to appear in *Natural Language & Linguistic Theory*

Prosodic identity in copy epenthesis: evidence for a correspondence-based approach*

Juliet Stanton[1] and Sam Zukoff[2]
[1]juliet.stanton@alum.mit.edu [2] MIT, szukoff@mit.edu

ABSTRACT. This paper focuses on languages that exhibit processes of *copy epenthesis*, specifically those where the similarity between a copy vowel and its host extends to prosodic or suprasegmental resemblance. We argue that copy vowels and their hosts strive for identity in all prosodic properties, and show that this drive for prosodic identity can cause misapplication in the assignment of properties such as stress, pitch, and length. To explain these effects, we argue that any successful analysis of copy epenthesis must involve a correspondence relation (following Kitto and de Lacy 1999). Our proposal successfully predicts the extant typology of prosodic identity effects in copy epenthesis; alternative analyses of copy epenthesis relying solely on featural spreading (e.g. Kawahara 2007) or gestural realignment (e.g. Hall 2003, 2006) do not naturally capture the effects discussed here.

KEYWORDS: copy epenthesis, phonology, correspondence, misapplication, prosody

1 Introduction

The term *copy epenthesis* describes a class of patterns in which the quality of an epenthetic vowel depends on the quality of one of its vocalic neighbors. For example: in a language where underlying /pri/ is realized as [pi̲ri] but /pra/ as [pa̲ra] (not *[pi̲ra]), the vowel that appears in the unexpected position (the *copy*) is featurally identical to the vowel that appears in the expected position (its *host*).

How the dependence between a copy vowel and its host should be formalized is a matter of debate. Traditional autosegmental analyses treat copy epenthesis as the result of feature spreading: the epenthetic vowel, lacking features of its own, obtains them through autosegmental association with a nearby host (e.g., Clements 1986, 1991; Gafos and Lombardi 1999; Halle et al. 2000; Shademan 2002; Kawahara 2007). Analyses in the framework of Articulatory Phonology (Browman and Goldstein 1986) claim that (at least some) cases of copy epenthesis result from gestural realignment: in the mapping from underlying /pra/ to surface [pa̲ra], the constellation of gestures that yield an [r] migrate to the middle of the nucleus, allowing underlying /a/ to be heard on both sides (e.g., Steriade 1990; Hall 2003, 2006). Others propose extensions of Correspondence Theory (McCarthy and Prince 1995) in which a copy and its host stand in correspondence with one another; faithfulness constraints render the corresponding pair identical (Kitto and de Lacy 1999, Yu 2005, Kim 2008).

*Authors' names are in alphabetical order. We are grateful to Adam Albright, Donca Steriade, Eric Bakovic, Edward Flemming, Gunnar Ólafur Hansson, Bruce Hayes, Junko Ito, Michael Kenstowicz, Ezer Rasin, Nina Topintzi, Eva Zimmerman, and audiences at MIT, CLS 51, and 24mfm for helpful discussion. Earlier portions of this research have been published as Stanton and Zukoff (2016); comments from three *NLLT* reviewers (Paul de Lacy, Shigeto Kawahara, and one anonymous reviewer) and the associate editor (Rachel Walker) have helped shape the paper into its present form. All remaining mistakes are ours.

The predictions of these analyses diverge when we consider the potential for prosodic identity (or, agreement for suprasegmental features such as stress and length) between a copy vowel and its host (see Kitto and de Lacy 1999:17, also Kawahara 2007:16–17). Under a spreading-based theory in particular, there is no reason to predict that a copy vowel and its host should exhibit prosodic identity, as properties such as stress and length are commonly assumed not to undergo autosegmental spreading. And under a theory of copy epenthesis as gestural intrusion, there is no *a priori* reason to expect that the two halves of the underlying vowel should seek to resemble one another prosodically (see Sections 3.4 and 4.6 for discussion). Under a correspondence-based theory, however, prosodic identity effects are predicted: faithfulness constraints can refer to prosodic properties such as stress and length on a number of different correspondence dimensions (see e.g. McCarthy and Prince 1988 on length identity in reduplication), so we would expect to find effects of these constraints on the copy-host dimension as well. Thus the (non-)existence of prosodic identity effects has the potential to inform our understanding of the nature of the dependence between a copy vowel and its host.

In this paper, we analyze a number of cases in which the similarity between copy vowels and their hosts extends beyond featural identity to prosodic identity, and show that the data are best-captured under an analysis in which copy vowels and their hosts stand in correspondence with one another. Although the claim that copy-host identity extends to prosodic properties is not novel, and the particular languages discussed here have been previously identified as significant for correspondence-based theories of copy epenthesis (see Kitto and de Lacy 1999), the main contribution of this paper is its detailed explication and analysis of a larger set of prosodic misapplication effects than has previously been acknowledged. We show, for example, that, in Scottish Gaelic, copy vowels and their hosts must be durationally equivalent; when this requirement cannot be satisfied, copy epenthesis is *blocked* (Sections 3.2–3.4). To give another example, we show that the unusually short durations of both copy and host vowels in Ho-Chunk can be explained if epenthetic vowels in Ho-Chunk must be short, and the host must shorten to match the copy (Section 4.4); this is then an instance of *back-copying* in copy epenthesis (cf. Kitto and de Lacy 1999:16).

We conclude by showing that a correspondence-based theory of copy epenthesis is capable of predicting the full attested typology of prosodic misapplication effects, a result that spreading-based and gestural-intrusion analyses of copy epenthesis cannot match. While our proposal is not yet sufficient to account for the full range of generalizations that hold over the typology of copy epenthesis more generally (see Section 5.1 for discussion, as well as Kawahara 2007), we believe that the current proposal's ability to account for prosodic misapplication effects provides a solid argument that any successful analysis of copy epenthesis on the whole must involve a correspondence relation.

1.1 Representational Assumptions

Before proceeding to the analyses in this paper, it is necessary to lay out some of our assumptions regarding the phonological representation of copy epenthesis and the constraints that derive it. Following Kitto and de Lacy (1999), we assume that copy vowels and their hosts stand in correspondence with one another; we refer to this as the Host-Epenthetic (HE) correspondence relation.[1] Beyond this, the only correspondence relation that we assume to hold over this domain is the one between the surface host vowel and its underlying correspondent. The representational structure we adopt is thus the one schematized in (1). Whether or not the analysis of copy epenthesis benefits

[1] We use "Host-Epenthetic" throughout the paper instead of Kitto and de Lacy's "Base-Epenthetic" so as to avoid confusion with morphological bases.

from some additional correspondence relation *z*, holding between the host's underlying correspondent and its copy (as proposed in Stanton and Zukoff 2016), is a question we leave for future work.

(1) Representation of copy epenthesis

/ **ɔ**$_{x}$ r m /

IO-Correspondence

[**ɔ**$_{x,y}$ r **ɔ**$_{y}$ m]

HE (Host-Epenthetic) Correspondence

Cases of copy epenthesis can be differentiated from cases of *default* epenthesis, where the quality of the inserted vowel is not dependent on the quality of the neighboring vowels (in Tongan, for example, 'camel' is adapted as [kameli̱] and 'colonel' as [keːnali̱]; Kitto 1997). Following Kitto and de Lacy (1999), we assume that default epenthetic vowels, just like copy vowels, stand in correspondence with a host. The difference between copy and default epenthesis amounts to a difference in the ranking of a collection of faithfulness constraints, which require an epenthetic vowel to be identical to one of its vocalic neighbors (abbreviated below as HE-IDENT[F]); and another collection of faithfulness constraints, which prefer the insertion of a specific default vowel (given a language where [i] is the preferred default vowel, this collection is abbreviated as ∅→[i]).[2]

(2) HE-IDENT[F]: Assign one violation mark for each pair of vowels standing in HE correspondence that do not have identical values for [F].

(3) ∅→[i]: Assign one violation mark if any non-[i] output vowel lacks an input correspondent.

In cases of copy epenthesis, HE-IDENT[F] dominates ∅→[i] (4); in cases of default epenthesis, ∅→[i] dominates HE-IDENT[F] (5). For epenthesis to be motivated, it must be the case that an additional phonotactic constraint dominates either HE-IDENT[F] or ∅→[i]; below, we demonstrate this with the constraint NOCODA ('assign one violation mark for each coda consonant').

(4) Copy epenthesis: HE-IDENT[F] ≫ ∅→[i]

/a$_{x}$t/	NOCODA	HE-IDENT[F]	∅→[i]
a. [a$_{x}$t]	*!		
b. [a$_{x,y}$ti̱$_{y}$]		*!	
☞ c. [a$_{x,y}$ta̱$_{y}$]			*

(5) Default epenthesis: ∅→[i] ≫ HE-IDENT[F]

/a$_{x}$t/	NOCODA	∅→[i]	HE-IDENT[F]
a. [a$_{x}$t]	*!		
☞ b. [a$_{x,y}$ti̱$_{y}$]			*
c. [a$_{x,y}$ta̱$_{y}$]		*!	

The analyses that follow assume that copy epenthesis occurs when HE-IDENT[F] dominates the markedness constraints that prefer default epenthesis; the analytic differences between copy and default epenthesis are not relevant to the topics at issue here.

[2]We assume that HE-IDENT[F] stands for a collection of faithfulness constraints that require identity for individual feature values (i.e. HE-IDENT[±back], etc.). With Kitto & de Lacy (1999:5), we assume that cases of partial copy can be derived when HE-IDENT constraints are interleaved with markedness constraints that penalize certain feature values.

Before moving forward, we would like to acknowledge several implementational issues that we do not address here. The first is the issue of directionality: how does a copy vowel choose whether to resemble the vowel to its left (as in Selayarese, Section 2) or the vowel to its right (as in Scottish Gaelic and Ho-Chunk, Section 3–4)? We follow Kitto and de Lacy (2009:9) in assuming that directionality is regulated by two competing constraints: COPY-LEFT and COPY-RIGHT. In the analyses that follow, we simply assume the correct directionality and omit reference to these constraints. The second implementational issue concerns the rationale for correspondence between an epenthetic vowel and its host. We do not explore the question of *why* epenthetic vowels stand in correspondence with their hosts; we only assume that they must. For the purposes of this paper, we assume that the requirement that an epenthetic vowel and its host stand in correspondence is a constraint on GEN: candidates in which this correspondence relation does not hold are impossible candidates. (But for some evidence that it might be advantageous to treat the pressure to correspond as a violable constraint, see Kitto and de Lacy 1999:11–12, as well as pp. 36–37 of the present paper.) Our focus here is only on the evidence that positing a correspondence relationship is necessary.

1.2 Details and outline of the study

The empirical phenomena discussed in this paper were identified from a survey of 51 languages with copy epenthesis, drawn primarily from Kawahara's (2007) survey and other online resources. In conducting the survey, we sought to find cases of prosodic misapplication in copy epenthesis that could not be straightforwardly analyzed by appealing to a ban on stressing epenthetic vowels (e.g. McCarthy and Prince 1994 on Makassarese) or faithfulness to the source stress, in the case of loanwords (e.g. Kenstowicz 2007 on Fijian). We term these kinds of cases *complex misapplication effects*. The survey yielded three languages that display such effects: Selayarese (discussed in Section 2), Scottish Gaelic (discussed in Section 3), and Ho-Chunk (discussed in Section 4). For a list of languages consulted in the survey, see the Appendix.

2 Selayarese

In Selayarese (Malayo-Polynesian), stress misapplies when a copy vowel is present at certain positions within the word (see Mithun and Basri 1986, Piggott 1995, Kitto and de Lacy 1999, Broselow 2008, *a.o.*). We show that these data cannot be straightforwardly explained under an analysis where epenthetic vowels are invisible to stress assignment, but that they are easily explained under an analysis where copy-host pairs strive for stress identity. (We note that the proposed analysis closely resembles Kitto and de Lacy's 1999 analysis; see the end of Section 2.2 for discussion.)

2.1 The basic stress pattern

In Selayarese, stress normally falls on the penultimate syllable (Mithun and Basri 1986, *a.o.*). As shown in (6), all stressed syllables are heavy. If the stressed syllable is open, the vowel lengthens (a–d); if the stressed syllable is closed (by a geminate or a glottal stop), the vowel is short (e–f).

(6) Penultimate stress in Selayarese (Mithun and Basri 1986:212–219)

a. [sikúːra] 'how many?'
b. [sampúːlo] 'ten'
c. [kasúːᵐba] 'dye for coloring clothes or cake'
d. [káːsi] 'white cloth'
e. [kásːi] 'sour'
f. [séʔla] 'salt'

In what follows we assume that all stressed syllables must be heavy (i.e. we do not consider candidates in which the stressed syllable is light) and focus only on the placement of stress. We model the pattern in (6) with the constraints *LAPSERIGHT (abbreviated *LAPSER, (7a)) and NONFINALITY (7b). Together, these two constraints require that stress fall on the penult (8). (Throughout, we follow Prince 1983, Gordon 2002, and others in analyzing the distribution of stress without making reference to metrical constituency. The question of whether or not reference to metrical constituency is necessary is tangential to the topics at issue in this paper.)

(7) Constraints to generate penultimate stress (Gordon 2002)

a. *LAPSER: Assign one violation mark * if neither of the final two syllables is stressed.
b. NONFINALITY: Assign one violation mark * if the final syllable is stressed.

(8) Deriving penultimate stress

/sikura/	*LAPSER	NONFINALITY
a. [síːkura]	*!	
☞ b. [sikúːra]		
c. [sikuráː]		*!

2.2 Copy epenthesis and stress misapplication

There is a class of words in Selayarese that unexpectedly bear antepenultimate stress (9). All of these words share several properties (Mithun and Basri 1986:237): the last two vowels are identical; these vowels are separated by a coronal continuant (/s/, /r/, or /l/); and the final vowel present in the isolation form is absent when a vowel-initial suffix is added: for example, [láːᵐber**e**] 'long' → [laᵐbéːr-aŋ] 'longer' (Mithun and Basri 1986:238). This is significant because deletion under hiatus is not a normal phonological process in Selayarese: word-final vowels are permitted to surface before vowel-initial suffixes more generally – for example, [tínro] 'sleep' → [pa-tinró-aŋ] 'bed, bedroom' (Mithun and Basri 1986:238). These observations suggest that final vowels are epenthetic, and that their status as epenthetic vowels influences stress placement.[3]

(9) Examples of antepenultimate stress in Selayarese (Mithun and Basri 1986:237)

a. [sóːᵐbala] 'a sail' e. [kíːkiri] 'metal file'
b. [káːtala] 'itch' f. [hálːasa] 'suffer'
c. [bóːtolo] 'bottle' g. [láːᵐbere] 'long'
d. [béːrasa] 'rice' h. [túːlisi] 'write'

[3] Similarly, in Tahitian, there is exceptional antepenultimate stress when the final two vowels are identical and separated by [ʔ] (Bickmore 1995, Kitto and de Lacy 1999). However, there is no direct evidence that either of the vowels is epenthetic.

One possible analysis of (9) is that epenthetic vowels in Selayarese are invisible for the purposes of stress assignment (Mithun and Basri 1986): these forms have antepenultimate stress because the antepenult is the second non-epenthetic vowel, counting from the word's right edge. When more data is taken into consideration, however, it becomes clear that epenthetic vowels are not always invisible to stress. In (10–11), stress lands on the penultimate syllable despite the presence of an epenthetic vowel (underlined) in the penult (10), or in the antepenultimate and final syllables (11).

(10) Internal epenthesis in loanwords (Broselow 2008:3)
  a. [kar<u>áː</u>tu] 'card' not *[káːr<u>a</u>tu] (source: Indonesian [kártu])
  b. [sur<u>úː</u>ga] 'sugar' not *[súːr<u>u</u>ga] (source: Indonesian [súrga])

(11) Internal and external epenthesis in loanwords (Broselow 2008:4)
  a. [sol<u>o</u>déːr<u>e</u>] 'weld' not *[sóːlodere] (source: Indonesian [sólder])
  b. [kar<u>a</u>tíːs<u>i</u>] 'ticket' not *[káːratisi] (source: Indonesian [kárcis])

An analysis in which epenthetic vowels are invisible to stress assignment would predict a different outcome. The forms in (10) would have antepenultimate stress (*[káːr<u>a</u>tu]), because the penultimate vowel would be ignored. The forms in (11) would have pre-antepenultimate stress (*[káːr<u>a</u>tis<u>i</u>]), as both the antepenultimate and final vowels would be ignored. An analysis appealing to the invisibility of epenthetic vowels thus does not extend to the full set of data in any straightforward way.

We show now that positing a correspondence relation between a copy vowel and its host allows us to analyze this case of conditional invisibility. We assume, as outlined above, that an epenthetic vowel and its host stand in correspondence with one another; the constraint HE-IDENT[stress] (12) requires them to agree in stress. (For the purposes of this paper, we tacitly follow Prince 1983, Gordon 2002, *a.o.* in assuming a grid-based representation of stress. In (12), [stressed] refers to a vowel associated with at least two grid marks; [stressless] refers to a vowel associated with one.)

(12) HE-IDENT[stress]: Assign one violation mark * for each pair of vowels standing in HE correspondence that do not have identical values for stress (i.e. [stressed]↔[stressless] or [stressless]↔[stressed]).

To derive misapplication, we require two crucial rankings: (i) HE-IDENT[stress] must dominate *LAPSER; and (ii) at least one of NONFINALITY and *CLASH (13) must also dominate *LAPSER.[4]

(13) *CLASH: Assign one violation mark * for each sequence of two adjacent stressed syllables.

Recall that in (14) and the tableaux that follow, we analyze only the assignment of stress; the candidate set is limited to those candidates where copy epenthesis occurs.

(14) Misapplication of stress to satisfy HE-IDENT[stress]: [túːlis<u>i</u>] (9h)

/tulis/	NONFINALITY	*CLASH	HE-IDENT[stress]	*LAPSER
☞ a. [túːlis<u>i</u>]				*
b. [tulíːs<u>i</u>]			*!	
c. [tulis<u>íː</u>]	*!		*	
d. [tulíːs<u>íː</u>]	*!	*!		

[4]See e.g. Kager 1994 on *CLASH in Optimality Theory. It is also possible to replace *CLASH with a constraint that bans the presence of multiple stresses within a word (see e.g. Prince 1983 on culminativity). For the purposes of this analysis, the two options are equivalent.

In optimal candidate (14a), copy-host identity for stress is achieved by retracting stress to the antepenult, contrary to the default penultimate pattern. All other options are ruled out by high-ranked constraints. The default candidate with a single penultimate stress (14b) violates HE-IDENT[stress], because the host (the penult, stressed) and the copy (the final, unstressed) disagree in stress. A candidate with a single stress on the final (14c) violates both NONFINALITY and HE-IDENT[stress], while one that stresses both the penult and the final (14d) violates NONFINALITY and *CLASH.

Consider now the forms in (10), where a copy-host pair occupies the antepenult and the penult (e.g. [kará:tu]). While stressing the final syllable would result in satisfaction of HE-IDENT[stress], this candidate (15a) violates NONFINALITY; the fact that final stress is not the attested outcome indicates that NONFINALITY ≫ HE-IDENT[stress]. In this situation, stress defaults to the penultimate syllable, as stressing both the antepenult and the penult (15b) would result in a fatal violation of *CLASH (which thus must dominate HE-IDENT[stress]), and stressing the antepenult only (15d) would result in a fatal violation of low-ranked *LAPSER.

(15) Default stress with medial epenthesis: [kará:tu] (10a)[5]

/kartu/	NONFINALITY	*CLASH	HE-IDENT[stress]	*LAPSER
a. [karatú:]	*!			
b. [ká:rá:tu]		*!		
☞ c. [kará:tu]			*	
d. [ká:ratu]			*	*!

This same interaction yields default penultimate stress in the double-epenthesis forms in (11) (e.g. [karatí:si]), demonstrated in (16). As stressing any single syllable would result in a violation of HE-IDENT[stress] (16c–d), and stressing both members of a copy-host pair would result in the violation of higher-ranked *CLASH (16a–b), stress defaults to the penult.

(16) Default penultimate stress with medial and final epenthesis: [solodé:re] (11a)

/solder/	NONFINALITY	*CLASH	HE-IDENT[stress]	*LAPSER
a. [solodé:ré:]	*!	*!		
b. [só:ló:dere]		*!		
☞ c. [solodé:re]			*	
d. [soló:dere]			*	*!

While the data are scarce, it appears as though default stress also applies when the copy-host pair occupies the second and third syllables of a four-syllable word: [potoló-k:u] 'my pencil' (Mithun and Basri 1986:231), not *[pó:tolo-k(:)u], is the attested form.[6] This is easily incorporated into the current analysis if we assume that *EXTENDEDLAPSER (abbreviated *EXTLAPSER, (17); Gordon 2002) dominates HE-IDENT[stress] in Selayarese; that is to say, it is more important to keep stress within the trisyllabic right-edge window than it is for the copy-host pair to match in stress.

(17) *EXTLAPSER: Assign one violation mark * if none of the final three syllables are stressed.

[5]A reviewer asks us to clarify our assumptions regarding the analysis of loanword adaptation. We assume here that the input form for Selayarese loans is the Bahasa Indonesian surface form (e.g. /kartu/ → [kará:tu]). This assumption amounts to a claim that a Selayarese speaker would derive [kará:tu] from /kartu/ on-line.

[6]Note that in [potoló-k:u] the stressed vowel does not lengthen; rather, the suffixal consonant geminates. This peculiarity has to do with properties of the possessive suffix, which we do not discuss here (see Broselow 2008:22).

This pattern is analyzed in (18). For reasons of space, we omit NONFINALITY and candidates that violate it. This discussion exhausts the entirety of the available data on stress misapplication in Selayarese copy epenthesis. A summary of the analysis that generates them is given in (19).

(18) Default penultimate stress with medial copy

/potol/ + /ku/	*EXTLAPSER	*CLASH	HE-IDENT[stress]	*LAPSER
a. [póːtolokːu]	*!			*
b. [potóːlókːu]		*!		
☞ c. [potolókːu]			*	
d. [potóːlokːu]			*	*!

(19) Summary of Selayarese analysis (transitive arguments omitted)

$$\begin{tikzcd}
\text{NONFINALITY} \arrow[dr, "1"'] & \text{*EXTLAPSER} \arrow[d, "2"'] & \text{*CLASH} \arrow[dl, "3"] \\
& \text{HE-IDENT[stress]} \arrow[d, "4"] & \\
& \text{*LAPSER} &
\end{tikzcd}$$

[1] [karáːtu] ≻ *[karatúː] [3] [karáːtu] ≻ *[káːráːtu]
[2] [potolókːu] ≻ *[póːtoloku] [4] [túːlisi] ≻ *[tulíːsi]

The above analysis closely resembles the analysis of Selayarese developed by Kitto and de Lacy (1999), who argue that copy vowels and their hosts strive to be durationally (or moraically) equivalent. Their analysis, however, rests on the assumption that all epenthetic vowels are short, an assumption that has since been corrected by Broselow (2008:22) (and was reflected above, see especially (10) and (11)). While this correction to the data does not rule out a copy-host identity analysis based on length along the lines of that proposed by Kitto and de Lacy (1999) – it is possible to modify their analysis so as to account for Broselow's (2008) correction to the data – we have focused here on the role of stress in copy-host identity for the sake of consistency and compatibility with the majority of the literature on Selayarese copy epenthesis.

A reviewer asks how the correspondence-based analysis proposed here accounts for the fact that, in word-final position, copy epenthesis applies only across the coronal continuants /s/, /r/, and /l/. We follow Broselow (2000) in assuming that, in word-final position, epenthesis is a last-resort strategy. Selayarese allows two word-final coda consonants, [ŋ] and [ʔ], and an illicit final coda is converted to one of these segments in all cases where such a mapping would not result in a change in [±continuant] (so /p,t,k,b,d,g/ → [ʔ], and /m,n,ŋ/ → [ŋ]). But /r/, /l/, and /s/, as they are [+continuant], cannot be mapped to [-continuant] [ʔ] or [ŋ]. As epenthesis is preferable to deletion in Selayarese, the attested repair is copy epenthesis following the coronal continuant. For more general discussion of the role of intervening consonants in copy epenthesis, see Section 5.1.2.

2.3 The nature of correspondence in Selayarese copy epenthesis

Before moving on, we consider the nature of the correspondence relation that holds between an epenthetic vowel and its host. So far, we have assumed that an epenthetic vowel stands in correspondence with a neighbor. A potential alternative proposal for the data discussed thus far could be that pairs of identical vowels spanning [r], [l], or [s] must correspond; this requirement, combined with a requirement that corresponding vowels must agree in stress, would derive the misapplication effects analyzed here. This analysis differs from the one proposed above in that it makes crucial reference only to the *surface properties* of strings; the status of a vowel as epenthetic vs. non-epenthetic plays no role in determining whether it stands in correspondence with another. Such an analysis could be implemented in the framework of Agreement by Correspondence (Rose and Walker 2004, Hansson 2010, Bennett 2015a,b, *a.o.*), where a class of CORR constraints compels surface correspondence among featurally similar segments.

There is, however, a problem for this alternative: surface strings alone do not allow us to correctly predict which forms exhibit stress misapplication. In Selayarese, reference to the epenthetic vs. non-epenthetic status of a vowel is necessary. Evidence for this comes from the near-minimal pair in (20) and others like it; more examples can be found in Mithun and Basri (1986). Despite the fact that both forms end in a final [uru] string, stress only misapplies in (20a).

(20) Stress difference in near-minimal pair (Mithun and Basri 1986:217,238)

a.	Antepenultimate stress	já:guru̲	'to box, punch'
b.	Penultimate stress	kaʔmú:ru	'nose'

The relevant difference between (20a) and (20b) lies in the status of the final vowel. In (20a), where stress misapplies, the final vowel is epenthetic; in (20b), where stress does not misapply, the final vowel is not epenthetic. (Recall that a final vowel's epenthetic status can be confirmed independently by the diagnostics mentioned at the beginning of Section 2.2.) Thus whether or not misapplication applies cannot be predicted by reference to surface strings alone; reference to a segment's epenthetic vs. non-epenthetic status is necessary.

To the best of our knowledge, there is no proposal in the Agreement by Correspondence literature that allows a CORR constraint to reference whether or not a given segment is epenthetic. There is thus no proposal in this literature that can account for the data discussed here. In large part for this reason, we follow Kitto and de Lacy (1999) in assuming throughout that a special correspondence relation holds between an epenthetic vowel and its host (for reasons that remain unknown), rather than assuming that the correspondence relationship discussed here is one instantiation of a broader class of surface correspondence phenomena.[7] In addition, were we to pursue an analysis of prosodic misapplication that makes crucial reference to surface similarity, it would necessarily predict that a precondition for prosodic misapplication is featural similarity between an epenthetic vowel and its host. We have not fully explored this prediction, but suggestive evidence that it is incorrect comes from blocking in Irish Gaelic default epenthesis (discussion in Section 3.4).

3 Scottish Gaelic

Our next case of prosodic misapplication in copy epenthesis comes from Scottish Gaelic (Celtic), where heterorganic falling-sonority clusters are broken up by epenthesis. In the Barra dialect (as well as several others), the epenthetic vowel is generally identical to the preceding vowel ((21); Børgstrom 1937, Kenstowicz and Kisseberth 1979, Clements 1986, Bosch and de Jong 1997, Hall

[7] A reviewer asks if, under the theory we adopt, additional kinds of surface correspondence relations are possible. On this point, we are agnostic. The claim that epenthetic vowels stand in correspondence with a host does not exclude the possibility that other kinds of surface correspondence relations exist, but it also does not imply it.

2003, Hammond et al. 2014, *a.o.*).[8] Transcriptions in this section generally follow standard IPA convention, but we follow Børgstrom (1937) and Clements (1986) in their transcription of sonorants (lenited sonorants are [n], [r], [rʲ],[lʲ]; nonlenited sonorants are [N], [L], [R], [Nʲ], [Lʲ]).

(21) Copy epenthesis in Scottish Gaelic (from Clements 1986:328)

a.	/ɔrm/	→ [ɔrɔ̲m]	'on me'
b.	/aLpə/	→ [aLa̲pə]	'Scotland'
c.	/urpəL/	→ [uru̲pəL]	'tail'
d.	/faRkʲə/	→ [faRa̲kʲə]	'rough sea'
e.	/æmʃirʲ/	→ [æmæ̲ʃirʲ]	'time'
f.	/duNxəɣ/	→ [duNu̲xəɣ]	'Duncan'

The epenthetic status of the vowels in (21) can be determined by comparison with near-minimal pairs: they exhibit different prosodic properties than non-epenthetic vowels in the same environments (Bosch and de Jong 1997:2). We focus here on several unexpected interactions between copy epenthesis and other aspects of Scottish Gaelic phonology, especially concerning stress and vowel length. We show that positing a correspondence relation between a copy vowel and its host allows for a unified analysis of these effects. While aspects of the data discussed here have been treated in other recent work on the analysis of copy epenthesis (see esp. Hall 2003, Hammond et al. 2014), the analyses in this section are original to this paper.

3.1 Copy epenthesis and stress

Scottish Gaelic words generally have initial stress (22). But in words where a copy-host pair occupies the first two syllables (23), the initial and peninitial vowels both bear primary stress (see Børgstrom 1937:73,130 for the claim and Bosch and de Jong 1997 for instrumental evidence).

(22) Initial stress (Bosch and de Jong 1997:2; stress marks added)
  a. [áran] 'bread'
  b. [ʃílʲəɣ] 'to rain'
  c. [kálʲak] 'young girl'
  d. [fɛ́nak] 'crow'

(23) Initial and peninitial stress (from (21); stress marks added)
  a. [ɔ́rɔ̲́m] 'on me'
  b. [fáRá̲kʲə] 'rough sea'

To account for the preference for initial stress, we assume that StressLeft (abbreviated StressL, (24)) is undominated. (StressL is a categorical version of a foot-free constraint that prefers left-alignment of stress; for gradient formulations see Garrett's 1994 Edge-Left, or Gordon's 2002 Align($x_{\text{level 1}}$, L, level 0, PrWd).)

(24) StressL: Assign one violation mark * if the initial syllable is stressless.

In words where a copy-host pair occupies the initial and peninitial syllables, the combination of StressL and HE-Ident[stress] (defined in (12)) compels stress to fall on both the copy and its host. The fact that this is the attested result indicates that HE-Ident[stress] ≫ *Clash (defined

[8] There are cases where copy vowels do not completely resemble their hosts as a result of agreement for [±back] with a preceding sonorant (see esp. Clements 1986). As far as we are aware, these words have the same prosodic properties as do the words where copy and host are identical. For expositional simplicity, we do not discuss the cases of partial copy.

in (13)). The tableaux in (25) and (26) illustrate the analysis of stress for words without and with copy epenthesis, respectively.

(25) Initial stress; no copy epenthesis present

/kalʲak/	STRESSL	HE-IDENT[stress]	*CLASH
☞ a. [kálʲak]			
b. [kálʲák]			*!
c. [kalʲák]	*!		

(26) Initial and peninitial stress; copy epenthesis present

/ɔrm/	STRESSL	HE-IDENT[stress]	*CLASH
a. [ɔ́rɔ̱m]		*!	
☞ b. [ɔ́rɔ̱́m]			*
c. [ɔrɔ̱m]	*!		

The analysis as formulated above predicts that stress should apply normally if the first syllable is not a member of a copy-host pair (for example, the analysis predicts that hypothetical /aparm/ should be realized as [ápara̱m]). In all forms with copy epenthesis that we have seen, however, the copy-host pair occupies the initial and peninitial syllables. We do not know if this limitation reflects an accident of the lexicon, or a generalization that should be accounted for in a fuller analysis.

With this analysis of stress in place, we consider how the relationship between stress and vowel length contributes to further prosodic misapplication effects.

3.2 Copy epenthesis and vowel length

In Scottish Gaelic, copy epenthesis interacts with vowel length in two unexpected ways. The first suggests that there are limitations on the kinds of acceptable host vowels. Note that all forms displaying copy epenthesis (as in (21)) have a short host vowel. When the potential host vowel is a long vowel or a diphthong, however, copy epenthesis fails to apply (27). In other words, all attested heterorganic falling sonority clusters in Scottish Gaelic follow a long vowel or diphthong.

(27) Copy epenthesis fails with a potential long host vowel (Hammond et al. 2014:126)[9]

a. /miːrvəlʲəx/ → [miːrvəlʲəx], *[miːri̱vəlʲəx] 'marvelous'
b. /iːrmɛltʃ/ → [iːrmɛltʃ], *[iːri̱mɛltʃ] 'firmament'
c. /du͜əlxəs/ → [du͜əlxəs], *[du͜ələ̱xəs] 'tradition'
d. /ni͜alvər/ → [ni͜alvər], *[ni͜ala̱vər] 'cloudy'

The second unexpected interaction suggests that copy-host pairs display a desire to match each other in length, or moraic count. Bosch and de Jong (1997) show that in words where V_1 and V_2 are both short and present underlyingly, V_1 is longer than V_2 – which is unsurprising, if length is a correlate of stress in Scottish Gaelic. In words where V_1 is a host and V_2 is its copy, however, V_1 and V_2 match in length. (For further discussion of the instrumental data and statistical analyses that support these observations, see Bosch and de Jong 1997.)

[9]Transcriptions are from Hammond et al. (2014). We do not know whether or not the [l]'s and [r]'s in these examples are lenited, so we have not reconciled these with the transcription conventions of Børgstrom (1937) and Clements (1986).

We show that both interactions can be derived under the assumption that a copy vowel and its host strive for identity in length (although since stress and length interact, the analysis is in some ways overdetermined; we make note of this below). The length-matching effects documented by Bosch and de Jong occur when HE-IDENT[length] (defined below) can be satisfied. When high-ranked markedness constraints prevent its satisfaction, copy epenthesis is blocked (as in (27)).

3.2.1 Vowel length categories

To analyze the above patterns, we propose that vowels in Scottish Gaelic are associated with 1, 2, or 3 moras (abbreviated μ), as in (28). Short stressless vowels, as they are the shortest vowels, are monomoraic; short stressed vowels, which are intermediate in duration between short stressless vowels and long vowels, are bimoraic; and long vowels are trimoraic. For instrumental data supporting this division, see Bosch and de Jong (1997).[10] Note that we do not differentiate between long stressed vs. long stressless vowels, as long vowels are licensed only in initial position (and thus are necessarily stressed). This distributional asymmetry is analyzed below.

(28) Vowel lengths in Scottish Gaelic

Vowel type	Short stressless V	Short stressed V	Long V
Moraic Value	μ	$\mu\mu$	$\mu\mu\mu$

3.2.2 Distribution of vowel lengths

We begin by analyzing the restriction of certain vowel lengths to certain positions, putting aside for now interactions with copy epenthesis (see Section 3.2.3). Although we introduce more constraints in this short subsection than are locally necessary, all will be needed for the full analysis. Furthermore, we assume that stress is assigned based on the ranking developed in Section 3.1, and stress placement is not sensitive to mora count – in other words, STRESSL and HE-IDENT[stress] outrank all the length-related constraints introduced below.

In Scottish Gaelic, all stressless vowels are short, i.e. monomoraic. We assume here that this limitation on the length of stressless vowels is enforced by a class of markedness constraints that (stringently) penalize subsets of longer vowels: $*V_{\mu\mu+}$ (29) assigns a penalty to vowels that are at least bimoraic, and $*V_{\mu\mu\mu+}$ (30) assigns a penalty to vowels that are at least trimoraic. (We assume throughout that $*V_{\mu\mu\mu\mu+}$, which assigns a penalty to vowels associated with four or more moras, is undominated; this derives the limitation of Scottish Gaelic vowels to maximally three moras.)

(29) $*V_{\mu\mu+}$: Assign one violation mark * for each vowel that is associated with 2+ moras.

(30) $*V_{\mu\mu\mu+}$: Assign one violation mark * for each vowel that is associated with 3+ moras.

Stressed vowels are minimally bimoraic. We analyze this requirement as an effect of a variety of the STRESS-TO-WEIGHT Principle (SWP; Myers 1987, Riad 1992), which requires stressed vowels to be minimally bimoraic (31). Satisfaction of this constraint through vowel lengthening entails violation of DEP-μ (32), a faithfulness constraint that penalizes mora insertion.

[10] Bosch & de Jong (1997:7) also claim that "vowels previous to epenthetic vowels are shorter than their non-epenthetic counterparts", suggesting that the duration of V_1 depends on whether or not V_2 is epenthetic (a claim later picked up on by Hall 2003:97 and Stanton and Zukoff 2016). This claim, however, is not supported by the data: Bosch & de Jong's statistical analysis did not find a significant effect of the (non-)epenthetic status of V_2 on V_1's duration (see their p. 7).

(31) SWP$_{\mu\mu+}$: Assign one violation mark * for each stressed vowel that is not at least bimoraic.

(32) DEP-μ: Let *y* be an output vowel, and *x* its input correspondent. Assign one violation mark * for each μ in the output associated with *y* that lacks a corresponding μ in the input that is associated with *x*.

Additionally, in stressed syllables, a vowel can be long or short. We assume that this flexibility is due to high-ranked MAX-μ/σ_1 (33), a positional faithfulness constraint that militates against mora deletion in root-initial syllables (abbreviated σ_1; on root-initial positional faithfulness see esp. Beckman 1998:50ff). The more general MAX-μ, which militates against mora deletion in all positions, is defined in (34).

(33) MAX-μ/σ_1: Let *y* be an output vowel in σ_1, and *x* its input correspondent. Assign one violation mark * for each μ in the input associated with *x* that lacks a corresponding μ in the output that is associated with *y*.

(34) MAX-μ: Let *y* be an output vowel, and *x* its input correspondent. Assign one violation mark * for each μ in the input associated with *x* that lacks a corresponding μ in the output that is associated with *y*.

With these constraints, we can derive the attested range and distribution of vowel lengths in Scottish Gaelic words (again, putting aside the exceptional patterns induced by copy epenthesis). Words with a short stressed vowel allow us to infer three crucial rankings. The fact that the stressed vowel must be minimally bimoraic – even if it were to correspond to an input monomoraic vowel – indicates that SWP$_{\mu\mu+}$ dominates both DEP-μ and *V$_{\mu\mu+}$ (compare (35a–b)). In addition, the fact that the stressless vowel must be monomoraic – even it corresponds to a bimoraic input vowel – indicates that *V$_{\mu\mu+}$ dominates MAX-μ (compare (35b–c)).

(35) Tableau for [kálʲak] 'young girl' (with maximally unfaithful hypothetical input)

/ka$_\mu$lʲa$_{\mu\mu}$ak/	SWP$_{\mu\mu+}$	DEP-μ	*V$_{\mu\mu+}$	MAX-μ
a. [ká$_\mu$lʲa$_\mu$ak]	*!			*
☞ b. [ká$_{\mu\mu}$lʲa$_\mu$ak]		*	*	*
c. [ká$_{\mu\mu}$lʲa$_{\mu\mu}$ak]		*	**!	

Accounting for the distribution of trimoraic vowels allows us to infer an additional crucial ranking. The fact that a trimoraic vowel may surface in word-initial position indicates that MAX-μ/σ_1 dominates *V$_{\mu\mu\mu+}$ (36); in any other position, an input trimoraic vowel would be shortened to an output monomoraic vowel, as *V$_{\mu\mu+}$ dominates the general version of MAX-μ (established above).

(36) Tableau for [iːrmɛltʃ] 'firmament'

/i$_{\mu\mu\mu}$rmɛ$_\mu$ltʃ/	MAX-μ/σ_1	*V$_{\mu\mu\mu+}$
☞ a. [i$_{\mu\mu\mu}$rmɛ$_\mu$ltʃ]		*
b. [i$_{\mu\mu}$rmɛ$_\mu$ltʃ]	*!	

For an interim summary, see (37). The reader may wonder, at this point, about the predictions of an analysis that refers to three degrees of vowel length. Does the analysis predict that there should exist languages that make a contrastive three-way vowel length distinction? The answer is yes, and such languages do exist (see e.g. Prince 1980 on Estonian and Baal et al. 2012 on North Saami).

(37) Interim ranking summary

$$\begin{tikzcd}
\text{SWP}_{\mu\mu+} \arrow[d, "1"'] \arrow[dr, "2"] & & \textsc{Max}\text{-}\mu/\sigma_1 \arrow[d, "3"] \\
\textsc{Dep}\text{-}\mu & {*}\text{V}_{\mu\mu+} \arrow[d, "4"] & {*}\text{V}_{\mu\mu\mu+} \\
& \textsc{Max}\text{-}\mu &
\end{tikzcd}$$

1 [ká$_{\mu\mu}$lʲa$_{\mu}$k] ≻ *[ká$_{\mu}$lʲa$_{\mu}$k]
2 [ká$_{\mu\mu}$lʲa$_{\mu}$k] ≻ *[ká$_{\mu}$lʲa$_{\mu}$k]
3 [ká$_{\mu\mu}$lʲa$_{\mu}$k] ≻ *[ká$_{\mu\mu}$lʲa$_{\mu\mu}$k]
4 [í$_{\mu\mu\mu}$rmɛ$_{\mu}$ltʃ] ≻ *[í$_{\mu\mu}$rmɛ$_{\mu}$ltʃ]

3.2.3 Interactions with copy epenthesis

We analyze now the interactions between copy epenthesis and vowel length. As background, we assume here that epenthesis in Scottish Gaelic is motivated by *FALLCC (38), a cover constraint for whatever factors disprefer heterorganic falling-sonority clusters.[11] Satisfying *FALLCC through epenthesis entails a violation of DEP-SEG (39).

(38) *FALLCC: Assign one violation mark * for each heterorganic falling-sonority cluster.

(39) DEP-SEG: Assign one violation mark * for each output segment that does not have an input correspondent.

The fact that epenthesis occurs in Scottish Gaelic demonstrates that *FALLCC dominates DEP-SEG (40). (Note that while default and copy epenthesis both violate DEP-SEG equally, for expositional simplicity we only consider epenthetic candidates with copy vowels.)

(40) Copy epenthesis in Scottish Gaelic

/arm/	*FALLCC	DEP-SEG
a. [arm]	*!	
☞ b. [ara̱m]		*

We focus first on the generalization that copy epenthesis is blocked when the potential host vowel is long (/iːrmɛltʃ/ → [iːrmɛltʃ], *[iːri̱mɛltʃ] 'firmament'; see (27) for others). To analyze this, we assume that an epenthetic vowel and its host must be moraically equivalent: this pressure is encoded as HE-IDENT[length] (41).

(41) HE-IDENT[length]: Assign one violation mark * if an epenthetic vowel and its host are associated with a different number of moras.

When the potential host vowel is long, satisfaction of HE-IDENT[length] would entail inserting a long epenthetic vowel, to match a long host vowel, e.g. *[í$_{\mu\mu\mu}$rí̱$_{\mu\mu\mu}$mɛ$_{\mu}$ltʃ]; or shortening the host to a bimoraic vowel and inserting a bimoraic epenthetic vowel, e.g. *[í$_{\mu\mu}$rí̱$_{\mu\mu}$mɛ$_{\mu}$ltʃ]. (An option in which both vowels are monomoraic, e.g. *[í$_{\mu}$rí̱$_{\mu}$mɛ$_{\mu}$ltʃ], is completely ill-formed given the analysis in (37): since both vowels are stressed, the ranking SWP$_{\mu\mu+}$ ≫ *V$_{\mu\mu+}$ decrees that both must be minimally bimoraic.)

The fact that neither of these options is attested, however, indicates that *FALLCC is subordinated to other markedness and faithfulness constraints. In order for a candidate without epenthesis

[11] We do not discuss here why these clusters in particular might be subject to epenthesis; the issue of why some clusters are broken up by epenthesis and others are not is orthogonal to the current topic of discussion (though see Hall 2003:22ff for one idea as to why only rising-sonority clusters would be targeted, and Fleischhacker 2005 for another.)

(42a) to be preferable to a candidate in which the copy vowel is long (42b), it must be the case that $*V_{\mu\mu\mu+}$ dominates *FALLCC: leaving the cluster unrepaired is preferable to licensing a long vowel outside of initial position. And in order for a candidate without epenthesis (42a) to be preferable to a candidate in which both the host and copy vowel surface with two moras (42c), it must be the case that MAX-$\mu/\sigma_1 \gg$ *FALLCC: leaving the cluster unrepaired is preferable to shortening the initial vowel. Finally, the fact that insertion of a short vowel (with no concomitant change to the length of the host vowel) is not employed indicates that HE-IDENT[length] dominates *FALLCC: if the activity of high-ranked markedness constraints makes it impossible for the copy and host to match in moraic count, it is better to leave the cluster unrepaired. (In (42), the ranking MAX-$\mu/\sigma_1 \gg *V_{\mu\mu\mu+}$ is carried over from the analysis of the non-epenthetic data in (37).)

(42) Copy epenthesis blocked when the potential host vowel is long

/$i_{\mu\mu\mu}$rm$\varepsilon_\mu$ltʃ/	HE-IDENT[length]	MAX-μ/σ_1	$*V_{\mu\mu\mu+}$	*FALLCC
☞ a. [$\acute{i}_{\mu\mu\mu}$rm$\varepsilon_\mu$ltʃ]			*	*
b. [$\acute{i}_{\mu\mu\mu}$r$\underline{\acute{i}}_{\mu\mu\mu}m\varepsilon_\mu$ltʃ]			**!	
c. [$\acute{i}_{\mu\mu}$r$\underline{\acute{i}}_{\mu\mu}m\varepsilon_\mu$ltʃ]		*!		
d. [$\acute{i}_{\mu\mu\mu}$r$\underline{\acute{i}}_{\mu\mu}m\varepsilon_\mu$ltʃ]	*!		*	

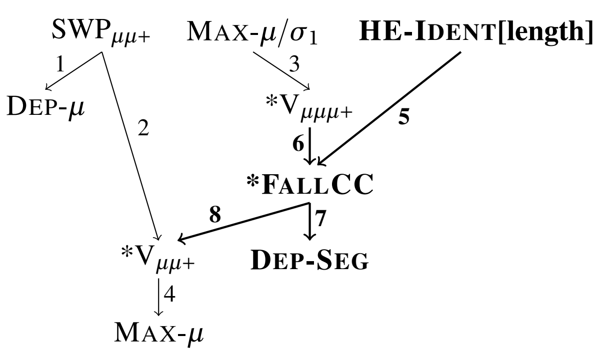
The behavior of copy-host sequences in which the host vowel is short motivates one further constraint ranking. As both copy and host bear stress (due to undominated HE-IDENT[stress], Section 3.1), both must surface as bimoraic due to high-ranked $SWP_{\mu\mu+}$ (compare (43a–c)). The analysis of length-matching in these forms is thus overdetermined; both $SWP_{\mu\mu+}$ and HE-IDENT[length] prefer (43c), where both vowels are bimoraic, to (43b). In any case, the fact that we see repair of the cluster when the host vowel is short thus indicates that inserting a short (stressed) vowel is preferable to licensing a falling-sonority cluster; *FALLCC dominates $*V_{\mu\mu+}$ (compare (43c–d)).

(43) Length-matching when the host is short

/$a_\mu$rm/	$SWP_{\mu\mu+}$	HE-IDENT[length]	*FALLCC	$*V_{\mu\mu+}$
a. [$\acute{a}_\mu$r$\underline{\acute{a}}_\mu$m]	*!*			
b. [$\acute{a}_{\mu\mu}$r$\underline{\acute{a}}_\mu$m]	*!	*		*
☞ c. [$\acute{a}_{\mu\mu}$r$\underline{\acute{a}}_{\mu\mu}$m]				**
d. [$\acute{a}_{\mu\mu}$rm]			*!	*

In sum, appealing to a desire for moraic equivalence between an epenthetic vowel and its host yields an account of interactions between copy epenthesis and vowel length in Scottish Gaelic. A summary of the full analysis is in (44). Transitive rankings are omitted; constraints introduced and ranking arguments established through consideration of the copy epenthesis data are bolded.

(44) Summary of Scottish Gaelic analysis



1 [ká$_{\mu\mu}$l^{j}a$_{\mu}$k] ≻ *[ká$_{\mu}$l^{j}a$_{\mu}$k]
2 [ká$_{\mu\mu}$l^{j}a$_{\mu}$k] ≻ *[ká$_{\mu}$l^{j}a$_{\mu}$k]
3 [ká$_{\mu\mu}$l^{j}a$_{\mu}$k] ≻ *[ká$_{\mu\mu}$l^{j}a$_{\mu\mu}$k]
4 [í$_{\mu\mu\mu}$rmɛ$_{\mu}$ltʃ] ≻ *[í$_{\mu\mu}$rmɛ$_{\mu}$ltʃ]
5 [í$_{\mu\mu\mu}$rmɛ$_{\mu}$ltʃ] ≻ *[í$_{\mu\mu\mu}$r$\underline{\text{í}}_{\mu\mu}$mɛ$_{\mu}$ltʃ]
6 [í$_{\mu\mu\mu}$rmɛ$_{\mu}$ltʃ] ≻ *[í$_{\mu\mu\mu}$r$\underline{\text{í}}_{\mu\mu\mu}$mɛ$_{\mu}$ltʃ]
7 [á$_{\mu\mu}$r$\underline{\text{á}}_{\mu\mu}$m] ≻ *[á$_{\mu\mu}$rm]
8 [á$_{\mu\mu}$r$\underline{\text{á}}_{\mu\mu}$m] ≻ *[á$_{\mu\mu}$rm]

3.3 Blocking in Irish Gaelic

In Irish Gaelic, like Scottish Gaelic, heterorganic falling-sonority clusters are broken up by epenthesis (for further restrictions and exceptions to this generalization, see Ní Chiosáin 1999). Unlike epenthesis in Scottish Gaelic, in Irish Gaelic a default vowel is inserted: [ə] after a non-palatalized consonant, and [i] after a palatalized consonant. But as in Scottish Gaelic, epenthesis only occurs if the preceding vowel is short (45); if the preceding vowel is long, epenthesis is blocked (46). (In (46), all examples involve potential epenthesis into a non-palatalized cluster. We were unable to find parallel examples involving a palatalized cluster, but we expect the same result would obtain.)

(45) Default epenthesis occurs after a short vowel (Ní Chiosáin 1999:560, 565)
  a. /bar^{j}b^{j}r^{j}i/ → [bar$^{j}\underline{\text{i}}$b^{j}r^{j}i] 'Bairbre (name)'
  b. /gorem/ → [gor$\underline{\text{ə}}$m] 'blue'
  c. /ar^{j}g^{j}əd/ → [ar$^{j}\underline{\text{i}}$g^{j}əd] 'money'
  d. /ʃk^{j}el^{j}g^{j}/ → [ʃk^{j}el$^{j}\underline{\text{i}}$g^{j}] 'crag'

(46) Default epenthesis is blocked after a long vowel (Ní Chiosáin 1999:565)
  a. /faːrbi/ → [faːrbi], *[faːr$\underline{\text{ə}}$bi] 'wrinkle'
  b. /t^{j}eːrmə/ → [t^{j}eːrmə], *[t^{j}eːr$\underline{\text{ə}}$mə] 'term'
  c. /l^{j}eːrgəs/ → [l^{j}eːrgəs], *[l^{j}eːr$\underline{\text{ə}}$gəs] 'insight'
  d. /duəlgəs/ → [duəlgəs], *[duəl$\underline{\text{ə}}$gəs] 'duty'

If all epenthetic vowels (copy and default) stand in correspondence with a host (as assumed in Section 1.1), the analysis developed for Scottish Gaelic blocking directly extends to the Irish Gaelic data in (45–46). Under an analysis where all epenthetic vowels correspond with a vocalic host, the only relevant difference between Irish and Scottish Gaelic is the relative ranking of the HE-IDENT constraints that refer to vocalic features (e.g. HE-IDENT[low]) and the collection of constraints that prefer insertion of a default vowel. As the constraints that compel prosodic identity are distinct from those that compel featural identity, HE-IDENT constraints can compel prosodic misapplication regardless of the epenthetic vowel's featural relationship to its host. Thus the fact that we find identical blocking effects in Scottish Gaelic (with copy epenthetic vowels) and Irish Gaelic (with default epenthetic vowels) provides some novel support for Kitto and de Lacy's (1999) hypothesis that all epenthetic vowels, whether default or copy, stand in correspondence with a host.

Note that, if this analysis is correct, we would expect Irish Gaelic default epenthetic vowels to be durationally equivalent to their host vowels. More broadly, given the claim that all epenthetic vowels stand in correspondence with a neighbor, cross-linguistically, we would expect to find a range of prosodic misapplication effects in default and copy epenthesis alike. We do not know if either prediction is borne out.

3.4 Discussion

This section has argued that misapplication effects involving length and stress in Scottish Gaelic can be attributed to a desire for prosodic identity between an epenthetic vowel and its host. Here we briefly consider whether or not alternative analyses of copy epenthesis can also generate the data.

We focus first on analyses that invoke *feature spreading* as a means of explaining the resemblance between a copy vowel and its host. Under these analyses (e.g. Clements 1986, 1991; Gafos and Lombardi 1999; Halle et al. 2000; Shademan 2002; Kawahara 2007), the copy obtains its vowel place (or V-Place) features from a nearby host (illustration in (47), following Kawahara (2007)).

(47) Copying as spreading (from Kawahara 2007:5)

/μ/	→	[μ	μ]
Rt		Rt	Rt
V-Place		V-Place	

An analysis along these lines accounts for the fact that Scottish Gaelic copy vowels featurally resemble their hosts, but such an analysis does not extend straightforwardly to the prosodic misapplication effects discussed in this section. The fact that both copy and host bear stress, for example, is not easily analyzed by a spreading-based theory: identity for segmental features does not automatically entail identity for suprasegmental features, and furthermore it is commonly assumed that stress is unlike segmental features (and some other suprasegmental features, i.e. tone) in that it cannot undergo autosegmental spreading. In addition, the fact that copy epenthesis is blocked following a long vowel in Scottish Gaelic is difficult to explain: there is no obvious reason why the decision to insert a new vowel and spread [V-Place] onto it should depend on the length of the preceding vowel. The correspondence-based analysis, however, provides a unified explanation of these facts.

In a study on the phonetics and phonology of vowel intrusion, Hall (2003) argues that the Scottish Gaelic phenomena discussed here can be analyzed if we assume that a single gesture underlies the copy-host pair. In the form [ɔrɔm] (21a), for example, the illusion of copy epenthesis emerges when the constellation of gestures that yield an [r] becomes superimposed on those that yield an [ɔ] (48). The resulting twin [ɔ]'s are thus both exponents of the same underlying gesture.

(48) Copy epenthesis as gestural reorganization (based on Steriade 1990)

Tongue body (/ɔ/):	[- - *ɔ* - - - - - - - - - *r* *m* -]	Tongue body (/ɔ/):	[- *ɔ* - - - - *r* - - - *ɔ* - - - *m* - -]
Tongue tip (/r/):	[- - -]	Tongue tip (/r/):	[- - -] ←
Lips (/m/):	[- - -]	Lips (/m/):	[- - -]

Hall (2003) proposes that the length-matching effect in forms with copy epenthesis (e.g. [ɔ$_{\mu\mu}$rɔ$_{\mu\mu}$m]) can be attributed to a desire for the vocalic and consonantal gestures to be centered at the same time point (for discussion and analysis see Hall 2003:111–116). The fact that copy epenthesis is blocked following a long vowel is attributed to another factor, namely the differing effects that short vs. long vowels have on the moraicity of the following sonorant. Adapting Hall's analysis to the current assumptions regarding the moraic values of different vowel lengths, the idea is that the Scottish Gaelic syllable is maximally trimoraic. A coda sonorant can bear a mora when the nucleus is short (as in /$\text{ɔ}_{\mu}\text{r}_{\mu}$m/), but not if the nucleus is long (as in hypothetical /$\text{ɔ}_{\mu\mu\mu}$rm/). Hall claims that if the sonorant is not moraic, it cannot migrate into the middle of the nucleus. Gestural intrusion is thus possible in /$\text{ɔ}_{\mu}\text{r}_{\mu}$m/ (yielding [ɔrɔm]), where the vowel is short – but not in /$\text{ɔ}_{\mu\mu\mu}$rm/, where the vowel is long.

There are open questions regarding the details and motivation of this analysis. Why, for example, is sonorant moraicity a prerequisite for gestural intrusion? And given that consonantal intrusion into a vowel requires a concomitant lengthening of that vowel (Hall 2003:114), why does the resulting VRV nucleus not violate the trimoraic maximum? What we wish to focus on, though, is a perceived loss of generalization. We have argued that both the blocking and the length-matching effects can be attributed to one motivating factor: copy vowels and their hosts *strive for prosodic identity*. Acknowledging this allows for a simple and economical analysis of the prosodic misapplication phenomena. But in a gestural intrusion theory, where such a demand for identity is not recognized, all identity effects are treated as separate phenomena and analyzed with different mechanisms. In short, the correspondence-based analysis directly accounts for a strong generalization that the gestural approach treats as incidental (for further discussion on this point, see Section 4.6).

Another argument for the correspondence-based approach comes from the parallel between the Scottish and Irish Gaelic blocking data (Section 3.3). As a reviewer notes, it is doubtful that Irish Gaelic epenthesis should be analyzed as a case of gestural intrusion. Hall's (2003, 2006) diagnostics for gestural intrusion include that the vowel be phonologically invisible, and optional or durationally variable. Regarding the invisibility of Irish Gaelic epenthetic vowels, Ní Chiosáin (1999) shows that whether or not epenthesis occurs in a given cluster depends on the metrical structure of the word that contains that cluster (see esp. pp. 565–570). That is to say, epenthetic vowels in Irish Gaelic crucially must be *visible*, as their distribution is constrained by metrical properties. Regarding the optionality and duration of Irish Gaelic epenthesis: the description provided by Ní Chiosáin (1999) suggests that epenthesis is obligatory, regardless of speech rate, and she does not suggest that they are durationally variable. Given that Hall's analysis of Scottish Gaelic blocking crucially relies on a sonorant's ability to participate in gestural intrusion, it is not clear that this analysis could extend to the Irish Gaelic data. But as discussed above, the correspondence-based analysis of Scottish Gaelic blocking extends naturally to the Irish Gaelic case: default vowels, like copies, seek durational equivalence with their hosts.[12]

In short, we have shown that invoking a correspondence relationship between an epenthetic vowel and its host allows us to account for cases of prosodic misapplication in Scottish Gaelic copy epenthesis, as well as analogous effects in Irish Gaelic default epenthesis. These results, combined with the difficulties faced by alternative analyses, provide support for the central claim of this paper: a correspondence relation is an essential component of any successful theory of copy epenthesis.

[12] Hall's (2003) position on Irish Gaelic is consistent with ours: she claims that the inserted vowels were historically intrusive, but have been reanalyzed as segments (e.g. p. 72, 95). It is worth noting however that Irish Gaelic epenthesis does satisfy several of Hall's other criteria for gestural intrusion: the vowel's quality is schwa (but can be influenced by the surrounding consonants), and it generally occurs in heterorganic sonorant-stop clusters. If it turns out that a gestural analysis of Irish Gaelic epenthesis is appropriate, this argument for the correspondence-based approach is invalid.

4 Ho-Chunk

In Ho-Chunk (Siouan; also known as Winnebago, Hocák, Hocank), a process of copy epenthesis known as *Dorsey's Law* often causes stress to misapply, though in complex and descriptively non-uniform ways (Miner 1979 *et seq.*, Hale and White Eagle 1980, Halle and Vergnaud 1987, Alderete 1995, Hayes 1995, Hall 2003, *a.o.*). We argue that this well-known interaction between Dorsey's Law and stress placement results from a drive for stress identity between a copy vowel and its host. We show, in addition, that the proposed analysis can be extended to capture additional prosodic misapplication effects in Ho-Chunk, including a case of *back-copying* of vowel length (Section 4.4), as well as unexpected patterns of vowel nasalization and ablaut (Section 4.5).

The data discussed in this section are for the most part well-known (with the exception of the length neutralization data in Section 4.4.2, which is a novel observation). The stress application facts have in particular played a major role in the literature on interactions between epenthesis and word-level prosody; for discussion, see the references cited in the above paragraph. For the most part, however, the relevance of these data to correspondence-based theories of copy epenthesis has not been recognized, nor have they yet been analyzed in such a framework (with the exception of the vowel nasalization and ablaut facts; see Kitto and de Lacy 1999:14–15, 22). Our contribution, then, is a demonstration of how all relevant misapplication phenomena receive a uniform analysis in a framework that explicitly acknowledges a drive for identity between a copy vowel and its host.

4.1 Transcription of stress

In Ho-Chunk, voiceless obstruent + sonorant clusters are broken up by copy epenthesis. Words containing a Dorsey's Law (DL) sequence can often be identified by the fact that they bear a different stress pattern than otherwise equivalent non-DL words (compare (49a–b)).

(49) Dorsey's Law and stress (examples from Miner 1989; stress marks in (a.i–ii) ours)

a. DL words: stress on 1st and 2nd syllable
   - (i) /kre/ → *ké̱ré* 'to leave returning'
   - (ii) /sretʃ/ → *sé̱rétʃ* 'long, tall'

b. Equivalent non-DL words: stress on 2nd syllable only
   - (i) /hiwãx/ → *hiwã́x* 'to ask'
   - (ii) /raǰox/ → *raǰóx* 'to break in the mouth'

Here and throughout, we assume that stress is shared between the members of a copy-host pair: if either the copy or host bears a stress, so must the other. While this is not what many recent sources and analyses assume, we believe there is sufficient motivation for such an assumption.

The motivation comes in part from instrumental data provided by Hall (2003:172–173). The three pitch tracks provided in Hall's study, reproduced (with her permission) as Figures 1–3 below, suggest that DL sequences have pitch patterns unlike those found in other bimoraic or disyllabic sequences. In disyllabic words where the two vowels comprise a DL sequence (Figure 1), both vowels bear high pitch; in other disyllabic words (Figure 2), however, only the second vowel bears high pitch. DL sequences also behave unlike long vowels, which bear high pitch on only their first constituent mora (Figure 3). As pitch is the primary (and perhaps only) cue to stress in Ho-Chunk (Miner 1989, Hall 2003), Figure 1 is consistent with a view in which both vowels in a DL sequence bear stress. (We realize that the pitch tracks are not ideal, as the axes are difficult to read and the

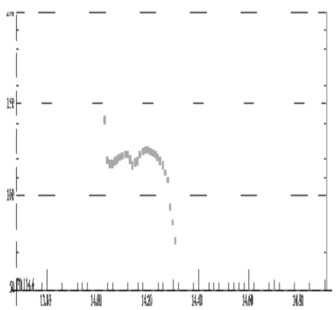


Figure 1: *se̱retʃ*
Hall 2003:173

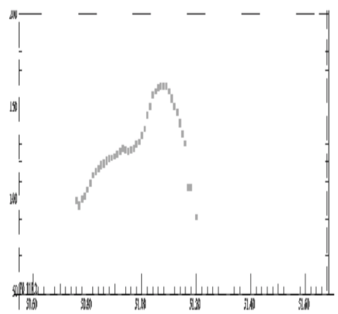


Figure 2: *warutʃ*
Hall 2003:172

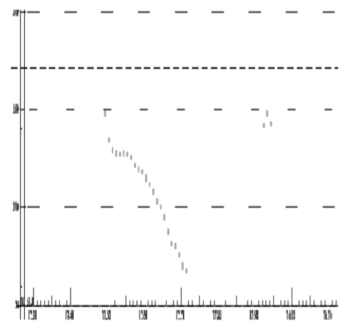


Figure 3: *seːp*
Hall 2003:173

pitch tracks are not segmented. What is important for the present purposes, however, is only the differences among the contours: the DL sequence in Figure 1 has a pitch plateau, which can be compared to the rise in Figure 2 and the fall in Figure 3. Due to the figures' ability to illustrate this contrast, we have elected to include them, despite their impoverishment in other domains.)

Further evidence that both vowels in a DL sequence can bear stress comes from our readings of the early sources. All sources on Ho-Chunk stress agree that, when stress falls in particular positions within the word, some part of a DL sequence must carry stress. The transcription of stress in such DL sequences has, however, varied across descriptions. Starting with Hale and White Eagle (1980), work on stress in Ho-Chunk has assumed that only one member of the copy-host pair bears stress. But in work prior to Hale and White Eagle (1980), there is consistent acknowledgment of prominence on both members of the DL sequence. Susman (1943:13) writes that, in DL sequences "secondary stress seems to attach equally to both syllables". Miner (1979) transcribes secondary stress on the copy and primary stress on the host (e.g. *hipè̱rés* 'know'), but notes that often "the secondarily accented syllable has almost as much accent as, or even as much as (but never more than) the primarily accented one" (p. 27). Miner's move in later work (1981, 1989) to transcribe stress on only one portion of the DL sequence is made without comment.

As we have not obtained access to high-quality recordings, we refrain from speculation as to the source of the difference in stress transcription between the pre-1980 and post-1980 sources. But based on these descriptions and Hall's (2003) pitch-tracking data, we believe there is sufficient motivation to assert that when one vowel in a DL sequence bears stress, so must the other.

4.2 Stress in non-DL words

Even setting aside its interactions with copy epenthesis, the Ho-Chunk stress system is fairly complex. To simplify matters, in what follows we focus only on the distribution of primary stress.

As is evident from (50a–b), stress prefers to fall on the third mora from the left (the post-peninitial mora): this means that stress falls on the third syllable if the first two syllables are light (notated as σ_μ), and on the second syllable if the first syllable is heavy (notated as $\sigma_{\mu\mu}$).[13] If the word contains only two light syllables, stress falls on the second mora/syllable (50c).

[13]Systems displaying post-peninitial stress are rare, but not unattested (cf. Gordon 2002): Kager (2012) notes that post-peninitial position is possible in Azkoitia Basque (Hualde 1998) and Choguita Rarámuri (Caballero 2008, 2011).

(50) Default stress pattern in Ho-Chunk

a. Examples of $\sigma_\mu\sigma_\mu\acute{\sigma}_\mu\ldots$ forms

(i)	*hipirák*	'belt'	Miner 1989:152
(ii)	*hiǰowíre*	'fall in'	Miner 1979:28
(iii)	*hokiwároke*	'swing (n.)'	Miner 1979:28

b. Examples of $\sigma_{\mu\mu}\acute{\sigma}_\mu\ldots$ forms

(i)	*ki:zá*	'fight'	Miner 1979:31
(ii)	*mã:sú*	'feather'	Miner 1979:31
(iii)	*ta:nĩ́žu*	'sugar'	Miner 1989:152

c. Examples of $\sigma_\mu\acute{\sigma}_\mu$ forms

(i)	*hiwã́k*	'to ask'	Miner 1989:152
(ii)	*hosgáč*	'playground'	Miner 1989:152

The pattern in (50) can be modeled with the foot-free constraints in (51–53). (These constraints follow the system proposed by Gordon 2002; see also Garrett 1994, Alderete 1995 for (52–53).)

(51) NONINITIALITY (NONINIT): Assign one violation mark * if stress falls on the first mora of the word.

(52) EXTENDEDNONINITIALITY (EXTNONINIT): Assign one violation mark * if stress falls on one of the first two moras of the word.

(53) *EXTENDEDLAPSEL (*EXTLAPSEL): Assign one violation mark * if no stress falls within the first three moras of the word.

Together, these three constraints correctly predict post-peninitial stress. As shown in (54c), a candidate with post-peninitial stress satisfies all of these metrical constraints. By contrast, a candidate with initial stress (54a) violates both NONINIT and EXTNONINIT; a candidate with peninitial stress (54b) violates EXTNONINIT; and a candidate with fourth-mora stress (54d) violates *EXTLAPSEL. These constraints also correctly predict that bimoraic words have peninitial stress: a candidate with initial stress (55a) is harmonically bounded by one with peninitial stress (55b).

(54) Third mora (post-peninital) stress in 3+ mora words

/hiǰowire/	NONINIT	EXTNONINIT	*EXTLAPSEL
a. hı́ǰowire	*!	*!	
b. hiǰówire		*!	
☞ c. hiǰowíre			
d. hiǰowiré			*!

(55) 2nd mora (peninitial) stress in bimoraic words

/hiwãx/	NONINIT	EXTNONINIT	*EXTLAPSEL
a. híwãx	*!	*	
☞ b. hiwã́x		*	

In addition to the preference for post-peninitial stress, there are several additional factors governing stress placement that involve syllable weight. For example, if a long vowel bears stress, stress must fall on its first mora; if a diphthong bears stress, stress must fall on the more sonorous member (see Miner 1989 on both of these points). We do not discuss these facts further because they are orthogonal to the behavior of stress in words with DL sequences, to which we now turn.

4.3 Copy epenthesis and misapplication of stress

As summarized below, whether or not (and how) DL sequences influence stress assignment depends on their position within the word, as well as the word's length. We focus first on patterns of stress assignment in four-syllable words, as these demonstrate many of the necessary crucial rankings, and then show how the analysis can be extended to generate the remainder of the data. We primarily restrict our discusssion to words consisting entirely of light syllables, so as to neutralize the distinction between moras and syllables (though see the end of Section 4.3.2). In the remainder of this section, we will mainly refer to position of stress in terms of syllables, for ease of comprehension. (In Table 1, all DL words are from Miner 1989; non-DL words are from (50).)

Table 1: Summary of stress differences in Ho-Chunk DL vs. non-DL words

Shape	Non-DL word	Dorsey's Law word (DL sequence(s) in [ ])	
$\sigma_1\sigma_2$	$\sigma_1\acute{\sigma}_2$ ex. *hiwã́x* 'to ask'	$[\acute{\sigma}_1\acute{\sigma}_2]$ ex. [*ké̲ré*]	(← /kre/) 'to leave returning'
$\sigma_1\sigma_2\sigma_3$	$\sigma_1\sigma_2\acute{\sigma}_3$ ex. *hipirák* 'belt'	$[\sigma_1\sigma_2]\acute{\sigma}_3$ [*xe̲re*]*hí*	(← /xrehi/) 'to boil'
		$\sigma_1[\acute{\sigma}_2\acute{\sigma}_3]$ ex. *hi*[*pé̲rés*]	(← /hipres/) 'to know'
$\sigma_1\sigma_2\sigma_3\sigma_4$	$\sigma_1\sigma_2\acute{\sigma}_3\sigma_4$ ex. *hiǰowíre* 'fall in'	$[\sigma_1\sigma_2]\acute{\sigma}_3\sigma_4$ ex. [*xo̲ro*]*ǰíke*	(← /xroǰíke/) 'hollow'
		$\sigma_1[\sigma_2\sigma_3]\acute{\sigma}_4$ ex. *hi*[*ko̲ro*]*hó*	(← /hikroho/) 'to prepare'
		$\sigma_1\sigma_2[\acute{\sigma}_3\acute{\sigma}_4]$ ex. *hiru*[*pĩ̲́nĩ́*]	(← /hirupni/) 'to twist'
		$[\sigma_1\sigma_2][\acute{\sigma}_3\acute{\sigma}_4]$ ex. [*ke̲re*][*pã̲́nã́*]	(← /krepna/) 'unit of ten'
$\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5$	$\sigma_1\sigma_2\acute{\sigma}_3\sigma_4\sigma_5$ ex. *hokiwároke* 'swing'	$\sigma_1[\sigma_2\sigma_3][\acute{\sigma}_4\acute{\sigma}_5]$ ex. *wa*[*ki̲ri*][*pá̲rás*]	(← /wakripras/) 'flat insect'

4.3.1 Stress in four-syllable words

The table in (56) shows the types of four-syllable words that contain only light syllables, categorized by the presence and position of a DL sequence. In (56a–b), where a DL sequence is either absent or occupies the first and second syllables, stress applies normally. In (56c), where the DL sequence occupies the second and third syllables, stress misapplies: it falls on the fourth syllable, rather than the default third. And in (56d), where the DL sequence occupies the third and fourth syllables, stress also misapplies (though in a different way): it falls on both the third and the fourth syllables.

(56) Stress in four-syllable words

Word type	Stress pattern	Example	Stress application type
a. Non-DL word	$\sigma_1\sigma_2\acute{\sigma}_3\sigma_4$	*hiǰowíre*	Normal application: Stress on σ_3
b. DL in $[\sigma_1\sigma_2]$	$[\sigma_1\sigma_2]\acute{\sigma}_3\sigma_4$	[*xo̱ro*]*ǰíke*	Normal application: Stress on σ_3
c. DL in $[\sigma_2\sigma_3]$	$\sigma_1[\sigma_2\sigma_3]\acute{\sigma}_4$	*hi*[*ko̱ro*]*hó*	Misapplication: Stress on σ_4
d. DL in $[\sigma_3\sigma_4]$	$\sigma_1\sigma_2[\acute{\sigma}_3\acute{\sigma}_4]$	*hiru*[*pĩ̱́nĩ́*]	Misapplication: Stress on σ_3 & σ_4

We argue here that the cases of misapplication in (56c–d) occur to render the two vowels in the DL sequence stress-identical: the behavior of (56c) and (56d) can be explained if HE-IDENT[stress] – the same constraint employed in the analyses of Selayarese and Scottish Gaelic above – dominates certain metrical constraints. Under our analysis, the misapplication patterns in (56c–d) differ because there are two ways to satisfy HE-IDENT[stress]: by stressing *both* the copy and the host (as in *hiru*[*pĩ̱́nĩ́*] (56d)), or by stressing *neither* the copy nor the host (as in *hi*[*ko̱ro*]*hó* (56c)). Misapplication occurs in words of these shapes because the position targeted for default stress assignment, σ_3, is part of a DL sequence. Due to the activity of HE-IDENT[stress], applying normal σ_3 stress would trigger the need for stress on the other member of the DL sequence. The pattern that results is determined by the position of the other member of the DL sequence, and whether or not the metrical constraints permit that position to bear a stress.

We focus first on words where the DL sequence occupies the third and fourth syllables, e.g. *hiru*[*pĩ̱́nĩ́*] ($\sigma_1\sigma_2[\acute{\sigma}_3\acute{\sigma}_4]$, (56d)). If stress applied normally, to σ_3 alone, HE-IDENT[stress] would be violated (57a). Stressing the other DL vowel in order to satisfy HE-IDENT[stress] places an additional stress on σ_4 (57c); this option violates *CLASH, but no other relevant constraints.[14] If HE-IDENT[stress] were satisfied by moving stress off the DL sequence to σ_2, this would incur a violation of EXTNONINIT (57b). The fact that (57c) is the preferred outcome indicates that, in Ho-Chunk, HE-IDENT[stress] and EXTNONINIT dominate *CLASH.

(57) Misapplication: 3rd and 4th syllable stress with DL in $[\sigma_3\sigma_4]$ (hiru[pĩ̱́nĩ́])

/hirupni/	HE-IDENT[stress]	EXTNONINIT	*CLASH
a. hirupĩ̱́nĩ	*!		
b. hirúpĩ̱nĩ		*!	
☞ c. hirupĩ̱́nĩ́			*

We turn now to words where the DL sequence occupies the second and third syllables, e.g. *hi*[*ko̱ro*]*hó* ($\sigma_1[\sigma_2\sigma_3]\acute{\sigma}_4$, (56c)). Normal application of stress to σ_3 alone would result in a violation of HE-IDENT[stress] (58a). If stress remained on σ_3, HE-IDENT[stress] could be satisfied by additionally stressing σ_2 (58b), but this option incurs a violation of EXTNONINIT. The other option is to push stress off the DL sequence to σ_4, which incurs a violation of *EXTLAPSEL. The fact that (58c) is the attested option therefore shows us that both HE-IDENT[stress] and EXTNONINIT dominate *EXTLAPSEL.

14 Note that what we call *CLASH here could just as well be defined as a constraint that disprefers high pitch plateaus.

(58) Misapplication: 4th syllable stress with DL in [$\sigma_2\sigma_3$] (hi[ko̲ro]hó)

/hikruni/	HE-IDENT[stress]	EXTNONINIT	*EXTLAPSEL	*CLASH
a. hiku̲rúni	*!			
b. hikú̲rúni		*!		*
☞ c. hiku̲runí			*	

In forms where the DL sequence occupies the first two syllables (e.g. [*xo̲ro*]*ǰíke*), stress applies normally: when the DL sequence is not targeted for default stress assignment, there is no reason for stress to misapply. As shown in (59), stressing the third syllable of [*xo̲ro*]*ǰíke* results in satisfaction of HE-IDENT[stress], as well as the full set of metrical constraints (NONINIT is excluded for space).

(59) Stress applies normally in [*xo̲ro*]*ǰíke*

/xorǰike/	HE-IDENT[stress]	EXTNONINIT	*EXTLAPSEL	*CLASH
☞ a. xoro̲ǰíke				
b. xoro̲ǰiké			*!	
c. xóró̲ǰike		*!		*
d. xoró̲ǰike	*!	*!		

In sum, the two types of misapplication in four-syllable words show us that HE-IDENT-[stress] must always be satisfied, even at the expense of *CLASH (as in (57)) and *EXTLAPSEL (as in (58)).

4.3.2 Stress in other word shapes

The other patterns in Table 1 are consistent with the rankings demonstrated in four-syllable words and, in addition, reveal several others. We focus first on disyllabic words. In disyllabic words without a DL sequence, stress falls on the second syllable (e.g. *hiwã́x*). This outcome follows from any ranking of the metrical constraints, as demonstrated in (55) above. In disyllabic words that do consist (entirely) of a DL sequence, however, both vowels are stressed, as in [*ké̲ré*]. This double-stress pattern violates both NONINIT and *CLASH, as well as EXTNONINIT. The fact that double-stress is nonetheless attested indicates that HE-IDENT[stress] dominates NONINIT and *CLASH (60). (To ensure that some stress surfaces in [*ké̲re*] – as opposed to the alternative [*kere*], with no stress – we assume that all Ho-Chunk words must have at least one stress.)

(60) 1st and 2nd syllable stress in two-syllable DL words ([*ké̲ré*] 'to leave returning')

/kre/	HE-IDENT[stress]	NONINIT	*CLASH
☞ a. ké̲ré		*	*
b. ké̲re	*!	*	
c. ke̲ré	*!		

We move on now to trisyllabic words. Trisyllabic words where a DL sequence occupies σ_2 and σ_3 reveal that NONINITIALITY dominates *CLASH: it is better to stress both the second and third syllables than it is to stress just the first (61). (Trisyllabic words in which σ_1 and σ_2 comprise a DL sequence, like [*xe̲re*]*hí*, display normal application. In these cases, there is no motivation for stress to misapply.)

(61) 2nd & 3rd syllable stress in three-syllable words with DL in [$\sigma_2\sigma_3$] (*hi*[*pé̲rés*] 'to know')

/hipres/	HE-Ident[stress]	NonInit	ExtNonInit	*Clash
☞ a. hipé̲rés			*	*
b. hípe̲res		*!	*	
c. hipe̲rés	*!			

In longer words that contain two adjacent DL sequences (e.g. [*ke̲re*][*pã̲́nã́*] and *wa*[*ki̲ri*][*pá̲rás*]), the attested stress patterns are predicted by the current ranking, as demonstrated in (62).

(62) Second DL receives stress (*wa*[*ki̲ri*][*pá̲rás*] 'flat insect')

/wakripras/	HE-Ident[stress]	ExtNonInit	*ExtLapseL	*Clash
☞ a. waki̲ripá̲rás			*	*
b. wakí̲rípa̲ras		*!		*
c. waki̲rípa̲ras	*!			
d. waki̲ripá̲ras	*!		*	
e. wáki̲ripa̲ras		*!		

The words we have discussed so far contain only light syllables, but our analysis makes predictions regarding words of other shapes as well. We close by discussing two kinds of words that have one heavy syllable. In the first kind, a long initial syllable is immediately followed by a DL sequence (e.g. *bo:*[*pṹ̲nṹs*] 'to hit at random'; Miner 1989:149). In such cases, we observe the same sort of double-stressing misapplication as is found in *hiru*[*pĩ̲́nĩ́*] (57). This is exactly as predicted: the two moras of the long vowel that precede the DL sequence in *bo:*[*pṹ̲nṹs*] are equivalent to the two moras of the two light syllables that precede the DL sequence in *hiru*[*pĩ̲́nĩ́*]. In the second kind, a DL sequence occupying the first two syllables is immediately followed by a long vowel (e.g. hypothetical [*xo̲ro*]*ǰí:*). We have been unable to locate such forms, but they are predicted to display normal application, as the third mora is not part of a DL sequence.

4.3.3 Summary of Ho-Chunk stress

This concludes the analysis of the Ho-Chunk stress misapplication data. The analysis is summarized in (63); transitive arguments are omitted.

(63) Ranking hierarchy for Ho-Chunk stress

HE-Ident[stress]
1 ↙ ↘ 2
NonInit   ExtNonInit
3 ↘ 4 ↙ ↘ 5
*Clash   *ExtLapseL

1 [ké̲ré] ≻ *[ke̲ré]
2 [hipé̲rés] ≻ *[hipe̲rés]
3 [hipé̲rés] ≻ *[hípe̲res]
4 [hirupĩ̲́nĩ́] ≻ *[hirúpĩ̲nĩ]
5 [hiku̲runĩ́] ≻ *[hikú̲rúnĩ]

At first glance, the Ho-Chunk misapplication pattern appears fairly complex, and many previous accounts of Ho-Chunk stress have needed to appeal to otherwise unmotivated analytic conventions to capture the full range of facts (e.g. the Domino Condition; Halle and Vergnaud 1987). Under our analysis, however, this complex misapplication pattern is just a combination of the simpler stress misapplication patterns attested in Selayarese (see Section 2.2) and Scottish Gaelic (Section 3.1).

In Selayarese, HE-IDENT[stress] is satisfied by stressing neither the copy vowel nor the host; in Scottish Gaelic, HE-IDENT[stress] is satisfied by stressing both the copy and the host. Ho-Chunk utilizes both strategies, distributed according to their interaction with other independently necessary metrical constraints.

4.4 Length matching, neutralization, and back-copying

The previous subsection argued that, in a Dorsey's Law sequence, the copy vowel and its host desire identity for stress. In this subsection, we consider evidence that a copy vowel and its host desire identity for length, or moraic count, as well. We identify two ways in which the normal durational patterns in Ho-Chunk are disrupted by Dorsey's Law, and argue that both should be viewed as instances of *back-copying*: restrictions on the maximum length of an epenthetic vowel hold not only over the copy vowel, but are transferred to its host as well.

As far as we are aware, these effects are the first reported examples of back-copying in copy epenthesis (cf. Kitto and de Lacy 1999:16). Insofar as the existence of back-copying provides evidence for the activity of correspondence constraints that hold over surface forms (see McCarthy and Prince 1995, esp. pp. 24–29), the existence of these phenomena underscores the major claim of our paper: any successful analysis of copy epenthesis must involve a correspondence relation.

4.4.1 Length matching

Many authors working with primary Ho-Chunk data have reported that one or both vowels in the DL sequences are noticeably short. For example, Susman (1943:9–10) writes that "…[DL sequences] can be identified […] usually by the fact that the vowels are very short"; Miner (1979:26) notes that "the sequences are spoken […] faster than other CVCV sequences"; and Hale and White Eagle (1980:117) note that the epenthetic vowel exhibits "extra brevity" (they do not remark on the length of the host vowel). From these descriptions, we infer that both the copy and the host vowel are shorter than the target duration of a normal short vowel (that is, one not in a DL sequence).

For an analysis, we propose that Ho-Chunk vowels belong to one of three durational categories: long, or bimoraic ($V_{\mu\mu}$); short, or monomoraic (V_{μ}); and overshort, or some quantity less than a mora (which we notate as V_{ι}).[15] In the general case, a ban on overshort vowels ($*V_{\iota}$, (64)) limits the vocalic inventory to short and long vowels. In contexts of epenthesis, however, overshort vowels arise due to a desire to minimize the length of an epenthetic vowel: the shorter the epenthetic vowel, the more minimal the perceptual change between the unrepaired and repaired clusters (see Steriade 2009 on the P-map). We formalize this desire for short epenthetic vowels as DEP-$V_{\mu+}$ (65).

(64) $*V_{\iota}$: Assign one violation mark * for each vowel that does not bear at least one mora.

(65) DEP-$V_{\mu+}$: Assign one violation mark * for every output vowel associated with one or more moras that lacks an input correspondent.

If both DEP-$V_{\mu+}$ and HE-IDENT[length] – the constraint ensuring copy-host length identity defined in (41) above in the analysis of Scottish Gaelic – dominate $*V_{\iota}$, the following situation arises. The epenthetic vowel is prevented from being associated with one or more moras, as inserting such a

[15] To be clear, we assume that the unit for stress assignment is the set of $\{\iota,\mu\}$. The difference in notation reflects a difference in duration and not a difference in stressability.

vowel would fatally violate DEP-$V_{\mu+}$ (66a). It is also impossible to insert an overshort copy vowel of a host which surfaces with one mora (i.e. the normal mora count for a short vowel), as a fatal violation of HE-IDENT[length] results (66b). Thus, the optimal choice in this situation is to shorten the host vowel (violating MAX-μ, as defined in (34) above in the analysis of Scottish Gaelic) and realize both vowels as overshort, despite the two violations of low-ranked *V_ι (66c). (Stress marks are omitted from tableaux from this point forward.) Under this analysis, the length-matching effect in Ho-Chunk is an example of back-copying: restrictions on the maximum length of the copy cause the length of the host to be altered, as well.

(66) Length-matching in Ho-Chunk

/kre$_\mu$/	HE-IDENT[length]	DEP-$V_{\mu+}$	MAX-μ	*V_ι
a. k$\underline{e}_\mu$re$_\mu$		*!		
b. k$\underline{e}_\iota$re$_\mu$	*!			*
☞ c. k$\underline{e}_\iota$re$_\iota$			*	**

A reviewer, picking up on the importance of Steriade's (2009) P-map to this analysis, asks how the existence of copy epenthesis is consistent with its assumptions. Steriade (2009:174–175) notes that the cross-linguistic preference for [ə] as a default epenthetic vowel (also Kitto and de Lacy 1999:4) follows from the assumption that phonological grammars prefer minimal input-output modification, as the insertion of a short, contextually variable vowel is 'the closest thing to no epenthesis at all'. Copy epenthesis by definition does not involve the insertion of a short default vowel, and when the P-map is taken into account, would appear to be a substandard repair strategy.

Under our analysis, the existence of copy epenthesis arises due to the ability of HE-IDENT[F] constraints to outrank IO-Correspondence constraints, i.e. those constraints that would prefer insertion of short [ə] to longer [a] (see Section 1.1 for discussion). We adopt, with Steriade, the hypothesis that IO-Correspondence constraints are organized so as to prioritize the least salient repair. But there is no reason that these constraints, as a unit, cannot be subordinated to HE-IDENT[F]. Put differently, the P-map is not a theory of the overall structure of phonological grammars, but rather a theory of the relationships among correspondence constraints along a single correspondence dimension. As HE-IDENT[F] constraints are not IO-Correspondence constraints, the relative ranking of these two different classes of faithfulness constraints can vary on a language-by-language basis.[16]

4.4.2 Length neutralization in DL

Recall that Ho-Chunk admits both short and long vowels. Available evidence suggests, however, that this contrast is neutralized in DL sequences: no DL sequence that we have found contains a long host. Since the ban on voiceless stop + sonorant clusters is surface-true, we find it unlikely that the limitation to short hosts is due to blockage of epenthesis when the potential host vowel is long: if this were the case, we would expect to find voiceless stop + sonorant clusters licensed when a long vowel follows (e.g.*[kreː]); this would be the Ho-Chunk equivalent of the Scottish Gaelic blocking pattern (see Section 3). Given that words of this shape are unattested, it is more likely that DL applies regardless of the underlying length of the host. Assuming richness of the base (Prince and Smolensky 2004), an underlyingly long vowel must then shorten when it functions as a host.

[16]For now, we leave aside discussion of cases in which copy epenthesis occurs across some consonants and default epenthesis occurs across others; see e.g. Kawahara (2007) on Japanese. The fact that both copy and default epenthesis are consistent with the P-map suggests that systems employing both repairs necessarily are, too.

Shortening of long hosts is predicted by the length-matching analysis above. If the copy vowel is forbidden from being associated with a full mora (DEP-V$_{\mu+}$ is undominated), and both host and copy must be associated with the same number of moras (HE-IDENT[length] is undominated), the host vowel must shorten to match its copy – even if that host is underlyingly long (67).

(67) Shortening of long hosts (hypothetical example)

/kre$_{\mu\mu}$/	HE-IDENT[length]	DEP-V$_{\mu+}$	MAX-μ	*V$_\iota$
a. ke̲$_{\mu\mu}$re$_{\mu\mu}$		*!		
b. ke̲$_\iota$re$_{\mu\mu}$	*!			*
☞ c. ke̲$_\iota$re$_\iota$			**	**

Is there evidence from alternations to support the prediction that underlyingly long host vowels shorten? Potentially. There are a number of roots of the shape RVː(C) (where R = sonorant, (C) = optional consonant); for example, *reː* 'to go'. When the second person singular marker *ʃ-* is prefixed to roots of this shape, DL is triggered, and the root's long vowel shortens: for example, *ʃe̲re* 'you (sg.) go' (Miner 1989:151). While these alternations could be interpreted as HE-IDENT[length]-motivated neutralization, there are other potential explanations for why the root vowel should shorten. For example, a preliminary examination of Miner (1984) suggests that inflectional prefixes may more generally cause a root-initial vowel to shorten (e.g. *kũːwá̃k* 'to dive' → *ha-kũwá̃k* 'I dive', Miner 1984: ex. 1925; *hóː* 'to howl' → *ha-hó* 'I howl', Miner 1984: ex. 1428).

Regardless of whether or not alternations like *reː* ↔ *ʃe̲re* provide evidence for the activity of HE-IDENT[length], the lack of long hosts in DL is an exceptionless phonotactic generalization, and as such should receive some analysis. Taking into account the length-matching data discussed above, the generalization is that a host, no matter its underlying length, shortens to match its copy.

4.5 Other evidence for HE-faithfulness in Ho-Chunk

Beyond stress and length misapplication, evidence consistent with the activity of HE-IDENT constraints comes from two additional phenomena: overapplication of predictable vowel nasalization and overapplication of ablaut (see also Kitto and de Lacy 1999:14–15, 22 on both). While the main focus of this paper is on misapplication effects that target *prosodic* properties, these phenomena involve the misapplication of *segmental* processes. Together, these two kinds of phenomena suggest that copy-host pairs in Ho-Chunk strive for identity in all predictable phonological properties.[17]

4.5.1 Overapplication of vowel nasalization

In Ho-Chunk, vowel nasality is contrastive for [+low] and [+high] vowels (68). In post-nasal position, however, only nasal vowels are permitted to occur (69).

(68) Minimal pairs demonstrating vocalic nasality contrast (Miner 1989:149)

a.	síː	'foot'	*vs.*	sí̃ː	'liver'
b.	háːk	'rear part'	*vs.*	há̃ːk	'woodchuck'
c.	gisú	'to husk'	*vs.*	gisú̃	'upset'

[17] An additional fact often cited in this context has to do with reduplication: Ho-Chunk reduplicants are generally monosyllabic (e.g. *gihu* 'to swing' → *gihu-hu* 'to wag its tail', Miner 1989:149), but DL sequences can exceptionally reduplicate in full (e.g. *ʃa̲ra* 'bald' → *ʃa̲ra-ʃa̲ra* 'bald in spots', Miner 1989:146). However, the transcription of stress in these forms is inconsistent across different works, so we refrain from making any claims regarding these data.

(69) Only nasal vowels permitted in post-nasal position (Miner 1989:149)
  a. mã́ː ‘earth, ground’ (*maː)
  b. nĩ́ː ‘water’ (*niː)
  c. wamãnṹke ‘thief’ (*wamanuke)

Following Kitto and de Lacy (1999:14–15), we model this pattern with three constraints: *NV, which prevents oral vowels from following nasal consonants (70); *Ṽ, which bans nasal vowels (71); and IO-Ident[±nasal], which demands faithfulness to input values for [±nasal] (72).

(70) *NV: Assign one violation mark * for each oral vowel which follows a nasal consonant.

(71) *Ṽ: Assign one violation mark * for each nasal vowel.

(72) IO-Ident[±nasal]: Assign one violation mark * for each IO pair disagreeing in [±nasal].

As a contrast in vocalic nasality is generally licensed in Ho-Chunk, we know that IO-Ident[±nasal] must dominate *Ṽ. To derive the generalization that oral vowels are forbidden following nasal consonants, it must be the case that *NV dominates IO-Ident[±nasal]. Below, we show that this ranking induces neutralization in post-nasal position (73) but generates contrast elsewhere (74).

(73) Contrast for [±nasal] neutralized in post-nasal position

	/maː/	*NV	IO-Ident[±nasal]	*Ṽ
☞	a. [mãː]		*	*
	b. [maː]	*!		
	/mãː/	*NV	IO-Ident[±nasal]	*Ṽ
☞	c. [mãː]			*
	d. [maː]	*!	*	

(74) Contrast for [±nasal] licensed elsewhere

	/gisu/	*NV	IO-Ident[±nasal]	*Ṽ
	a. [gisũ]		*!	*
☞	b. [gisu]			
	/gisũ/	*NV	IO-Ident[±nasal]	*Ṽ
☞	c. [gisũ]			*
	d. [gisu]		*!	

The one place where this distribution is not observed is in DL sequences. As demonstrated in (75), copy vowels always agree for [±nasal] with their hosts. If the host vowel follows a nasal consonant, both copy and host are always nasal, despite the fact that the copy vowel resides in an environment where both oral and nasal vowels are generally possible.

(75) Neutralization of nasality contrasts for copy vowels (Miner 1989:149)
  a. /sni/ → [sĩ̲nĩ] (*[si̲nĩ], *[si̲ni]) ‘cold’
  b. /knak/ → [kã̲nãk] (*[ka̲nãk], *[ka̲nak]) ‘to marry’
  c. /boːpnus/ → [boːpũ̲nũs] (*[boːpu̲nũs], *[boːpu̲nus]) ‘to hit at random’

As shown above, nasal vowels are generally permitted due to the ranking IO-IDENT[±nasal] ≫ *Ṽ. However, given our assumption that the epenthetic copy vowel lacks an input correspondent, IO-IDENT[±nasal] is irrelevant, and *Ṽ should force all copy vowels to surface as [-nasal] (since copy vowels only ever surface after non-nasal stops). If, however, HE-IDENT[±nasal] (76) also outranks *Ṽ (again following Kitto and de Lacy 1999:14–15), the copy is forced to resemble its predictably nasalized host, and the unexpected distribution in (75) is explained. This is illustrated in (77).

(76) HE-IDENT[±nasal]: Assign one violation mark * for each HE pair disagreeing for [±nasal].

(77) Insertion of a nasal copy vowel, to satisfy *NV and HE-IDENT[±nasal]

/knak/	*NV	HE-IDENT[±nasal]	IO-IDENT[±nasal]	*Ṽ
a. [kanak]	*!			
b. [kanãk]		*!	*	*
☞ c. [kãnãk]			*	**

4.5.2 Overapplication of ablaut

Ho-Chunk has a process of final vowel ablaut, in which root-final [e] maps to [a] when immediately followed by a certain kind of suffix. This process is demonstrated in (78), where the vowel that undergoes ablaut is bolded (data from Miner 1989:50). When the word-final [e] belongs to a DL sequence, however, both the copy and the host undergo ablaut, as demonstrated in (79) (data from Miner 1989:150). (Stress marks are omitted here.)

(78) Final [e] → [a] ablaut in Ho-Chunk
a. war**e** → war**a**-re 'work / imperative'
b. t**e**: → t**a**:-nã 'I go / I could go'

(79) DL [e] → [a] ablaut in Ho-Chunk
a. k**ere** → k**ara**-ire 'to leave returning / they leave returning'
b. gis**e**w**e** → gis**a**w**a**-nãk 'to calm down / to calm down sitting'

Kitto and de Lacy (1999:22) note that these data are compatible with a correspondence-based analysis; such an analysis, in sketch form, follows. Let us assume there is some constraint, ABLAUT, that requires word-final [e] to lower to [a] in certain morphological contexts; we do not attempt to offer an analysis of the ablaut pattern itself. In order to take effect, ABLAUT must dominate IO-IDENT[±low]. In words that do not contain a DL sequence, the result is that only the root-final vowel lowers as a result of ablaut (80b). Lowering any other vowels, as in (80c), results in unnecessary violations of IO-IDENT[±low]. (Note that the forms in (80) are hypothetical. Actual Ho-Chunk forms like [texe] are attested, e.g. [xete] 'big, large', Helmbrecht and Lehmann 2010:219, but examples of these stems undergoing ablaut are unattested in the work on Ho-Chunk available to us. We assume from the more general descriptions of ablaut that only the final vowel lowers.)

(80) Ablaut affects only the final vowel in non-DL words

/texe/ + /wi/	ABLAUT	IO-IDENT[±low]
a. [texe-wi]	*!	
☞ b. [texa-wi]		*
c. [taxa-wi]		**!

In words where the final [e] belongs to a DL sequence, however, both the final root vowel (the host) and the penultimate root vowel (the copy) lower. This pattern shows us that a faithfulness constraint enforcing identity for [±low] among copy-host pairs, HE-IDENT[±low], is active.

(81) Ablaut affects the penult and the final in DL words

/giswe/ + /nak/	ABLAUT	HE-IDENT[±low]	IO-IDENT[±low]
a. [gise̲we-nãk]	*!		
b. [gise̲wa-nãk]		*!	*
☞ c. [gisa̲wa-nãk]			*

4.6 Discussion of Ho-Chunk

These vowel length, nasality, and ablaut effects, taken together with the stress misapplication patterns described and analyzed above, provide significant support for the existence of a correspondence relationship between epenthetic vowels and their hosts in Ho-Chunk. In the remainder of this section, we briefly discuss several alternative approaches to the data discussed here.[18]

Past accounts of many of the phenomena in this section have treated the observed identity effects as a consequence of Dorsey's Law applying after other phonological rules. For example, Miner (1979, *et seq.*) shows that the nasalization facts can be derived if progressive vowel nasalization is ordered before vowel copy: given an input like /knak/, vowel nasalization will yield the intermediate form *knãk*, and then a copy of the nasalized vowel is inserted to break up the cluster (derivation in (82)). The ablaut facts are treated similarly: given the underlying form /kre-ire/, ablaut applies to yield intermediate *kra-ire*, and then a copy of the final vowel is inserted to break up the cluster (83).

(82) Derivational account of vowel nasalization in DL sequences

a.	Underlying Form		/knak/
	Rule (i)	*Nasalization*	*knãk*
	Rule (ii)	*Dorsey's Law*	*kã̲nãk*
b.	Surface Form		[kã̲nãk]

(83) Derivational account of ablaut in DL sequences

a.	Underlying Form		/kre-ire/
	Rule (i)	*Ablaut*	*kra-ire*
	Rule (ii)	*Dorsey's Law*	*ka̲ra-ire*
b.	Surface Form		[ka̲ra-ire]

Miner's (1979) analysis captures the facts, but misses a generalization: in both cases where DL follows another phonological process, the result is unexpected identity between a copy vowel and its host. It is thus an accident, under Miner's analysis, that copy-host pairs converge on total identity.

Hall's (2003) analysis of the Ho-Chunk facts is also subject to this criticism. Under her analysis, copy-host identity in Ho-Chunk emerges because both DL vowels are part of the same segment; however, additional assumptions required to make the observed facts follow appear in some cases to be arbitrary. To analyze the overapplication of vowel nasalization in [kã̲nãk], for example, Hall

[18] We shy away from discussing alternative approaches to Ho-Chunk stress, as they largely assume different facts about stress placement than we have, and it is difficult to compare analyses when they make different assumptions about the description of the facts. (Even Miner 1979, who acknowledges the existence of prominence on both members of the DL sequence, does not present an analysis that is consistent with his description of the data.)

assumes that predictably nasalized vowels are associated with their own velum lowering gesture (an assumption that, as far as we can tell, is not motivated elsewhere). As the two [ã]'s are part of the same segment, and thus linked to the same velum lowering gesture, both are nasalized (pp. 147–148). For the ablaut facts, Hall proposes a similar analysis: because the two [e]'s in [ke̱re] are part of the same segment, both map to [a] when an ablaut-inducing suffix is added (p. 148). Finally, in regard to the stress misapplication facts, Hall proposes that both vowels in a DL sequence bear high pitch because high pitch is linked to segments, and the two DL vowels are part of the same segment (p. 173). We note, however, that in order to square this proposal with the fact that a long vowel bears high pitch for only half of its duration (see Figure 3, also Miner 1989), it is necessary to assume that high pitch is actually linked to the accented *mora*, not the accented *segment*, and that consequently both halves of the DL sequence are linked to a single mora (as both bear high pitch). In addition, to account for the fact that words composed of only a single DL sequence satisfy the bimoraic minimal word requirement, Hall is forced to assume that the sonorant member of the DL sequence also bears a mora; (e.g. the [n] in [kã̱nãk]) is moraic. How exactly it is possible for two halves of the same mora to be split up by a different mora, without violating conditions on line-crossing, is unclear.

Our point is this: existing alternative analyses of the Ho-Chunk data account separately for each misapplication effect, and in doing so, miss the generalization that unites them – all promote identity between a copy vowel and its host. The correspondence-based analysis, by contrast, captures this generalization directly: misapplication occurs to promote identity. In addition, the correspondence-based analysis extends to the length-matching facts in Section 4.4 which, while others have discussed (see e.g. Miner 1979:26, Hall 2003:149), none have analyzed. In particular, it is unclear to us how the lack of DL sequences with long vowels could be accounted for by either of the existing analyses described above.

5 Discussion and Conclusion

We have argued that positing a correspondence relation between a copy vowel and its host yields a natural, unified treatment of prosodic misapplication effects in copy epenthesis. Faithfulness constraints that act on this relation, demanding identity between an epenthetic vowel and its host, can explain otherwise unexpected prosodic patterns in four languages: Selayarese (Section 2), Scottish Gaelic (Section 3), Irish Gaelic (Section 3.3), and Ho-Chunk (Section 4). (See Table 2 for a summary of all cases.) In this way, our results bolster Kitto and de Lacy's (1999) claim that a correspondence relationship holds between an epenthetic vowel and its host. In Section 5.1, we discuss how these results bear on Kawahara's (2007) claim that correspondence-based analyses of copy epenthesis overgenerate. Section 5.2 concludes.

5.1 Discussion of Kawahara (2007)

The claim that a copy vowel and its host correspond with one another has precedent (esp. Kitto and de Lacy 1999), but it is also controversial. In a paper on the similarities and differences between copy epenthesis and reduplication, Kawahara (2007) argues that a number of generalizations underlying the typology of copy epenthesis can only be captured in an analysis that does *not* make use of correspondence. This section discusses Kawahara's generalizations. We first show (Section 5.1.1) that the phenomena discussed in this paper cast doubt on Kawahara's assertion that copy epenthesis

Table 2: Summary of effects observed

Language	*Source(s)*	*Effect*	*Type of prominence*
§2: Selayarese	Mithun & Basri 1986 Kitto & de Lacy 1999 Broselow 2008	stress misapplication	stress/length
§3: Scottish Gaelic (Barra dialect)	Børgstrom 1937, 1940 Clements 1986 Bosch & de Jong 1997 Hall 2003 Hammond et al. 2014	stress matching length matching blocking	stress length
§3.3: Irish Gaelic	Ní Chiosáin 1999	blocking	length
§4: Ho-Chunk	Miner 1979 *et seq.* Hale & White Eagle 1980 Hall 2003	stress misapplication length matching nasality matching ablaut	stress/pitch length

never involves length matching. Following this, we discuss the rest of Kawahara's generalizations, and, where we have such ideas, suggest possible ways forward.

5.1.1 Length-matching

Length transfer effects (Levin 1985), where a short vowel is copied as short and a long vowel copied as long, are well-attested in the typology of reduplication (see (84) for examples from Kihehe). However, Kawahara (2007:16) finds that equivalent patterns in copy epenthesis – such as (85), where a short host begets a short copy and a long host a long copy – are unattested.

(84) Length transfer in Kihehe reduplication (Odden and Odden 1985:500)
  a. kú-ceéng-a → kú-<u>c**ee**nga</u>-ceénga, *<u>c**e**nga</u>-ceénga 'build/a bit'
  b. kú-ceeng-él-a → kú-<u>ceeng**e**la</u>-ceengéla, *<u>ceeng**ee**la</u>-ceengela 'build for/a bit'

(85) Length transfer in copy epenthesis is unattested (Kawahara 2007:16)
  a. /takt/ → [tak<u>a</u>t]
  b. /taakt/ → [taak<u>aa</u>t]

Kawahara shows that a spreading-only analysis of copy epenthesis is unable to generate length transfer patterns like (85), while a correspondence analysis predicts their existence. If it could be shown that patterns like (85) are systematically absent from the typology, this gap would provide a strong argument in favor of spreading-based analyses of copy epenthesis. But given the presence of other varieties of length matching in copy epenthesis (see Section 3 on Scottish Gaelic, and Section 4 in Ho-Chunk), we believe that the lack of patterns exactly equivalent to (85) could be an accidental gap, and therefore does not argue in favor of spreading-based approaches. We illustrate this through factorial typology below.

The analysis required to derive the length transfer effects in Scottish Gaelic (Section 3) and Ho-Chunk (Section 4.5) involve five crucial ingredients, as laid out in (86). Additional factors may be involved in the actual patterns, but these suffice to explore the broad typology.

(86) Components of length transfer analyses

a. Input-Output faithfulness to length (IO-IDENT):
Output vowels must preserve their input correspondents' length specifications.

b. Copy-Host faithfulness to length (HE-IDENT):
Copy vowels and their hosts must have identical length specifications.

c. Cluster markedness constraints (*BADCLUSTER):
Certain cluster types are not permitted in the output.

d. Constraints against long epenthetic vowels (DEP-LONGV):[19]
The epenthetic copy vowel is not permitted to exceed a certain length threshold.
(cf. DEP-$V_{\mu+}$ (65) for Ho-Chunk)

e. Input-Output faithfulness against epenthesis (DEP-SEG)
No epenthetic segments (irrespective of length).

A factorial typology of these five constraints yields five possible systems, summarized in Table 3. (The factorial typology was explored with OTSoft; Hayes et al. 2013). Where applicable, each system type is annotated with an example of a language that instantiates it.

Of the five predicted patterns in Table 3, four are attested. Type (1), in which copy epenthesis does not occur, is attested in many of the world's languages, including American English. Type (2), in which copy epenthesis applies when the host is short but not when the host is long, is attested in Scottish Gaelic (Section 3.2). Type (4), in which only short copy vowels are permitted and underlyingly long vowels shorten to match their copy, is instantiated in the neutralization pattern of Ho-Chunk (Section 4.4.2). And type (5), in which length matching does not occur, is attested in Cook Islands Maori (Kitto and de Lacy 1999:186), among others. The only unattested pattern, then, is type (3), where short vowels are copied as short and long vowels as long (as in (85)).

It is important to note, though, that the incidence of length transfer effects is fairly low. Of the 51 languages surveyed for this paper (see the Appendix), only two (Scottish Gaelic and Ho-Chunk) exhibit them. Given the survey's composition, the fact that type (3) systems are unattested is a fairly probable outcome, as there are three kinds of predicted length transfer effects but only two languages that exhibit such patterns. The fact that types (2) and (4) are attested, but type (3) is not, could thus be due to chance.

In short, we have argued throughout this paper that certain kinds of length transfer effects are attested, and that these attested effects necessitate a correspondence-based theory of copy epenthesis. Although our survey did not uncover the particular kind of length transfer discussed by Kawahara (i.e. the type (3) systems in Table 3), the fact that such cases are unattested is unsurprising given the composition of the current survey, and should be treated as an accidental gap. The result, then, is that the attested typology of length transfer effects is a strong argument *in favor* of a correspondence-based theory of copy epenthesis (contra Kawahara), not an argument against it.

[19] This preference could also be implemented as a markedness constraint that bans long vowels, as in the analysis of Scottish Gaelic (Section 3.2.2). For the purposes of this demonstration, the two options make equivalent predictions.

Table 3: Factorial typology of length matching effects

(1)	**Ranking:**	DEP-LONGV, IO-IDENT, HE-IDENT, DEP-SEG ≫ *BADCLUSTER
	Pattern:	No copy epenthesis
		Short host: /bard/ → [bard] *Long host:* /baːrd/ → [baːrd]
	Example:	American English
(2)	**Ranking:**	DEP-LONGV, IO-IDENT, HE-IDENT ≫ *BADCLUSTER ≫ DEP-SEG
	Pattern:	Copy epenthesis for short hosts only; cluster unrepaired with long hosts
		Short host: /bard/ → [barad] *Long host:* /baːrd/ → [baːrd]
	Example:	Scottish Gaelic
(3)	**Ranking:**	IO-IDENT, HE-IDENT, *BADCLUSTER ≫ DEP-LONGV, DEP-SEG
	Pattern:	Canonical length transfer: short copied as short, long copied as long
		Short host: /bard/ → [barad] *Long host:* /baːrd/ → [baːraːd]
	Example:	**unattested**
(4)	**Ranking:**	DEP-LONGV, HE-IDENT, *BADCLUSTER ≫ IO-IDENT, DEP-SEG
	Pattern:	Copy epenthesis always yields short copy; long hosts shorten to match copy
		Short host: /bard/ → [barad] *Long host:* /baːrd/ → [barad]
	Example:	Ho-Chunk
(5)	**Ranking:**	DEP-LONGV, IO-IDENT, *BADCLUSTER ≫ HE-IDENT, DEP-SEG
	Pattern:	Copy epenthesis always yields short copy; length of host does not change
		Short host: /bard/ → [barad] *Long host:* /baːrd/ → [baːrad]
	Example:	Cook Islands Maori (Kitto and de Lacy 1999:186)

5.1.2 Other arguments

In addition to the generalization that copy epenthesis never induces length transfer, Kawahara (2007) presents three further generalizations that appear to elude explanation under a correspondence-based theory of copy epenthesis. We discuss each in turn.

Lack of consonant copy

Kawahara (2007:8) asserts that one characteristic of copy epenthesis is that it only targets vowels. For example, while copy vowel epenthesis often occurs to break up an illicit consonant cluster, patterns in which copy consonant epenthesis is employed to satisfy some other phonotactic – for example, ONSET, as in the hypothetical pattern in (87) – are claimed to be unattested.

(87) Consonant copy in epenthesis: unattested (example from Kawahara 2007:9)
  a. /apa/ → [papa]
  b. /ata/ → [tata]
  c. /aka/ → [kaka]

Under a correspondence-based account, there is no obvious way to account for this generalization, as HE-IDENT constraints should be expected to hold over both consonants and vowels. (For an explanation of why patterns like (87) are excluded under a spreading-based analysis, the reader is directed to Kawahara 2007:11–14.) We have no solution to this problem, except to deny that it exists: we believe that the descriptive generalization is not as strong as Kawahara claims. For several potential cases of consonant copy, see Yu (2005), Inkelas and Zoll (2005: Ch. 1).[20]

Lack of non-local copy

A further generalization that unites the known cases of copy epenthesis is that the copy and its host are always tier-adjacent. Kawahara (2007:13–14) notes that under a correspondence-based analysis, however, it is possible to derive patterns in which the copy and host are more distant. Let us assume, following Kitto and de Lacy (1999), that adjacency between a copy vowel and its host is enforced by the constraint HE-ADJACENCY (in their terms, BE-ADJACENCY; 'assign one violation mark * for each segment intervening between an epenthetic segment and its host'). But if a language prefers more sonorous nuclei (cf. HNUC; Prince and Smolensky 2004), for example, the markedness constraints that enforce this preference could dominate HE-ADJACENCY, resulting in a pattern like (88) – where the epenthetic vowel chooses to correspond with a more distant, but more sonorous, host (88b–c). (The reader is referred to Kawahara 2007:15 for a demonstration.)

(88) Unattested pattern of sonority-sensitive copying (after Kawahara 2007:14)
  a. /tematk/ → [tematak]
  b. /tametk/ → [tametak]
  c. /temitk/ → [temitek]

We have found no pattern that resembles (88). And while it is possible to exclude (88) and other similar patterns by stipulating that HE-ADJACENCY must dominate all markedness constraints (following a suggestion by Kitto and de Lacy 1999:20–21, and assumed throughout this paper), we find this solution unsatisfying. At present, then, we agree with Kawahara (2007) that the lack of non-local copy is problematic for a correspondence-based theory of copy epenthesis.

Segmental blocking

The final argument offered by Kawahara (2007) against correspondence-based theories is the existence of blocking effects, where copy epenthesis only applies across certain types of consonants. In Japanese, for example, copy epenthesis applies across laryngeal [h] (e.g. [bahha] 'Bach', [mazohho] 'Masoch'); but across supralaryngeal consonants a default vowel (its identity to some extent dependent on the preceding consonant) is inserted instead (e.g. [bazɯ] 'buzz', [sɯpɯraito] 'Sprite'; Kawahara 2007:19 for all forms). While there has not yet been a thorough study of the typology of interveners in copy epenthesis, it appears that most languages that exhibit such restrictions tend to restrict copying to cross-sonorant or cross-laryngeal contexts (Hall 2003, Kawahara 2007:20).[21] Under a spreading-based analysis, accounting for blocking effects is straightforward if we assume

[20] A reviewer asks whether or not, in cases of consonant copy, the inserted segment and its host stand in correspondence with one another. We assume they do, and that the absence of attested misapplication effects is an accidental gap: consonant copy is rarer than vowel copy, and misapplication effects of the sort discussed here are uncommon.

[21] With the exception of Selayarese, which allows copying only across coronal continuants. As discussed in Section 2.2, however, epenthesis in Selayarese should be seen as a last-resort strategy that occurs only when the word-final consonant cannot be modified. It thus does not constitute an example of segmental blocking.

that feature spreading is strictly local (e.g. Ní Chiosáin and Padgett 2001): blocking occurs when an intervening consonant is unable to host the features in transit from the host vowel to an epenthetic vowel (see Kawahara 2007:21–24 for further discussion and analysis). Under a correspondence-based analysis, however, such effects are not easy to capture, as in the general case whether or not α can correspond with γ typically does not depend on whether or not some segment β intervenes. We grant, then, that segmental blocking effects are currently a problem for correspondence-based analyses of copy epenthesis. We sketch the beginning of a possible solution below.

Some current work in the Agreement by Correspondence literature (see esp. Rose and Walker 2004, Bennett 2015b, Shih and Inkelas 2015), proposes that there are adjacency requirements on correspondence. Bennett's (2015b:61) CC·SYLLADJ, for example, requires that corresponding consonants must inhabit neighboring syllables (see also Rose & Walker's 2004 PROXIMITY): if some segment α residing in some syllable σ_a stands in correspondence with some segment γ in σ_c, there must be some some segment β that stands in correspondence with both α and γ in each syllable between σ_a and σ_c. Let us assume that it is possible to implement a version of this constraint, HE·SEGADJ, that refers to adjacency among host-epenthetic pairs (as in (89)).

(89) HE·SEGADJ: If some epenthetic segment α stands in correspondence with a host γ, then every segment β that intervenes between α and γ must stand in correspondence with α and γ. Assign one violation mark * if this condition is not met.

When this constraint is high-ranked, the implication is the following: if an epenthetic vowel and its host stand in correspondence, the consonant that intervenes between the two must stand in correspondence with them as well. It could then be possible to limit the consonantal contexts in which copy epenthesis applies by placing restrictions on the kinds of consonants that can stand in correspondence with vowels. The constraint HE·IDENT[±sonorant], for example, requires corresponding segments to agree for [±sonorant], and would penalize a [-sonorant] consonant standing in correspondence with a [+sonorant] vowel; if satisfaction of HE·IDENT[±sonorant] is paramount, then copy epenthesis across an obstruent would be dispreferred.

Although this proposal provides a possible way forward, it raises questions about other aspects of the analysis, namely the assumption that host-epenthetic correspondence is mandatory (see Section 1.1). In cases like Japanese, the fact that default epenthesis occurs when copy epenthesis is blocked suggests that one way of avoiding violation of HE·SEGADJ is to avoid correspondence altogether; it is difficult to explain, under other assumptions, why an epenthetic vowel should featurally match its host across some consonants but not others. And as noted by a reviewer, the above proposal also raises questions about the activity of similar constraints in reduplication: if HE·SEGADJ has a Base-Reduplicant analog, we would wrongly expect to find segmental blocking effects in reduplication (see Kawahara 2007 for discussion). For now, we leave these questions open. Our intention here is only to show that, although the issue of blocking in copy epenthesis is a problem for current correspondence-based theories, the problem is likely solvable.

5.2 Conclusions

In sum, we have argued from the typology of prosodic misapplication effects that a correspondence relationship is a necessary component of any successful analysis of copy epenthesis (following Kitto and de Lacy 1999, *contra* Kawahara 2007). While the relevance of some of the phenomena discussed here to correspondence-based theories of copy epenthesis has been previously noted (see esp. Kitto and de Lacy 1999), this paper goes beyond the previous contributions in that it provides a unified analysis of all attested misapplication effects.

We do not pretend, however, that this paper ends the debate on spreading- vs. correspondence-based theories of copy epenthesis, as there are still several domains in which the spreading-based theory makes more accurately restrictive predictions than does the correspondence-based theory (see Section 5.1.2, as well as Kawahara 2007 for more in-depth discussion). In addition, we do not pretend to have a case for the appropriateness of a correspondence-only analysis of copy epenthesis, as the fact that a correspondence relation links epenthetic segments and their hosts does not mean that other kinds of relations (e.g. feature spreading) cannot link them as well. All we hope to have shown here is that a correspondence relation is a necessary component of any descriptively adequate theory of copy epenthesis: a spreading-only analysis cannot account for the prosodic misapplication effects documented in this paper. Having established this, we hope that future work can explore how the correspondence-based theory developed here and in Kitto and de Lacy (1999) can be developed to yield sufficiently restrictive typological predictions in all relevant domains.

References

Akinlabi, Akin. 1993. Underspecification and the phonology of Yoruba /r/. *Linguistic Inquiry* 24: 152–156.

Alber, Birgit, and Ingo Plag. 2001. Epenthesis, deletion, and the emergence of the optimal syllable in Creole: The case of Sranan. *Lingua* 111: 811–840.

Alderete, John. 1995. Winnebago accent and Dorsey's Law. In *UMOP 18: Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 21–52. Amherst, MA: GLSA Publications.

Awbery, Gwenllian. 2009. Welsh. In *The Celtic languages*, eds. Martin J. Ball and Nicole Müller. *Routledge Language Family Descriptions*, 359–426. London/New York: Routledge.

Baal, Berit Anne Bals, David Odden, and Curt Rice. 2012. An analysis of North Saami gradation. *Phonology* 29: 165–212.

Beckman, Jill N. 1998. Positional Faithfulness. PhD diss, UMass Amherst.

Bender, Bryon. 1968. Marshallese phonology. *Oceanic Linguistics* 8: 16–35.

Bennett, William G. 2015a. Assimilation, dissimilation, and surface correspondence in Sundanese. *Natural Language and Linguistic Theory* 33: 371–415.

Bennett, Wm. G. 2015b. *The Phonology of Consonants: Harmony, Dissimilation, and Correspondence*. Cambridge: Cambridge University Press.

Bhaskararao, Peri. 1998. Gadaba. In *The Dravidian Languages*, 328–358. New York: Routledge.

Bickmore, Lee S. 1995. Refining and Formalizing the Tahitian Stress Placement Algorithm. *Oceanic Linguistics* 34: 410–442.

Blevins, Juliette. 2003. Antigemination: Natural or Unnatural History? In *Proceedings of the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations*, 499–522.

Børgstrom, C. H. 1937. The dialect of Barra in the Outer Hebrides. *Norsk tidsskift for sprogvidenskap* 8: 71–242.

Børgstrom, C. H. 1940. *A linguistic survey of the Gaelic dialects of Scotland, vol 1: the dialects of the Outer Hebrides*. Oslo: Norwegian Universities Press.

Bosch, Anna, and Kenneth de Jong. 1997. The prosody of Barra Gaelic epenthetic vowels. *Studies in the linguistic sciences* 27: 2–15.

Bowern, Claire. 2012. *A Grammar of Bardi*. Mouton de Gruyter.

Brasington, Ron. 1978. Vowel epenthesis in Rennellese and its general implications. In *Work in Progress 2, Phonetics Laboratory, University of Reading*.

Broselow, Ellen. 2000. Stress, Epenthesis and Segment Transformation in Selayarese Loans. In *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, 311–325. Washington, DC: Linguistic Society of America.

Broselow, Ellen. 2008. Stress-epenthesis interactions. In *Rules, Constraints, and Phonological Phenomena*, eds. Bert Vaux and Andrew Nevins, 121–148. Oxford University Press.

Browman, Catherine P., and Louis Goldstein. 1986. Towards an Articulatory Phonology. *Phonology Yearbook* 3: 219–252.

Caballero, Gabriela. 2008. Choguita Rarámuri (Tarahumara) Phonology and Morphology. PhD diss, UC Berkeley.

Caballero, Gabriela. 2011. Morphologically conditioned stress assignment in Choguita Rarármuri. *Linguistics* 49: 749–790.

Campbell, Lyle. 1974. Theoretical implications of Kekchi phonology. *International Journal of American Linguistics* 40: 269–278.

Campbell, Lyle. 2013. *Historical Linguistics: An Introduction, Third Edition*. Cambridge, MA: MIT Press.

Clements, Nick. 1986. Syllabification and epenthesis in the barra dialect of gaelic. In *The Phonological Representation of Suprasegmentals*, eds. K. Bogers, H. van der Hulst, and M. Mous, 317–336. Dordrect: Foris Publications.

Clements, Nick. 1991. Place of articulation in consonants and vowels: a unified theory. In *Working Papers of the Cornell Phonetics Laboratory*, Vol. 5, 77–123.

de Lacy, Paul. 2002. The Formal Expression of Markedness. PhD diss, University of Massachusetts, Amherst, MA.

de Lima Silva, Wilson. 2012. A Descriptive Grammar of Desano. PhD diss, University of Utah, Salt Lake City.

Fleischhacker, Heidi Anne. 2005. Similarity in Phonology: Evidence from Reduplication and Loan Adaptation. PhD diss, University of California, Los Angeles.

Furbee-Losee, Louanna. 1974. *The Correct Language: Tojolabal*. New York: Garland.

Gafos, Adamantios, and Linda Lombardi. 1999. Consonant transparency and vowel echo. In *Proceedings of NELS 29*, 81–95.

Garrett, Edward. 1994. Minimal words aren't minimal feet. ROA-1031.

Gordon, Matt. 2002. A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20: 491–552.

Haas, Mary R. 1940. *Tunica*. New York: J. J. Augustin.

Hale, Kenneth, and Josie White Eagle. 1980. A preliminary metrial account of winnebago accent. *International Journal of American Linguistics* 46 (2): 117–132.

Hall, Nancy. 2003. Gestures and Segments: Vowel Intrusion as Overlap. PhD diss, University of Massachusetts, Amherst, MA.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23: 387–429.

Halle, Morris, and Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, MA: MIT Press.

Halle, Morris, Bert Vaux, and Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31: 387–444.

Hammond, Michael, Natasha Warner, Andréa Davis, Andrew Carnie, Diana Archangeli, and Muriel Fisher. 2014. Vowel insertion in Scottish Gaelic. *Phonology* 31: 123–153.

Hansson, Gunnar Ólafur. 2010. *Consonant Harmony: Long-Distance Interaction in Phonology*. Berkeley, CA: University of California Press.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago and London: The University of Chicago Press.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2013. OTSoft 2.3.2, software package. http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Helmbrecht, Johannes, and Christian Lehmann. 2010. *Hocak Teaching Materials, Volume 1: Elements of Grammar/Learner's Dictionary*. Albany, NY: SUNY Press.

Hualde, José Ignacio. 1998. A gap filled: postpeninitial accent in Azkoitia Basque. *Linguistics* 36: 99–117.

Inkelas, Sharon, and Cheryl Zoll. 2005. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University Press.

Jensen, John. 1977. *Yapese Reference Grammar*. Honolulu: University Press of Hawaii.

Ka, Omar. 1994. *Wolof phonology and morphology*. University Press of America.

Kager, René. 1994. Ternary rhythm in alginment theory. Ms., Utrecht University.

Kager, René. 2012. Stress in windows: Language typology and factorial typology. *Lingua* 122: 1454–1493.

Kawahara, Shigeto. 2007. Copying and spreading in phonological theory: Evidence from echo epenthesis. In *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory III*, eds. Leah Bateman, Michael O'Keefe, Ehren Reilly, and Adam Werle. Amherst, MA: GLSA Publications.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, Michael. 2001. The role of perception in loanword phonology: A review of Les emprunts linguistiques d'origine européenne en fon by flavien gbéto. *Linguistique Africaine*.

Kenstowicz, Michael. 2007. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences* 29: 316–340.

Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology*. San Diego, CA: Academic Press.

Kim, Yuni. 2008. Topics in the Phonology and Morphology of San Francisco del Mar Huave. PhD diss, UC Berkeley.

Kitto, Catherine. 1997. Epenthesis and Polynesian Loanwords. Master's thesis, The University of Auckland.

Kitto, Catherine, and Paul de Lacy. 1999. Correspondence and epenthetic quality. *Proceedings of AFLA* 4: 181–200.

Kumagai, Gakuji. 2016. Resolving the Issue of the Target of Vowel Copy in Fijian Loanwords. In *Proceedings of the 2015 Annual Meeting on Phonology*, eds. Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin, and Douglas Pulleyblank. Washington, DC: Linguistic Society of America.

Levin, Juliette. 1985. A metrical theory of syllabicity. PhD diss, Massachusetts Institute of Technology, Cambridge, MA.

Lopes, Aurise Brandao, and Steve Parker. 1999. Aspects of Yuhup phonology. *International Journal of American Linguistics* 65: 324–342.

Lynch, John. 1978. *A Grammar of Lenakel*. Canberra: Australian National University.

McCarthy, John J. 1994. The Phonetics and Phonology of Semitic Pharyngeals. In *Phonological Structure and Phonetic Form - Papers in Laboratory Phonology III*, ed. Patricia Keating, 191–233. Cambridge: Cambridge University Press.

McCarthy, John J. 1998. Morpheme structure constraints and paradigm occultation. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society 32*, 123–150. Chicago: Chicago Linguistic Society.

McCarthy, John J., and Alan Prince. 1988. Quantitative transfer in reduplicative and templatic morphology. In *Linguistics in the Morning Calm 2*, ed. The Linguistic Society of Korea. Hanshin Publishing Co..

McCarthy, John J., and Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. Available online at https://rucore.libraries.rutgers.edu/rutgers-lib/41845/.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 249–384. Amherst, MA: GLSA.

Metcalfe, C. D. 1975. *Bardi Verb Morphology*. Canberra: Australian National University.

Miller, Marion. 1999. *Desano Grammar*. Arlington, TX: SIL International and University of Texas.

Mills, Roger F., and John Grima. 1980. Historical developments in Lettinese. In *Austronesian studies, Papers from the Second Eastern Conference on Austronesian Languages*, ed. Paz Buenaventura, 273–283. University of Michigan, Center for South and Southeast Asian Studies.

Miner, Kenneth. 1984. *Winnebago field lexicon*. University of Kansas.

Miner, Kenneth L. 1979. Dorsey's law in winnebago-chiwere and winnebago accent. *International Journal of American Linguistics* 45 (1): 25–33.

Miner, Kenneth L. 1981. Metrics, or Winnebago Made Harder. *International Journal of American Linguistics* 47: 340–342.

Miner, Kenneth L. 1989. Winnebago accent: The rest of the data. *Anthropological Linguistics* 31: 148–172.

Mithun, Marianne, and Hasan Basri. 1986. The Phonology of Selayarese. *Oceanic Linguistics* 25: 210–254.

Mous, Maarten. 2004. *A grammar of Iraqw*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Myers, Scott. 1987. Vowel shortening in english. *Natural Language and Linguistic Theory* 5: 485–518.

Nelson, Nicole. 2003. Asymmetric Anchoring. PhD diss, Rutgers University.

Ní Chiosáin, Máire. 1999. Syllables and phonotactics in Irish. In *The syllable: Views and Facts*, 551–575. Berlin / New York: Walter de Gruyter.

Ní Chiosáin, Máire, and Jaye Padgett. 2001. Markedness, segment realization, and locality in spreading. In *Segmental phonology in Optimality Theory: constraints and representations*, ed. Linda Lombardi, 118–156. Cambridge: Cambridge University Press.

Nichols, Lynn. 1994. Vowel copy and stress in Northern Tiwa. In *Harvard Working Papers in Linguistics*, 133–140. Harvard University.

Odden, David, and Mary Odden. 1985. Ordered reduplication in Kihehe. *Linguistic Inquiry* 16: 497–503.

Olson, Kenneth S. 2005. *The Phonology of Mono*. Dallas, TX: SIL International.

Paradis, Carole. 1996. The inadequacy of filters and faithfulness in loanword adaptation. In *Current Trends in Phonology: Models and Methods*, eds. Jacques Durand and Bernard Laks. Salford: University of Salford Publications.

Paradis, Carole, and Jean-François Prunet. 1989. On coronal transparency. *Phonology* 6: 317–348.

Parker, Steve. 2001. Non-optimal onsets in Chamicuro: an inventory maximised in coda position. *Phonology* 18: 361–386.

Piggott, Glyne L. 1995. Epenthesis and syllable weight. *Natural Language and Linguistic Theory* 13: 283–326.

Postal, Paul M. 1969. Mohawk Vowel Doubling. *International Journal of American Linguistics* 35: 291–298.

Prince, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11: 511–562.

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

Prince, Alan S. 1983. Relating to the grid. *Linguistic Inquiry* 14 (1): 19–100.

Rehg, Kenneth L., and Damian Sohl. 1981. *Ponapean Reference Grammar*. Honolulu: University Press of Hawaii.

Riad, Tomas. 1992. Structures in Germanic prosody: a diachronic study with special reference to Nordic languages. PhD diss, Stockholm University, Stockholm.

Rose, Sharon. 1996. Variable laryngeals and vowel lowering. *Phonology* 13: 73–117.

Rose, Sharon, and Rachel Walker. 2004. A Typology of Consonant Agreement as Correspondence. *Language* 80: 475–531.

Safir, Ken. 1979. Metrical structure in Capanahua. In *MIT Working Papers in Linguistics 1: Papers on Syllable Structure, Metrical Structure and Harmony Processes*, 95–114.

Shademan, Shabnam. 2002. Epenthetic Vowel Harmony in Farsi. Master's thesis, University of California, Los Angeles.

Shih, Stephanie, and Sharon Inkelas. 2015. Autosegmental Aims in Surface Optimizing Phonology. Ms., UC Merced and UC Berkeley.

Stanton, Juliet, and Sam Zukoff. 2016. Prosodic effects of segmental correspondence. In *Proceedings of CLS 51 (2015)*, 501–515. Chicago: Chicago Linguistic Society.

Steriade, Donca. 1987. Locality Conditions and Feature Geometry. In *Proceedings of the 17th North East Linguistic Society*, eds. Joyce McDonough and Bernadette Plunkett, 595–617. Somerville, MA: Cascadilla.

Steriade, Donca. 1990. Browman and Goldstein's paper. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, eds. J. Kingston and M. E. Beckman, 382–397.

Steriade, Donca. 2009. The Phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The Nature of the Word: Studies in Honor of Paul Kiparsky*, eds. Kristin Hanson and Sharon Inkelas, 151–179. MIT Press.

Susman, Amelia. 1943. The accentual system of Winnebago. PhD diss, Columbia University, New York.

Uffmann, Christian. 2003. A typology of epenthetic vowels in lonwords. Ms., Philipps-Universitat Marburg.

Werle, Adam. 2002. The Southern Wakashan one foot word. In *UBC Working Papers in Linguistics*. Vancouver: University of British Columbia.

Yu, Alan C. L. 2005. Toward a Typology of Compensatory Reduplication. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. John Alderete et al., 397–405. Somerville, MA: Cascadilla Proceedings Project.

Appendix: Copy Epenthesis Survey

The table below summarizes the survey of 51 copy epenthesis processes from which the data and generalizations in this paper are drawn. The survey includes those languages cited in Kawahara's (2007) survey whose sources were available either online or at MIT's Hayden Library, as well as assorted other cases available online. Kawahara distinguishes between cases of "diachronic" and "synchronic" copy epenthesis, and that distinction is recorded here. The term "Complex Misapplication" refers to instances of misapplication that do not receive a straightforward analysis based on the avoidance of stressing epenthetic vowels (e.g., McCarthy and Prince 1994 on Makassarese), or faithfulness to the source stress in the case of loanwords (e.g., Kenstowicz 2007 on Fijian).

Language	*Type*	*Complex Misapplication?*	*Source(s)*
Ho-Chunk	Synchronic	Yes (§2.4)	Miner (1979) *et seq.* Hale and White Eagle (1980) Hall (2003)
Selayarese	Synchronic	Yes (§2.3)	Mithun and Basri (1986) Kitto and de Lacy (1999) Broselow (2008)
Scottish Gaelic (Barra dialect)	Synchronic	Yes (§2.2)	Børgstrom (1937, 1940) Clements (1986) Bosch and de Jong (1997) Hall (2003) Hammond et al. (2014)
Arabic, Bedouin	Synchronic	None mentioned	McCarthy (1994)
Bardi	Synchronic	None mentioned	Metcalfe (1975) Bowern (2012)
Capanahua	Synchronic	None mentioned	Safir (1979)
Chamicuro	Synchronic	None mentioned	Parker (2001)
Desano	Synchronic	None mentioned	Miller (1999) de Lima Silva (2012)
East Cushitic	Diachronic	None mentioned	Blevins (2003)
Farsi	Synchronic	None mentioned	Shademan (2002)
Fijian	Synchronic	None mentioned	Kenstowicz (2007) Kumagai (2016)
Finnish, Eastern	Diachronic	None mentioned	Campbell (2013)
Fula	Synchronic	None mentioned	Paradis (1996) Paradis and Prunet (1989)
Futankoore	Synchronic	None mentioned	Paradis and Prunet (1989)
Gadaba	Synchronic	None mentioned	Bhaskararao (1998)
Hebrew	Synchronic	None mentioned	McCarthy (1994)
Hawaiian	Synchronic	None mentioned	Kitto (1997) Kitto and de Lacy (1999)
Iraqw	Synchronic	None mentioned	Rose (1996) Mous (2004)
Japanese	Synchronic	None mentioned	Kawahara (2007)

Language	*Type*	*Complex Misapplication?*	*Source(s)*
Kekchi	Synchronic	None mentioned	Campbell (1974) Hall (2003)
Kinyarwanda	Synchronic	None mentioned	Uffmann (2003)
Kolami	Synchronic	None mentioned	Clements (1991)
Latin, Late	Diachronic	None mentioned	Steriade (1990)
Lenakel	Synchronic	None mentioned	Lynch (1978)
Lettinese	Diachronic	None mentioned	Mills and Grima (1980)
Maga Rukai	Synchronic	None mentioned	de Lacy (2002)
Makah	Synchronic	None mentioned	Werle (2002)
Makassarese	Synchronic	None mentioned	McCarthy and Prince (1994) McCarthy (1998)
Maori	Synchronic	None mentioned	Kitto (1997) Kitto and de Lacy (1999)
Marshallese	Synchronic	None mentioned	Bender (1968)
Mawu	Synchronic	None mentioned	Kenstowicz (2001)
Moa	Diachronic	None mentioned	Mills and Grima (1980)
Mohawk	Synchronic	None mentioned	Postal (1969)
Mono	Synchronic	None mentioned	Olson (2005)
Northern Tiwa (Picuris, Taos)	Synchronic	None mentioned	Nichols (1994)
Ponapean	Synchronic	None mentioned	Rehg and Sohl (1981)
Rennallese	Synchronic	None mentioned	Brasington (1978)
Sardinian	Diachronic	None mentioned	Steriade (1990)
Shona	Synchronic	None mentioned	Uffmann (2003)
Samoan	Synchronic	None mentioned	Uffmann (2003)
Slavic, Eastern	Diachronic	None mentioned	Steriade (1990)
Somali	Synchronic	None mentioned	Kenstowicz (1994)
Sranan	Synchronic	None mentioned	Uffmann (2003) Alber and Plag (2001)
Tigre	Synchronic	None mentioned	McCarthy (1994) Rose (1996)
Tojolobal	Synchronic	None mentioned	Furbee-Losee (1974) Steriade (1987)
Tunica	Synchronic	None mentioned	Haas (1940)
Welsh	Synchronic	None mentioned	Awbery (2009)
Wolof	Synchronic	None mentioned	Ka (1994)
Yapese	Synchronic	None mentioned	Jensen (1977)
Yoruba	Synchronic	None mentioned	Akinlabi (1993) Nelson (2003)
Yuhup	Synchronic	None mentioned	Lopes and Parker (1999)