness and stress level) as predictors. In this regard, the models in this section differ from those in §2.4, which have the unaffixed stem tenseness as their sole independent variable.

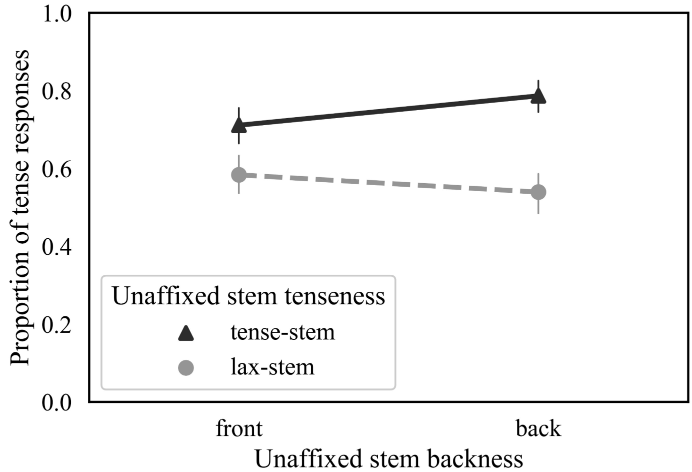

Figure 3: Back vowels amplify the retention effect.

Fixed effects	***Estimate***	***Std. error***	***p***
Intercept	1.2597	.2160	<.001***
Tenseness = lax (*vs.* tense)	−.9184	.2532	<.001***
Backness = front (*vs.* back)	−.6135	.2558	.017*
Tenseness = lax (*vs.* tense) & Backness = front (*vs.* back)	.7979	.3502	.023*

Note: *p <.05, **p <.01, ***p <.001

Table 4: Fixed effects (Reference level = tense, back & secondary-stressed stems).

The retention effect is likewise amplified for stems whose target vowels were primary-stressed in the unaffixed form (Figure 4). I reformatted the model with reference levels now set to "tense-stem, front, and primary stress". The intercept was significant ($\beta_0 = 1.32,\ p < .001$). In isolation, laxness ($\beta = -1.12,\ p < .001$) and secondary stress ($\beta = -.68,\ p = .01$) each reduced the proportion of tense responses. However, being both lax and secondary stressed led to a smaller reduction than expected ($\beta = 1.00,\ p = .005$), thus confirming the interaction between the unaffixed stem's tenseness & stress level. In particular, having a secondary-stressed target vowel in the unaffixed stem resulted in a smaller retention effect. This model found no other main effects or interactions. Model parameters that were significant at α = .05 are shown in Table 5.

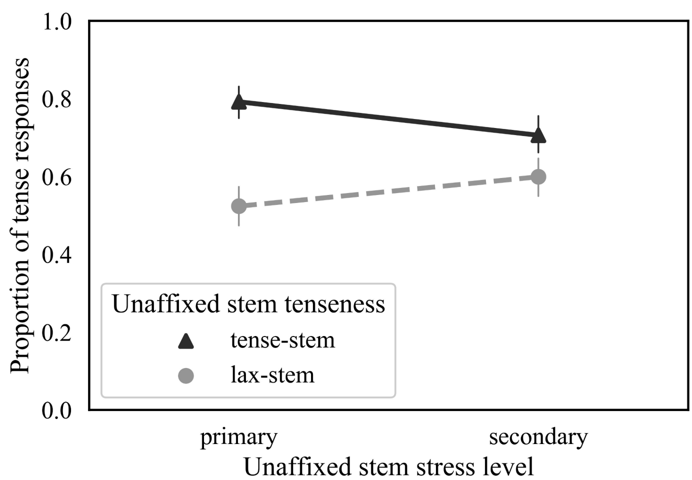

Figure 4: Primary-stressed target vowels amplify the retention effect.

Fixed effects	***Estimate***	***Std. error***	***p***
Intercept	1.3216	.2182	<.001***
Tenseness = lax (*vs.* tense)	−1.1156	.2545	<.001***
Stress = secondary (*vs.* primary)	−.6754	.2576	.009**
Tenseness = lax (*vs.* tense) & Stress = secondary (*vs.* primary)	.9951	.3511	.005**

Note: **p* <.05, ***p* <.01, ****p* <.001

Table 5: Fixed effects (Reference level = tense, front & primary-stressed stems).

We have seen that having a target vowel that is back or that bears primary stress in the unaffixed stem amplifies[9] the retention effect. Nevertheless, backness & stress level only affect the strength of the retention effect; they never reverse or eliminate them. The proportion of tense responses for tense-stems always remains higher than that of their lax-stem counterparts under all combinations of conditions.

3. CiV Lengthening as the Emergence of the Unmarked

In this section, I show that it is possible to model CiV Lengthening without any reference to derived environment effects. That is to say, I make reference to neither morphological derived environments (*e.g.* morpheme boundaries) nor to phonological derived environments (*e.g.* newly stressed vowels). Instead, my analysis hinges on the informativeness (or the lack thereof) of reduced vowels, and what it means to be faithful to a neutralized vowel.

I propose that CiV Lengthening is driven by a low-ranked markedness constraint: the Stress-to-Weight Principle (SWP). The SWP states that all stressed syllables must be heavy. The effect of this markedness constraint only emerges when the target vowel of the unaffixed form is a reduced vowel. When the target vowel of the unaffixed form is a full vowel, a higher ranking faithfulness constraint hides the effect of this markedness constraint.

Let us first consider the case of full-voweled stems. CiV Lengthening suffixes require primary stress to land on the antepenultimate syllable. I assume the presence of an undominated markedness constraint that rules out candidates that do not have primary stress on the antepenult, and so will not include this constraint or its corresponding candidates in the tableaux. The FAITHFULNESS constraint IDENT-OO [tense] requires that the tenseness value of a vowel in the BASE (here: the unaffixed stem) be identical to the tenseness value of its correspondent in the suffixed form. Thus, when lax-stems and tense-stems are suffixed, IDENT-OO [tense] acts to ensure that the target vowel remains unchanged, as in *Chadian* in (4) and *Ukrainian* in (5). The lower-ranked markedness constraint, SWP, has no say in picking the winning candidate. Hence, full-voweled stems retain the tenseness value of their unaffixed form upon affixation with a CiV Lengthening suffix.

(4)

[ˈtʃæd] + *ian*$_{suff}$	*ə́	IDENT-OO [tense]	SWP
a. ˈtʃeɪdiən		*!	
☞ b. ˈtʃædiən			*
c. ˈtʃədiən	*!	*	*

(5)

[juˈkɹeɪn] + *ian*$_{suff}$	*ə́	IDENT-OO [tense]	SWP
☞ a. juˈkɹeɪniən			
b. juˈkɹæniən		*!	*
c. juˈkɹəniən	*!	*	*

Let us now turn our attention to reduced-voweled stems. When a ə-stem word, like Can[ə]da, is suffixed, the CiV Lengthening suffix again requires primary stress to land on the antepenult. For ə-stems, this antepenultimate syllable's nucleus is [ə] in the unaffixed form (*i.e.* the BASE). Candidate (6c) merely shifts stress onto the target vowel while remaining faithful to the vowel quality. Candidate (6c) is ill-formed and ruled out by an undominated markedness constraint against stressed schwas. The two remaining candidates (6a) and (6b) do equally badly on the next highest ranking constraint, IDENT-OO [tense]. The tie between these two candidates is broken by the emergent markedness constraint, SWP, which requires that stressed syllables be heavy. Thus, the unmarked heavy stressed syllable emerges only when the target vowel was [ə] in the unaffixed surface form.

(6)

		[ˈkænədə] + *ian*$_{suff}$	*ə́	IDENT-OO [tense]	SWP
☞	a.	kəˈneɪdiən		*	
	b.	kəˈnædiən		*	*!
	c.	kəˈnədiən	*!		*

In English, reduced vowels like [ə] result from neutralizing-type vowel reductions. That is to say, the [+/− tense] distinction is lost when a vowel gets reduced to [ə]. Hence, I propose that the correspondence between [ə] and [+tense] is as equally good (or bad) as the correspondence between [ə] and [−tense]. One way to formalize this is to treat [ə] is [0tense], so corresponding [+tense] and [−tense] vowels are equally bad with respect to [ə]. (An alternative, if [ə] is treated as [−tense], is to replace IDENT-OO [tense] with IDENT-OO-V, a constraint that is satisfied only if the vowel in the suffixed form is featurally identical to the vowel in the BASE.)

A final set of words remains to be accounted for: the monomorphemic words. CiV Lengthening is reported to not apply to monomorphemic words, which means that there is no restriction on lax vowels showing up in the _CiV environment for these words. OO-Corr constraints are irrelevant for monomorphemic words, so IDENT-OO [tense] has been replaced with IDENT-IO [tense] in (7). Since IDENT-IO [tense] outranks SWP, the underlying lax /æ/ surfaces faithfully in ˈc[æ]meo.

(7)

		/ˈkæmiˌoʊ/	*ə́	IDENT-IO [tense]	SWP
	a.	ˈkeɪmiˌoʊ		*!	
☞	b.	ˈkæmiˌoʊ			*

4. General Discussion

I have provided a model of CiV Lengthening that falls out from the interactions of traditional markedness and faithfulness constraints. Nevertheless, my analysis does face a limitation. While my model of CiV Lengthening provides an account of vowel tenseness in the _CiV environment, it cannot predict other vowel qualities such as backness and height. For example, whilst the following target vowels are all tense, why do we get the mid front [eɪ] in Can[eɪ]dian, the high front [i] in Hand[i]lian, and the mid back [oʊ] in cust[oʊ]dial? Orthography is likely to play an important role in determining the backness and height of the target vowel since surface [eɪ, i, oʊ] are most closely associated with orthographic <a, e, o> respectively. A fuller account of CiV Lengthening would likely have to integrate orthographic effects into the analysis.

In my proposed analysis, CiV Lengthening is modeled via the interactions of simple markedness and faithfulness constraints. This contrasts with a more traditional derived environment effect analysis of CiV Lengthening (Steriade 2019). In constraint-based models such as Optimality Theory (Prince & Smolensky 1993/2004), derived environment effects have been analyzed via constraint conjunction (Lubowicz 2002) or comparative markedness (McCarthy 2003). Both constraint conjunction and comparative markedness increase the complexity of learning the relevant constraints of one's language. In the case of constraint conjunction, the child needs to learn which constraints should be conjoined. Comparative markedness constraints are more complex than traditional markedness constraints. While traditional markedness constraints only look at the output[9] in order to assess constraint violations, comparative markedness constraints need to look at both the output and the input in order to assess violations.

Nevertheless, simplifying the constraint machinery required to model CiV Lengthening does not come for free. Under my analysis, the featural representation increases in complexity. More specifically, the [tense] feature goes from being a binary-valued feature (+/−) to being a ternary-valued feature (+/−/0). In the case of reduced vowels, I have argued for the [0tense] status of [ə], given that it is a neutralized vowel in which the tenseness distinction is lost. That is to say, given a surface [ə], the underlying tenseness value cannot be recovered. Regarding non-reduced vowels, I forward that non-reduced (*i.e.* "full") vowels may only bear [+/−tense] (*i.e.* not [0tense]) because tenseness is contrastive in English. When affixation with a CiV Lengthening pushes stress onto [ə], [ə] must be repaired to a non-reduced vowel since stress is incompatible with [ə]. The ternary-valued [tense] allows for [ə] ([0tense]) to be faithful to neither [+tense] nor to [−tense]. Since faithfulness is unable to decide between the [+tense] and [−tense] full vowels, the winning [+tense] vowel (*e.g.* ˈCan[ə]da∼Caˈn[eɪ]dian) provides a rare opportunity to see the low-ranking SWP picking the winning candidate. In other words, for ə-stem words like *Canada*, the ternary-valued [tense] feature renders the higher-ranking faithfulness constraint impotent, thus enabling the oft-hidden preference for heavy stressed syllables to emerge.

In sum, my analysis shifts the complexity away from the complex constraints that drive the CiV Lengthening process and onto the featural representation.

References

Baković, Eric. 2013. *Blocking and complementarity in phonological theory*. Sheffield & Bristol, Conn.: Equinox Publishing.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. en. *Journal of Statistical Software* 67(1). https://doi.org/10.18637/jss.v067.i01.

Becker, Michael & Jonathan Levine. 2014. *Experigen – an online experiment platform*. http://becker.phonologist.org/experigen.

Burzio, Luigi. 2005. Sources of paradigm uniformity. In L. J. Downing, T. A. Hall & R. Raffelsiefen (eds.), *Paradigms in phonological theory*, 65–106. Oxford: Oxford University Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Lubowicz, Anna. 2002. Derived environment effects in Optimality Theory. *Lingua* 112. 243–80.

McCarthy, John. 2003. Comparative markedness (long version). In *Papers in Optimality Theory II [University of Massachusetts occasional papers in linguistics 26]*.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Steriade, Donca. 2019. CiV Lengthening and the weight of CV. In T Major M Bowler P Duncan & H Torrence (eds.), *Schuhschrift, papers in honor of Russell Schuh*. University of California.

[9] Here I use "output" to refer to the candidates in a tableau. The "input" refers to the representation in the top-left cell of a tableau. The "input" may be variously an underlying representation (as in the tableau in (7)) or the Base+suffix (as in the tableaux in (4-6)).

Proceedings of the 41st West Coast Conference on Formal Linguistics

edited by Nikolas Webster, Yağmur Kiper, Richard Wang, and Sichen Larry Lyu

Cascadilla Proceedings Project Somerville, MA 2025

Copyright information

ISBN 978-1-57473-483-6 hardback

Ordering information

Orders for the printed edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Tan, Adeline. 2025. CiV Lengthening: Productivity and the Emergence of the Unmarked. In *Proceedings of the 41st West Coast Conference on Formal Linguistics*, ed. Nikolas Webster et al., 574-583. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3792.