Partial Dependency of Vowel Reduction on Stress Shift —Evidence from English *-ion* Nominalization

Yuhan Zhang

Harvard University

1 Introduction

With rich investigation looking into the pattern of English stress assignment through multiple affixations (e.g., Burzio, 1994; Chomsky & Halle, 1968; Marvin, 2013; Shwayder, 2015), one long-lasting puzzle at this morphophonological interface is vowel reduction and its relation to stress shift, i.e., stress reassignment. A famous example regarding this topic is the comparison between *com.p[ə]n.sá.tion* and *con.d[ɛ]n.sá.tion* in *The Sound Pattern of English* (SPE, Chomsky & Halle, 1968) where the unstressed [ɛ] in *cóm.p[ɛ]n.sate* reduces to schwa via nominalization but the stressed [ɛ] in *con.d[ɛ́]nse* does not, as illustrated in (1).

(1) *con.d[ɛ́]nse* → *con.d[ɛ]n.sá.tion* (the stressed vowel does not reduce)
     cóm.p[ɛ]n.sate → *com.p[ə]n.sá.tion* (the unstressed vowel gets reduced)

This case has been taken to suggest that stressed vowels are immune to reduction in multiple morphological operation cycles. That is, the primary stress in previous cycles is preserved as some form of non-primary stress which is reflected as the impossibility for the stress-bearing vowel to reduce. On the other hand, unstressed vowels tend to reduce (see more discussion in Burzio, 1994; Chomsky & Halle, 1968; Kiparsky, 1979; Marvin, 2002). Nevertheless, exceptions are found against the above claims (e.g., Kenyon & Knott, 1953) and two of them are exhibited in (2). Later, Pater (2000) and Burzio (2007) note that those exceptions are more than a few and thus cannot be ignored. They also propose constraints under the Optimality Theory (OT) framework (Prince & Smolensky, 1993) to account for phenomena like stress assignment pattern and vowel reduction in English.

(2) *re.d[ú]ce* → *re.d[ʌ]c.tion* (the stressed vowel gets reduced)
     in.f[ɔ́]rm → *in.f[ɚ].má.tion* (the stressed vowel gets reduced)

While the existing OT constraints are versatile enough to cover numerous kinds of derived words, the particular paradox brought up in (1) and (2) awaits a definitive explanation and it is also worthwhile to see if specific examples like *-ion* and *-ation* nominalizations can offer us more insights into the general relation between vowel reduction and stress assignment. Zooming into this paradox, the two instances in (2) target phonological rules about the stressed vowels instead of the unstressed ones. Therefore, underlyingly, it is the "immunity" claim that needs a revisit. To test if the "immunity" claim holds as a universal constraint, the most direct means is to see whether the counterevidence in (2) is representative among all English cases where *-ion*/*-ation* nominalizations take place[1]. Luckily, the prevalence of corpus data makes it possible for

[1] Apart from the focus on the "immunity" of stressed vowels, another target is when and how unstressed vowels get reduced, as in the case of *cóm.p[ɛ]n.sate* → *com.p[ə]n.sá.tion*, which would be set aside for now for another paper. Besides this, the reason to include nominals ending with both *-ion* and *-ation* is that both are attached to verbs to form nouns and exhibit similar phonological patterns as shown in Section 3. From the morphology perspective, according to *Online Etymology Dictionary*, -ation can be decomposed as *-ate + -ion*, making it allomorphic to *-ion*. I will leave it to another paper to discuss why some verbs go with *-ion* and the others go with *-ation*.

such endeavor and the aim of this paper is to analyze an exhaustive list of nominals ending with *-ion/-ation* and prove that having primary stress actually does not guard the vowel from reduction through nominalization.

In addition, this paper also presents some preliminary theoretical analysis to account for the observed partial dependency of vowel reduction on stress assignment under the OT framework. It adopts Pater (2000)'s ranking hierarchy as a template (see the crucial ranking in (3) and the detailed interpretation in the Appendix), because when Pater comes up with this hierarchy to address the non-uniformity of English secondary stress assignment in various kinds of words, he also hints a sophisticated relation between vowel quality and stress assignment. Even though the sophistication is not explicitly addressed in his paper, the existing hierarchy might already be sufficient to account for the vowel reduction phenomenon related to the *-ion/-ation* nominalization. If so, there would be additional support for the domination of stress assignment on vowel reduction. However, if more constraints other than those in (3) are needed to specifically account for vowel reduction, there would be accumulated evidence that the relation between the two phonological phenomena is just partial. In fact, the latter prediction is what the paper ends up revealing.

(3)   FTBIN, TROCH, NON-FIN ≫ ALIGN-HEAD ≫ *CLASH-HEAD-S_2, ID-STRESS-S_1 ≫ PARSE-σ,
        *OBSNUC ≫ WEIGHT-TO-STRESS ≫ *CLASH-HEAD ≫ *SONNUC, ID-STRESS ≫ ALIGN-L
        where S_1 = {*condensation, apartmental, chimpanzee, …*},
        S_2 = {*admire, companion, Atlanta, Kilimanjaro, representation, …*}

To achieve both the quantitative and analytical goal, this paper is structured in the following way: Section 2 introduces how the empirical data were collected from the dictionary corpus CELEX2 (Baayen et al., 1995) and were prepared for the quantitative analysis. Section 3 presents the analysis and shows how various phonological features, such as vowel tenseness, the adjacency between stressed syllables of the verb and of the nominal, correlate with each other to predict vowel reduction. The quantitative interaction is then translated into ranked constraints in Pater (2000)'s hierarchy, which brings the whole observation closer to a theoretical understanding. For those observations that cannot fit in this hierarchy, further discussion is provided. Section 4 shows how the critical phonological features fit into a logistic regression model to predict vowel reduction, confirms the statistical importance of these features and reinforces the partial dependency of vowel reduction on stress assignment. Finally, Section 5 concludes with the prospect of future research.

2   Data collection

This paper aims to collect an exhaustive list of verb-noun pairs in English for a complete understanding of the interaction between vowel reduction and stress assignment. Therefore, I chose the CELEX2 corpus that is comprised of 53,178 non-overlapping lexical entries documented in *Oxford Advanced Learner's Dictionary* (1974) and *Longman Dictionary of Contemporary English* (1978). The linguistic features include orthographic spelling, word class, phonetic transcriptions, syllable boundaries, and primary stress pattern, etc., which facilitate feature extraction and construction in the following analysis.

After utilizing a self-built regex program in R with a follow-up manual check, I abstracted 1,047 verb-noun pairs where the nouns are nominalizations of the verbs after the suffixation of *-ion* or *-ation*. I further constructed variables such as "stress step" (i.e., the primary stress shifts to right in the nominal by *n* syllables compared with the verb root), vowel tenseness (i.e., whether the stressed vowel in the verb is tense or lax[2]), and vowel change (i.e., whether this stressed vowel exhibits any quality change after nominalization), etc.[3] By classifying the pairs according to their different feature values, I managed to obtain both the qualitative and the quantitative relation between vowel reduction and stress assignment.

---

[2] Vowel tenseness follows the classification criterion in Fromkin et al. (2017) where /ɪ ɛ ʊ ʌ æ ə/ are lax and all the other vowels and diphthongs are tense. The criterion is that lax vowels cannot occur in a coda-less syllable in the word final position (e.g., *[sə]) while tense vowels can (e.g., [si] as *see*).
[3] The corpus data, the data preprocessing code, and a user-friendly analyzable database can all be accessed through the open access platform https://osf.io/52hrv/.

3   Data pattern & Analysis

3.1   *Overview*   Among the 1,047 nominals, 656 take the suffix *-ion* and 391 take the suffix *-ation*. All nominals have their primary stress fall on the penult, corroborating the standard observation in the literature (e.g., Chomsky & Halle, 1968) and the OT hierarchy FtBin, Troch, Non-Fin ≫ Align-Head. 96.5% of the stressed penults are either heavy (i.e., (C)VV, (C)VC) or super heavy syllable (i.e., (C)VNC, (C)VVC), with 37 exceptions (3.5%) where the light nucleus /ɪ/ is stressed (e.g., *de.fi.n[í].tion*, *ad.mo.n[í].tion*).

To achieve the stressed penult in the nominals when this morphological process adds one or two syllables to the verb stem, 857 (81.8%) verbs experience the reassignment of primary while only 190 (18.2%) verbs whose primary stress falls on the ultima, plus *con.tra.di.stín.guish*, *di.stín.guish*, *ex.tín.guish*, have their nominal stress fall the same syllable. Table 1 classifies the verbs into four groups by their primary stress position and exhibits the stress shift patterns in each group. Just because *-ation* has one more syllable than *-ion*, verbs that go with *-ation* naturally "jump one more step to the right".

At first glance, the numerical value of the constructed "stress step" variable seems nothing more than a simple tally. But in fact, depending on the suffix, "stress step" is a proxy of the adjacency between the stressed syllable in the verb (coded as σ́-V) and that in the nominal (coded as σ́-N). With *-ion*, when stress step = 0, σ́-V and σ́-N are the same; when stress step = 1, σ́-V immediately precedes σ́-N; when stress step ≥ 2, there is at least one syllable intervening σ́-V and σ́-N. The interpretation is the same with *-ation* except that it does not make sense to have stress step equal 0 because σ́-N falls on the initial syllable of *-átion* and there is no chance for σ́-V and σ́-N to be the same.

Table 1 Stress shift distribution by stress position

	STRESS STEP: stress shift to the right by *n* syllables					
STRESSED SYLLABLE IN VERB	**0**	**1**	**2**	**3**	**4**	**total (N)**
ultima	<u>187</u>	*168*	*3*			358
penult	<u>3</u>	<u>33</u>	*36*			72
antepenult			<u>428</u>	*161*	*1*	590
preantepenult				<u>5</u>	*22*	27
total (N)	190	201	467	166	23	1047

Note: Nominals ending with *-ion* are underlined; nominals ending with *-ation* are italicized. The dash line demarcates these two groups.

Furthermore, whether there is stress shift classifies the verbs into two groups: those that experience stress shift (annotated as [+shift]) are perfect empirical evidence to study stress preservation and its effect on vowel reduction (annotated as [+reduce]), while those that do not (i.e., [-shift]) offer new insights to whether vowel reduction could be independent from stress shift. If predictions from SPE and relevant claims (e.g., Burzio, 1994; Kiparsky, 1979; Marvin, 2002) were correct, there should be no case with the feature [+shift][+reduce] because of the "immunity" claim. Nor should there be a [-shift][+reduce] condition because the vowel still takes the primary stress in the nominal and thus should not be reduced. However, Table 2 debunks this prediction with a sizable proportion of the critical pairs that are of [+shift][+reduce] (N = 99) and even, more surprisingly, of [-shift][+reduce] (N = 35). This shows the intricacy between vowel reduction and stress shift.

Table 2 Correlation between stress shift & vowel reduction

	STRESS SHIFT		
VOWEL QUALITY	**no shift**	**shift**	**total (N)**
reduction	35 (18.4%)	99 (11.6%)	134
no reduction	155 (81.6%)	758 (88.4%)	913
total (N)	190	857	1047

By combining the information in Table 1 and 2 and considering the potential effect of vowel tenseness on vowel reduction, I classify the pairs into fine-grained groups in Table 3 for case by case interpretation in the following subsections (see the seven exceptions in Section 3.5).

Table 3 Vowel change by suffix type, vowel tenseness, and stress step

		VOWEL REDUCTION		
STRESS STEP	SYLLABLE ADJACENCY	reduction	no reduction	total (N)
-ion		**50**	**603**	**653**
lax vowel in σ-V		**1**	**380**	**381**
0	σ-V = σ-N		93	93
1	σ-V σ-N	1	14	15
2	σ-V σ_1^2 σ-N		272	272
3			1	1
tense vowel in σ-V		**49**	**223**	**272**
0	σ-V = σ-N	35	59	94
1	σ-V σ-N	14	4	18
2	σ-V σ_1^2 σ-N		156	156
3			4	4
-ation		**81**	**307**	**388**
lax vowel in σ-V		**4**	**162**	**166**
1	σ-V σ-N	4	38?	42
2			21	21
3	σ-V σ_1^3 σ-N		83	83
4			20	20
tense vowel in σ-V		**77**	**145**	**221**
1	σ-V σ-N	76	50	126
2			15	15
3	σ-V σ_1^3 σ-N		78	78
4			2	2
total (N)		**130**	**910**	**1040**

3.2 $\sigma\text{-}V\ \sigma_1^n\ \sigma\text{-}N$ By first looking at what features can best predict vowel reduction, I found that when the stressed syllable in the verb (σ́-V) and the stressed one in the nominal (σ́-N) are not adjacent to each other (coded as "σ-V σ_1^2 σ-N" or "σ-V σ_1^3 σ-N" in Table 3 under the "syllable adjacency" tab which is calculated from "stress step"), the stressed vowel do not get reduced. This "long distance" effect is also independent of vowel tenseness and the type of suffix. Examples in this group are in (4).

(4)   σ-V σ_1^n σ-N (1 ≤ n ≤ 3)
        Verbs: *f[ǽ].bri.cate, il.l[ú].mi.nate, p[ɛ́].tri.fy, r[ɛ́].cog.nize, [ɔ́]l.ter, d[ɛ́].stine*
        Nouns: *f[æ].bri.cá.tion, il.l[u].mi.ná.tion, p[ɛ].tri.fác.tion, r[ɛ].cog.ní.tion, [ɔ]l.te.rá.tion,*
            d[ɛ].sti.ná.tion

This pattern can be explained by the faithfulness constraint IDENT-STRESS that says stressed syllables had better retain the stress. Even though the constraint *CLASH-HEAD dominates IDENT-STRESS in Pater (2000)'s hierarchy, the medial syllable(s) in between σ́-V and σ́-N in this condition prevents the nominal from violating the markedness constraint. If we assume that the preserved stress prevents the vowel from reduction with the goal to keep the syllable's prominence, the retained vowel is a natural consequence of *CLASH-HEAD ≫ IDENT-STRESS and this group of words also aligns with the classic SPE claim.

3.3 *σ-V σ-N*   When the medial intervening syllable disappears, the patterns become less clear. First, it seems that tense vowels are more subject to reduction than lax vowels shown by Table 4 (Fisher's Exact Test for *-ion*: odds ratio = 0.024, p < .001; for *-ation*: odds ratio = 0.070, p < .001).

Table 4 Vowel reduction pattern of σ-V σ-N

| SYLLABLE ADJACENCY = σ-V σ-N | | VOWEL REDUCTION | | |
STRESS STEP = 1		reduction	no reduction	total (N)
-ion	lax vowel in σ-V	1	14	15
	tense vowel in σ-V	14	4	18
-ation	lax vowel in σ-V	4	38	42
	tense vowel in σ-V	76	50	126
total (N)		**95**	**106**	**201**

Second, the cases that are the most compatible with the analysis in Section 3.2 are when tense vowels get reduced (N(*-ion*) = 14, N(*-ation*) = 76) because this phenomenon can be explained by the constraint rank *CLASH-HEAD ≫ IDENT-STRESS. More specifically, in order to avoid two adjacent syllables from having two prominent stresses (see a similar constraint "avoid medial clash" in Halle and Vergnaud (1987) and Burzio (2007)), *CLASH-HEAD requires the syllable that used to bear the primary stress in the verb form to destress in the nominal. Just because this *CLASH-HEAD dominates IDENT-STRESS, the adjacent stresses would be penalized even though this would be against the subsequent faithfulness constraint.

While the majority of tense vowels are reduced (90 out of 144, 62.5%), a sizable proportion of exceptions with tense vowels do not (N(*-ion*) = 4, N(*-ation*) = 50). These exceptions are most likely to violate *CLASH-HEAD since the retained tense vowel probably indicates some form of stress preservation that could trigger the stress clash. I've explored whether some hidden phonological features can explain this puzzle here, but all the attempts have failed except that front vowels in the *-ation* condition unanimously get reduced. Before I find the definitive answer, it seems that the most plausible analysis for now is that those that do not exhibit vowel reduction in the tense vowel condition (N = 4 + 50 = 54) and violate *CLASH-HEAD actually belong to the S_1 set where the constraint ID-STRESS-S_1 dominates *CLASH-HEAD. It is the preservation of stress, which is mandated by ID-STRESS-S_1, that leads to the retainment of vowel quality (see examples in (5)).

(5)   Verbs: *de.h[ái].drate, str[ái].ate, c[ɔ́]se, re.f[ú]te* …
            Nouns: *de.h[ai].drá.tion, str[ai].á.tion, c[ɔ].sá.tion, re.f[u].tá.tion* …

Turning to the lax vowel condition, most vowels in the verb do not get reduced in the nominal (N = 14 for *-ion*, N = 38 for *-ation*). If the retainment of vowel quality also suggests some sort of stress preservation, *CLASH-HEAD could be violated again. To address the retainment of lax vowel in this condition, it might help to propose a new faithfulness constraint ID-V[+LAX] that says lax vowels should remain as the same qualitatively throughout morphological transformations. Subsequently, this constraint could dominate *CLASH-HEAD. Despite this attempt, there is an alternative explanation for the unreduced cases: due to the limitation of CELEX2, usually only one version of transcription is provided for each word and variations are not documented completely. This leads to more cases in the unreduced condition and fewer cases in the reduced condition than what it would be otherwise given the actual variation data. For instance, *adaptation* is only transcribed as *a.d[æ]p.tá.tion* in CELEX2 while the *Oxford English Dictionary* provides both

/ˌædæpˈteɪʃ(ə)n/ and /ˌædəpˈteɪʃ(ə)n/ (Adaptation, 1989) where the second version has the once stressed [æ] reduce to a schwa. Because of this lack of variations in CELEX2, it is difficult to pin down the cases whose once stressed lax vowels genuinely remain the same throughout the nominalization. This issue remains one of the limitations of this research. Yet, no matter what percentage of lax vowels that do not exhibit reduction when they immediately precede the stressed syllable in the noun, they could all belong to the S_1 set whose faithfulness constraint ID-STRESS-S_1 dominates *CLASH-HEAD—the word *con.d[ɛ́]nse* is a well-known member. For those that do experience reduction, the majority of lax vowels reduce to schwa with only two or three exceptions (see Zhang (2020) for a summary of the vowel reduction patterns).

3.4   *σ-V = σ-N*   The last group consists of verbs where the primary stress fall on the ultima in the verb of the *-ion* condition and this stress-bearing syllable does not change in their nominals. From a quantity sensitivity perspective, the syllable should retain its strength and thus should require the vowel not to reduce (WEIGHT-TO-STRESS). Surprisingly, the subgroup where tensed vowels still get reduced (N = 35 with the famous example of *reduce*) directly goes against the above prediction. This is also the most salient evidence that vowel reduction is only partially dependent on stress assignment.

Table 5 Vowel reduction pattern of σ-V = σ-N

SYLLABLE ADJACENCY = σ-V = σ-N		VOWEL REDUCTION		
STRESS STEP = 0		reduction	no reduction	total (N)
-ion	lax vowel in σ-V		93	93
	tense vowel in σ-V	35	59	94
total (N)		**35**	**152**	**187**

   In the analysis, I found one factor that is highly correlated with vowel reduction. That is the coda type of σ́-V. As shown in Table 6, it seems clear that once the stressed syllable ends with the coda /t/, the vowel is immune to reduction (the exception of coda /b/ is *ab.s[ɔ́]b → ab.s[ɔ́]p.tion*). This coda /t/ matches an underlying lexical pattern—verbs that end with *-ate* and have a Latin origin with examples like *e.qu[é]te*, *mi.gr[é]te*, and *vi.br[é]te*. While these cases can be captured by Pater (2000)'s existing hierarchy (e.g., ID-STRESS), there seems to be additional constraints at play that specifically characterize the lexical exceptionality of *-ate*. On the other hand, for the 35 reduced cases, while none of their reduced form is schwa (Zhang, 2020), there has not been sufficient account to the best of my knowledge that addresses motivation behind the reduction—there must be additional constraints that override the influence of stress and the quantity-sensitivity requirement and directly drive the reduction of tense vowels. While these hypotheses point to exciting directions to extend the current OT, I will leave the exploration for future.

Table 6 Vowel reduction based on coda type in the stressed syllable

CODA	voiced plosive		nasal		fricative			ending with *t*				no coda	total (N)
	b	**d**	**m**	**n**	**s**	**v**	**z**	**t**	**st**	**pt**	**kt**		
reduction	7	1	5	7	6	4	1	2				2	35
no reduction	1							55	1	1	1		59

3.5   *The 7 exceptions*   Due to the rigidity of the preprocessing algorithm, 7 "exceptions" were left out from Table 3. These are three cases where the nominals are formed by deleting the ultimate syllable in the verb and adjoining the suffix *-ion* (as Group 1 in Table 7) and four cases where the final syllable in the verb forms a light syllable with the nucleus [ɪ] before adjoining the suffix *-ation* (as Group 2 in Table 6). In Group 1, the preservation of the primary stress drives no reduction, given the underlying assumption that the stressed tense vowel [i] needs to retain its quality to keep the syllable weight (WEIGHT-TO-STRESS). In Group 2, while the pattern is underlyingly "σ-V σ σ-N", the stressed syllable in the verb still gets reduced, which is against the observation in Section 3.2 but similar to the 35 reduced cases in Section 3.4. Both the words in Group 2 and the 35 cases pose challenges to Pater (2000)'s hierarchy and implie that the OT system that aims to capture the secondary stress distribution is probably not sufficient to explain the vowel reduction. It becomes even clearer that constraints independent from stress are needed for a complete account of vowel reduction.

Table 7 "Exception" words

GROUP 1		GROUP 2	
Verb	**Noun**	**Verb**	**Noun**
con.tra.di.st[í]n.guish	*con.tra.di.st[í]nc.tion*	*mis.pro.n[áu]nce*	*mis.pro.n[ʌ]n.ci.á.tion*
di.st[í]n.guish	*di.st[í]nc.tion*	*pro.n[áu]nce*	*pro.n[ʌ]n.ci.á.tion*
ex.t[í]n.guish	*ex.t[í]nc.tion*	*re.n[áu]nce*	*re.n[ʌ]n.ci.á.tion*
		re.con.c[ái]le	*re.con.c[ɪ].li.á.tion*

3.6   *Interim summary*   In this section, I show how verb-noun pairs can be classified into subgroups based on linguistic features such as syllable adjacency and vowel tenseness, and how the features interact with each other to predict vowel reduction. I conclude that the observation largely corroborates with what the OT constraint hierarchy in (3) predicts, but the 35 reduced vowels under primary stress (in Table 5), the peculiarity of Latinate verbs (in Table 6), and the exceptions in Group 2 (in Table 7) indicate that more independent and lexically specified constraints are needed. Crucially, vowel reduction is only partially dependent on stress assignment.

4   Statistical modeling

The above analysis is based on a tally of verb-noun pairs that exhibit different features including but not limited to stress reassignment, vowel tenseness, and vowel reduction. Since those parameters were selected via my visual inspection, chance is that there might be other factors at play that I fail to review. Besides, even if the existing features are all that matter with respect to predicting vowel reduction, the analysis still lacks evidence to confirm their statistical significance. In order to ensure that the essential features discussed in the previous section are truly significant and to uncover potential hidden factors, I fit the 1,047 samples into a logistic regression model in R with the simplest syntax as "logit(vowel reduction) ~ suffix type + stress step + coda /t/ in σ́-V + vowel tenseness + vowel frontness + vowel height + verb syllable count"[4].

The result reveals that only the first four parameters are significant (see Table 8 for the detailed statistics) with the model achieving an overall prediction accuracy of 95.8%. The interpretation goes as (1) when the suffix type is -ation or there are exceptions like *pronunciation*, the vowel in σ́-V is more likely to be reduced; (2) when the stress step is larger, i.e., σ́-V and σ́-N are more distant from each other, the target vowel is more likely to stay the same; (3) when the σ́-V ends with the coda /t/, the target vowel is more likely to stay the same; (4) when the vowel in σ́-V is tense, it is more likely to be reduced.

While findings of (2) – (4) are consistent with the analysis in Section 3, the first one related to the suffix type does provide more information that I have not focused on and this is exactly where the value of exploratory statistical analysis lies. This additional feature indicates that from a morphological perspective,

---

[4] I should have looked into more interaction terms between these predictions but due to the limited sample size and the good-enough prediction accuracy of the simple model, I assume that having more interaction terms in the model does not alter the nature of the most crucial finding. After all, the goal of the statistical model is not to achieve its highest predictive power because there are always exceptions, but to detect the significance of the various predictors.

the suffix type influences the probability of vowel reduction and at the morphophonological interface, it could be the case that adding the two syllables of *-ation* rather than only one syllable of *-ion* is more likely to drive vowel shortening in order to restrain the word duration within a reasonable length. Altogether, these four important features can well predict vowel reduction after this nominalization process and the different predictors here again reemphasize that stress assignment is only one of the impacting factors: the suffix type alludes to some underlying requirement of word duration; the coda /t/ constraint provides the motivation to consider coda features and lexical specificity (i.e., the Latinate words); the vowel tenseness indicates the endogenous factor of vowel quality.

Table 8 Statistics of the logistic regression model

PARAMETERS	Estimate	Std. Error	z value	p value
(Intercept)	-20.28	1414.89	-0.014	0.989
Suffix type (numerical)	3.59	0.59	6.045	< .001
Stress step (numerical)	-3.87	0.50	-7.793	< .001
Coda /t/ in σ́-V (= true)	-3.96	0.48	-8.252	< .001
Vowel tenseness (= tense)	3.05	0.63	4.880	< .001

Note: As the response variable, "vowel reduction" is coded as 1, while the alternative is 0.

5   Conclusion

To sum up, this paper presents the first corpus-based quantitative case study that looks at the relation between vowel reduction and stress shift in verb-noun pairs formed via the English *-ion* nominalization. The discussion of this intricate relation starts from *The Sound Pattern of English* (Chomsky & Halle, 1968) with the famous comparison between *com.p[ə]n.s[é].tion* and *con.d[ɛ]n.s[é].tion*. While the classic theory is that stressed vowels are immune to reduction in subsequent cyclic affixations, inspired by the exceptions against this claim (e.g., Burzio, 2007; Pater, 2000), this study provides concrete examples that suggest that vowel reduction only partially depends on the stress assignment or reassignment. This is also supported by the analysis that the existing OT framework that captures the non-uniformity of English second stress (Pater, 2000), while being able to explain quite a lot of the phenomena about vowel reduction in this paper, is still not sufficient to cover all the nuanced cases. Furthermore, the classic claim is also opposed to by the exploratory statistical analysis that features like the suffix subtype and the vowel tenseness are at play.

Even though the quantitative data provides rich soil for the current investigation, this is still preliminary work before we getting a complete picture of the relation between vowel reduction and stress shift. First, we need have a complete summary of what kinds of phonological constraints affect vowel reduction and to what extent. Then, the attempt to achieve a satisfactory OT analysis should still be continued. Second, empirical data that provide more variations of words are highly needed since the dictionary entries provided by CELEX2 do not completely reflect the language profile of native speakers of American English, which is also a common limitation of using dictionary data (Stanton, 2019). Last but not least, since the Latinate words hint a layer of lexical specificity, it might also help to incorporate a diachronic perspective to see how words evolve with certain characteristics that relate to this vowel reduction phenomenon.

6   Appendix

The detailed interpretation of the OT hierarchy in Pater (2000) (ordered by column):

FTBIN (FootBinarity):
Feet must be minimally bimoraic.

PARSE-σ:
All syllables belong to feet.

TROCH:
Feet are trochaic.

*OBSNUC ≫ *SONNUC:
Obstruent nuclei are dispreferred to sonorant nuclei.

NON-FIN (Non-Finality):
The head of the Prosodic Word must not be final.

WEIGHT-TO-STRESS:

ALIGN-HEAD (Align (PrWd-R, Head(PrWd)-R)):
Align the right edge of the Prosodic Word with
the right edge of the head of the Prosodic Word.

*CLASH-HEAD-S2:
Pretonic stresslessness is preferred, even though
this results in an extra Parse-σ violation, for
words in S2, where S2 = {admire, companion,
Atlanta, Kilimanjaro, representation, …}

ID-STRESS-S1:
ID-STRESS constraint that only apply to words in
S1, where S1 = {condensation, apartmental,
chimpanzee, …}

Heavy syllables be placed in head position of a
foot.

*CLASH-HEAD:
No stressed syllables may be adjacent to the head
syllable of the Prosodic Word.

ID-STRESS:
If α is stressed, then $f(α)$ must be stressed, where
f is the correspondence relation between input
and output strings of segments.

ALIGN-L:
Align all feet with the left edge of the prosodic
word.

References

Adaptation. (1989). In *OED Online*. Oxford University Press. Retrieved January 29, 2021, from http://www.oed.com/view/Entry/2115

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *CELEX2 LDC96L14*. Philadelphia: Linguistic Data Consortium. https://catalog.ldc.upenn.edu/LDC96L14

Burzio, Luigi. (1994). *Principles of English stress*. Cambridge University Press.

Burzio, Luigi. (2007). Phonology and phonetics of English stress and vowel reduction. *Language Sciences*, *29*(2–3), 154–176. https://doi.org/10.1016/j.langsci.2006.12.019

Chomsky, Noam, & Halle, Morris. (1968). *The sound pattern of English*. Harper & Row.

Fromkin, Victoria, Rodman, Robert, & Hyams, Nina M. (2017). *An introduction to language* (11th edition). Cengage.

Halle, Morris, & Vergnaud, Jean-Roger. (1987). Stress and the cycle. *Linguistic Inquiry*, *18*(1), 45–84.

Kenyon, Johnson S., & Knott, Thomas A. (1953). *A pronouncing dictionary of American English* (4th edition). Merriam.

Kiparsky, Paul. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, *10*(3), 421–441.

Marvin, Tatjana. (2002). *Topics in the stress and syntax of words*. PhD thesis. Massachusetts Institute of Technology.

Marvin, Tatjana. (2013). Is word structure relevant for stress assignment? In Ora Matushansky, Alec Marantz, & Morris Halle (Eds.), *Distributed morphology today: Morphemes for Morris Halle* (pp. 79–93). The MIT Press.

Pater, Joe. (2000). Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology*, *17*(2), 237–274. http://www.jstor.org/stable/4420170

Prince, Alan, & Smolensky, Paul. (1993). *Optimality theory: Constraint interaction in generative grammar*. Wiley-Blackwell.

Shwayder, Kobey. (2015). *Words and subwords: Phonology in a piece-based syntactic morphology*. PhD thesis. University of Pennsylvania. https://repository.upenn.edu/cgi/viewcontent.cgi?article=2946&context=edissertations

Stanton, Juliet. (2019). Stress in–*ative* and the role of pre-at-segments: Evidence from dictionary studies. In Richard Stockwell et al. (Eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics* (pp. 326-332). Cascadilla Proceedings Project.

Zhang, Yuhan. (2020). *Stress shift is proportional and vowel reduction is not deterministic—A corpus case study of English -tion nominalization* [Poster]. The 94th Annual Meeting of the Linguistic Society of America, New Orleans, LA. https://scholar.harvard.edu/files/yuhanzhang/files/lsa2020-stress-vowel.pdf