To appear in C.Kitto & C.Smallwood (eds.) *Proceedings of AFLA VI*. Holland: Holland Academic Graphics.

Correspondence and epenthetic quality[*,†]

Catherine Kitto
University of Toronto

Paul de Lacy
University of Massachusetts, Amherst

In some languages, epenthetic segments are realized with unmarked features while in others they are copies of nearby segments. To account for this variation, we propose that epenthetic elements can be in a relation of Correspondence with other output segments, analogous to reduplication. This approach is shown to account for both cross-linguistic and language-internal variation in epenthetic quality.

Epenthetic quality – the featural content of epenthetic segments – varies from language to language. In some cases, epenthetic segments are copies of nearby elements, while in others they have default features:

(1) (i) Copy epenthesis:
Selayarese (Western Malayo-Polynesian)[1] – Mithun & Basri (1985)
/sahal/ → [sahala] 'profit'
/potol/ → [potolo] 'pencil'
/lamber/ → [lambere] 'long'

(ii) Default segmentism:
Tongan loanword epenthesis (Polynesian) – Kitto (1997)
[ke:nali] 'colonel'
[telefoni] 'telephone'
[kameli] 'camel'

Copy epenthesis and default segmentism are the two endpoints of a continuum: intermediate cases also exist, where the epenthetic element copies some features and defaults to others. For example, Ponapean epenthetic vowels must be [+high] but otherwise agree with the features of nearby segments (Rehg & Sohl 1981:94). There are analogous examples for consonants: epenthetic consonants are (1) copies of adjacent vowels in Southern Tati (Indo-Iranian) (i.e. they are the vowel's glide counterparts – Yar-

[*] We are particularly grateful to John McCarthy for his comments and advice. We would also like to thank to the organizers of AFLA VI for inviting us to present this paper, and to the audience for their input. Paul de Lacy's work was supported by the National Science Foundation under grant SBR-9420424.

[†] Note for the Rutgers Optimality Archive version: This version differs from the published one in that we have added a number of sections. These sections are marked by a vertical double line running down the left edge of the page.

[1] On first mention of a language, we include its genetic affiliation in brackets.

Shater 1969), (2) partial copies of adjacent vowels in Dakota (Siouan) (Shaw 1980), and (3) default [t] in Asheninca Campa (Arawakan) (Payne 1981).

Epenthetic qualities not only vary across languages – there are cases where there is variation in different contexts within the same language. In addition, cases of copy epenthesis differ in terms of what is copied. In some cases, epenthetic segments copy elements to their left, and in others to their right. There is also variation in terms of proximity – whether the copied element is adjacent to the epenthetic segment or not. A comprehensive theory of epenthetic quality must account for all these phenomena.

One possible way to account for differences in epenthetic quality would be to employ feature sharing: copy epenthesis is where an epenthetic segment shares features with another segment, while default features are inserted if feature sharing is blocked.

The aim of this paper is to explore an alternative to feature sharing: **Correspondence** (McCarthy & Prince 1995). We propose that epenthetic elements can correspond to other output segments, analogous to reduplication. Featural agreement is regulated by constraints that hold over this correspondence relation. The details of this proposal are presented in section 1.

In section 2, constraints that hold between the epenthetic element (hereafter 𝔼) and the element it copies – its Base – are shown to interact with markedness constraints to produce the variety of attested epenthetic qualities. Particular attention is given to cases of contextually-conditioned quality – where epenthetic segments are copies and defaults in different environments within the same language.

The focus of section 3 is on the Base of epenthesis – the elements that epenthetic segments correspond to. Variation in the identity of the Base is shown to result from permuting the rankings of constraints that regulate the correspondence relations between 𝔼 and its Base. We also demonstrate that the Base is determined **dynamically** –the Base of 𝔼 is the 'best' segment for 𝔼 to correspond to, where 'best' is determined by constraints.

Alternative approaches to explaining epenthetic quality are examined in section 4. Particular attention is paid to **feature spreading**. Another correspondence account, employing **breaking**, is also considered. Conclusions and implications for future research are discussed in section 5.

1.0 Correspondence Theory and epenthesis

McCarthy & Prince (1995) propose that a relation called 'correspondence' holds between segments on a number of different dimensions (e.g. input and output, base and reduplicant). Constraints regulate the similarity of corresponding segments. For example, in the hypothetical form [patapata] the segments of the reduplicant [pata] are in correspondence with those in the root [pata]. More precisely, [p] corresponds to [p], [a] to [a], and so forth. Exactness of copying – whether the correspondent of [p] ends up as faithfully copied [p] or unfaithful [t] – is determined by the ranking of constraints such as BR-IDENT-F "Reduplicant segments and their Base correspondents must have identical values for feature F."

Our proposal is that correspondence can hold between an epenthetic element and another output segment, paralleling the situation in reduplication.[2,3] In the Winnebago (Siouan) wordform [ʃa$_1$ra$_1$] 'bald' (Miner 1992:29), for example, the epenthetic segment corresponds to the nearest vowel to its right, as indicated by subscript $_1$'s. There can be as many correspondence relations as epenthetic segments in a form. Epenthetic segments can even copy other epenthetic elements. This is illustrated in the Bardi form below, with correspondence relations indicated by arrows:

(2) /i+ŋ+m+ŋurid+i+inđi+na/ → [iŋumuŋuridiinđina] 'he painted himself'
(Metcalfe 1975:150)

Just as in reduplication, constraints regulate featural similarity between Ɛ and its Base, most importantly BƐ-IDENT-F "Ɛ and its Base have identical values for feature F". BƐ-IDENT-F interacts with markedness constraints to produce the spectrum of attested epenthetic qualities (discussed in detail in section 2).

We propose that there are no restrictions on the positing of correspondence relations. GEN produces many different candidates, all differing simply in correspondence. For example, among its many different outputs, an Input /αβγ/ produces the following set of candidates {[α$_1$βγƐ$_1$], [αβ$_1$γƐ$_1$], [αβγ$_1$Ɛ$_1$], [αβγƐ]}. The only difference between these forms is the element that Ɛ corresponds to. Note that Ɛ need not correspond to anything at all, as in the last candidate. The significance of this possibility is discussed in §3.1.2.

This proposal raises the issue of how the placement of correspondence relations is evaluated. More concretely, what decides that the epenthetic vowel [u] in (2) should correspond to the nearest vowel to its right? Why can it not correspond to the vowel to its left, or to an adjacent consonant?

The answer is that the identity of the Base is determined by constraints: the Base of Ɛ is the 'best' segment available, where 'best' is determined by constraint interaction. This 'dynamically-determined' method of identifying the Base is discussed in section 3.

Before examining the typological implications of using correspondence theory for epenthesis, we pause to note the similarity between the formal mechanism proposed here and the one used for reduplication (McCarthy & Prince 1995). This similarity predicts that many of the processes and qualities found in reduplication will also be found in epenthesis. This is indeed the case, and we will note these parallels when they arise.

2.0 A typology of epenthetic quality

Correspondence constraints and markedness constraints are often in conflict since the former require marked structures to be retained while the latter militate against them. For

[2] We reject proposals that epenthetic quality is determined by either phonetic interpolation or post-phonological fill-in rules. These approaches predict that phonological processes cannot affect epenthetic quality (since epenthetic quality is an entirely post-phonological affair). However, there are a significant number of counter-examples to this: e.g. in Asheninca Campa epenthetic [t] undergoes a regular process of palatalisation before [i] (e.g. /wai+ia/ → [waiṯia] → [waiči̱a]) (Payne 1981). For other examples and relevant discussion see Akinlabi (1993) and Steriade (1995).

[3] The present proposals are a significant extension of a theory originally proposed in Kitto (1997).

epenthesis, this conflict is between BE-correspondence constraints – those constraints that hold over the correspondence relation between epenthetic elements (E) and their Bases (B) – and markedness constraints, which prefer featurally unmarked segments.[4] Varying resolutions of this conflict produce the spectrum of attested epenthetic qualities.

The two ends of the spectrum are shown in the tableaux below. BE-IDENT-F requires featural identity between epenthetic elements and the segments they correspond to. M(V→i) stands for the set of constraints that conspire to prefer [i] as the most harmonic vowel. With BE-IDENT-F outranking M(V→i), copy epenthesis results, as shown in tableau (3); the opposite ranking – in tableau (4) – produces the unmarked quality.

(3) Copy Epenthesis: ||BE-correspondence » Markedness||

/at/	BE-IDENT-F	M(V→i)
☞ (a) $a_1 t \underline{a}_1$		x x
(b) $a_1 t \underline{i}_1$	x!	x

(4) Unmarked epenthesis: ||Markedness » BE-correspondence||

/at/	M(V→i)	BE-IDENT-F
(a) $a_1 t \underline{a}_1$	x x!	
☞ (b) $a_1 t \underline{i}_1$	x	x

There is one final point about these tableaux: why can't a candidate such as [it<u>i</u>] win, so satisfying both M(V→i) and BE-IDENT-F? The reason is that faithfulness to underlying vowel quality, imposed by IO-IDENT-F, outranks M(V→i).

We note that there is another situation which can result in default segmentism: when E does not correspond to anything at all. This will be discussed in section 3.1.2. For the purposes of both this and the following sections, it will be assumed that E must correspond to some other output segment.

As expected, the same effects obtain in reduplication. In many languages, reduplicants are featurally faithful to their bases, so ||BR-IDENT-F » M||. However, in languages such as Lushootseed (Salish), vowels in reduplicants must be default [i], regardless of the Base vowel's quality: e.g. [t<u>i</u>təlaw'il] 'jog', [s<u>i</u>du:k^{w}] 'small knife' (Urbanczyk 1996, Alderete et al. in press) – hence ||M(V→i) » BR-IDENT-F||.

[4] In our survey of 67 languages with default epenthesis, we found that if a language allows [ə] to appear in its output forms, it was most likely to be the epenthetic vowel. In languages without [ə], [i] and [a] were common defaults, although [i] was twice as likely as [a] to be chosen. We found eight languages with epenthetic [e]. On inspection, though, [e] was found to almost always appear inside closed syllables. That [e] is less marked than [ə] in closed syllables is supported by phonological processes that convert [ə] to [e] in precisely this environment (e.g. French, Sekani – Hargus 1988:290). A telling case is found in Temiar: [ə] is epenthesized in open syllables, but [e] appears in closed syllables (McCarthy 1982). Other qualities ([ɨ], [o], [u], [ɯ]) occurred in a minority of languages which were usually genetically related. For consonant epenthesis, [ʔ] is most common, with [t] generally chosen in languages without [ʔ]. Epenthesis of [r] is attested in Japanese, English, and Southern Tati. In these cases, we consider [r] a type of glide, and suggest that it is best seen as a partial copy of adjacent vowels (as [j] is of [i], for example). See Kahn (1974) for similar conclusions.

However, epenthetic quality is not simply a parametric choice between copying and default segmentism. In a number of languages, epenthetic vowels copy only some features while markedness constraints dictate the quality of others. This is found in Ponapean epenthesis (Rehg & Sohl 1981:55,94). Epenthetic vowels are always [+high], but copy the roundness of the vowel to their right: e.g. [pwihk<u>i</u>men] *no gloss*, [<u>i</u>nta] 'blood', [<u>u</u>ŋkopw] 'species of crab', [ak<u>u</u>suwei] *no gloss*.[5]

Such cases of partial copy epenthesis can be accounted for by intermingling markedness and B𝔼-correspondence constraints. For Ponapean, the markedness constraint *[-high] outranks B𝔼-IDENT-F. This means that epenthetic vowels will always be [+high] in deference to *[-high], but otherwise agree with their correspondents, satisfying B𝔼-IDENT-F as much as possible.

(5)

/ŋkopw/	*[-high]	B𝔼-Ident-F	M(V→i)
(a) <u>o</u>$_1$ŋko$_1$pw	x x!		x x
(b) <u>i</u>$_1$ŋko$_1$pw	x	x x!	x
☞ (c) <u>u</u>$_1$ŋko$_1$pw	x	x	x x

Candidate (a) has a fully faithful epenthetic vowel but by doing so fatally incurs an extra *[-high] violation (the other violation is caused by the non-epenthetic Base vowel). This leaves the unmarked epenthetic vowel in candidate (b) or the partially faithful one in candidate (c). Candidate (b) is eliminated since it disagrees in both backness and roundness with the preceding vowel, gratuitously violating B𝔼-IDENT-F. This leaves candidate (c), in which 𝔼 necessarily disagrees with its correspondent in terms of [high] (so incurring one violation of B𝔼-IDENT-F) but is otherwise faithful.

Other examples of partial copying are found in Chadic vowel epenthesis (Frajzyngier & Koops 1989) and in Dakota consonant epenthesis (Shaw 1980).

In summary, permuting the ranking of B𝔼-correspondence and markedness constraints produces a range of possible epenthetic qualities, running from fully faithful copying through partial copy to default segmentism.

2.1 Contextually-conditioned quality

If two constraints $\mathbb{C}_1$ and $\mathbb{C}_2$ are in conflict, the effects of the lower ranked constraint can emerge when a higher ranked constraint $\mathbb{C}_3$ blocks the application of $\mathbb{C}_1$ (McCarthy & Prince 1994). This has implications for the present analysis: since B𝔼-correspondence and markedness constraints are in conflict, the effects of whichever constraint is lower-ranked could emerge in specific environments within the same language. In short, the mechanism of constraint ranking predicts that there could be language-internal variation in epenthetic quality, with copy epenthesis in some environments and default segmentism in others.

[5] In restricted environments, Ponapean epenthetic vowels can be fully faithful copies of vowels to their right. See Rehg & Sohl (1981:87) for details.

This prediction is borne out in Cook Islands Maori (CIM), a Polynesian language (Kitto 1997). CIM has five vowels /i e a o u/, and permits only (C)V(V) syllables. When foreign words with consonant clusters or word-final consonants are borrowed, they are made to conform by epenthesis. The quality of the epenthetic vowel varies depending upon its environment:

(6) (1) After /ir/, epenthesize an [a]
e.g. [pir<u>a</u>] 'bill
otherwise (2) After /r/, epenthesize a copy of the preceding vowel
e.g. [pe:r<u>e</u>] 'bail', [ʔa:mar<u>a</u>] 'hammer',
[po:r<u>o</u>] 'ball', [vu:r<u>u</u>] 'wool'
otherwise (3) epenthesize [i].
e.g. [kara:t<u>i</u>] 'carrot', [menet<u>i</u>] 'minute', [naeron<u>i</u>] 'nylon'

In other words, [i] is the default epenthetic quality. Only after [r] is there variation. After [ir], the epenthetic vowel is [a], while after [{e,a,o,u}r] it is a copy of the preceding vowel.

In brief, our analysis employs the constraint *ri, which bans [ri] sequences.[6] The constraint *ri blocks default insertion of [i] after [r]. This allows the next possible repair to emerge – copying of the preceding vowel. However, copying after an [ir] sequence would result in the illicit [ri] sequence, so copying is prohibited in just this environment; default [a] is inserted instead.

Since the default situation is for Ë to be [i], M(V→i) must outrank BË-IDENT-F, as established in the preceding section. However, as pointed out above, the subordination of BË-IDENT-F does not necessarily mean that the epenthetic vowel can never be a copy in the language. In fact, epenthetic vowels are copies in CIM when they are preceded by [r]. To put this in slightly different terms, the constraint M(V→i) is blocked by a higher constraint Ç, allowing BË-IDENT-F to emerge.

As stated above, the blocking constraint Ç is *ri. Default Epenthesis of [i] in non-[r] contexts is shown in example (1) below, while emergence of copying after [r] is shown in (2):

[6] The effect of this constraint can be seen in a number of languages: e.g. in Yoruba [r] must delete before [i] (Akinlabi 1993). It may be seen as an OCP constraint – a constraint against similar elements – since [i] and liquids seem to have some sort of affinity, as shown by the fact that they are in complementary distribution in Cibaeño Spanish (Harris 1983) and in Portuguese Spanish (Martinez-Gil 1997). The Peruvian language Chamicuro shows a similar relation (Steve Parker, p.c.).

(7)

❶	/menet/	*ri	M(V→i)	BE-IDENT
☞	meneti		x x	x
	menete		x x x!	
❷	/per/	*ri	M(V→i)	BE-IDENT
	peri	x!	x	x
☞	pere		x x	
	pera		x x	x!

Example (1) in the tableau above shows default insertion of [i]. Since the blocking constraint *ri is inapplicable here, the low-ranked BE-IDENT has no effect. This contrasts with example (2) where the high-ranked *ri blocks the possibility of default to [i]. The other option available is to copy, obeying BE-IDENT.

The final part of this system is when both M(V→i) and BE-IDENT-F are blocked by *ri. In such a situation, the effects of even lower ranked constraint emerge. These constraints – symbolized by M(V→a) – prefer [a] as the next most unmarked vowel:

(8)

	/pir/	*ri	M(V→i)	BE- IDENT	M(V→a)
	(a) piri	x!			x x
	(b) pire		x	x	x x!
☞	(c) pira		x	x	x

Candidate (a) – with epenthesis of [i] – is the only one that satisfies M(V→i) and BE-IDENT-F, but fatally violates *ri. All other candidates violate both constraints, so rendering them irrelevant in determining the optimal form. Candidate (c), with the next most unmarked vowel – [a] – emerges triumphant.

On the empirical side, CIM is a striking case of language-internal epenthetic variation: epenthetic quality varies from unmarked to copy to unmarked in increasingly restricted environments. On the the theoretical side, this case shows that complex variation in epenthetic quality can be explained by the interleaving of BE-correspondence constraints and markedness constraints, with relevant blocking constraints ranked high.

As a final note, CIM is not unique in the complexity of its variation. Similar cases are found in Sentani (Trans-New Guinea) vowel epenthesis (Cowan 1965 *cf* Elenbaas 1999) and in Faroese consonant epenthesis, discussed in the next section.

2.1.1 A note on global conditions

The case described above is significant in a more general theoretical respect. Effectively, the condition *ri triggers changes that can be arranged in a hierarchy of decreasing preference: |default to [i] > copy > default to [a]|. Situations where a condition C triggers a hierarchy of repairs $|R_1 > R_2 > \dots R_n|$ are easily captured in OT by the ranking $||C \gg R_1 \gg R_2 \gg \dots R_n||$, as illustrated above.

Such a situation is not easily captured in theories that employ serial derivations. For example, the CIM facts can be cast in terms of SPE formalism as follows:

(9) Ɛ → a / ir_
V_i / V_ir_
i *elsewhere*

This formulation misses the point that it is the sequence [ri] that is being avoided here. A similar point is made for epenthesis in Chukchi (Chukotko-Kamchatkan) by Krause (1973) and by Wilbur (1974) for certain cases of reduplication.

The CIM situation is not easy to capture in theories with localised constraints either – i.e. constraints that apply at a certain point in the derivation, triggering a repair:

(10) Process: ∅ → i
Constraint: *ri *triggers* Repair: ∅ → V_i/ V_iC_
Constraint: *ri *triggers* Repair: ∅ → a

The fact that *ri triggers both repairs is entirely incidental in this system. The problem is that *ri is a 'persistent' condition on well-formedness, holding at every point in the derivation. While extra mechanisms must be invoked to account for such persistence in serial theories (see e.g. Chafe 1968, Anderson 1974, Myers 1991), it falls out naturally in parallelist OT.

3.0 The Dynamic Base

In the examples presented so far, the Base – the segment that the epenthetic element corresponds to – was given without comment. However, determining the identity of the Base is far from trivial. For example, epenthetic vowels in Selayarese copy vowels to their left (Mithun & Basri 1985), while those in Bardi copy from their right (Metcalfe 1975). Proximity also varies: in some languages, the epenthetic vowel copies an adjacent segment, while in others non-adjacent segments are copied.

This variation in the identity of the Base precludes a 'static' definition, such as "The Base is the nearest segment to the right of Ɛ." (*cf* the static definition for the Base of reduplicants – McCarthy & Prince 1994). Instead, we propose that the identity of the Base of epenthetic elements is determined 'dynamically' – by constraint interaction. The Base, then, is simply that segment to which Ɛ corresponds.

As a simple example, take the hypothetical form /an/. An epenthetic vowel is added to this string in response to the triggering constraint NOCODA, which bans coda consonants, so producing [ana]. That the Base of Ɛ is [a] and not [n] is due to the constraint BƐ-IDENT-F, as illustrated in the tableau below:

(11)

/an/	NOCODA	BƐ-IDENT-F
(a) an_1n_1	x!	
(b) an_1i_1		x!
☞ (c) a_1na_1		

In candidate (a), the epenthetic element corresponds to the immediately preceding segment and faithfully copies it. The problem with this form is that it has gratuitous epenthesis – the epenthetic element fails to solve the problem posed by NOCODA: NOCODA can only be satisfied by adding vowels (in this language), not by adding consonants. While candidate (b) does satisfy NOCODA by epenthesizing a vowel, it necessarily incurs faithfulness violations since its correspondent is not a vowel. This leaves candidate (c), which satisfies both NOCODA and B-IDENT-F.

In short, there is no need to invoke a special mechanism for identifying the Base; the Base of E is simply the segment in the most harmonic form that corresponds to E.

The aim of the rest of this section is to explore the typological consequences of the dynamic Base. As shown in section 2, the mechanism of constraint ranking predicts not only cross-linguistic variation in the identity of the Base, but language-internal variation as well. For example, the Base may be to the left of E in some contexts yet to the right in others in the same language. This sort of variation will be shown to result from constraint ranking.

The set of cases examined below are divided into two groups: (1) those involving variation in direction – whether the Base is to the left or right of E, and (2) those involving variation in proximity – whether the Base is adjacent to E or further away. The former is discussed in section 3.1 and the latter in section 3.2.

3.1 Dynamic Direction

The following constraints regulate direction of copying:[7]

(12) COPY-LEFT "E corresponds to a segment to its left."
COPY-RIGHT "E corresponds to a segment to its right."

In terms of direction, cross-linguistic variation is straightforward: when COPY-LEFT outranks COPY-RIGHT copying from the left takes place; the opposite ranking produces the opposite result. Languages that employ left copying include Amele (Trans-New Guinea) (Roberts 1987:375), Barra Gaelic (Celtic) (Ni Chiosáin 1995), Kolami (Dravidian) (Zou 1991), most dialects of Ostrobothnian Finnish (Kettunen 1940), Selayarese (Western Malayo-Polynesian) (Mithun & Basri 1985), Takelma (southwest Oregon) (Sapir 1922), and Wolof (Niger-Congo) (Ka 1994). Copying from the right is found in Bardi (Metcalfe 1975), a dialect of Ostrobothnian Finnish (Kettunen 1940), Ponapean (Micronesian) (Rehg & Sohl 1981), and Winnebago (Miner 1974).

In the remainder of this section, attention will be focussed on cases where direction of copying varies within the same language. There are two general questions that arise in such cases: (1) what can block the normal direction of copying?, and (2) what are the possible responses to such blocking? To explore these issues, we will first focus on consonant epenthesis in Faroese.

[7] These constraints have their roots in McCarthy & Prince's (1993, 1995) ALIGN and ANCHOR constraints. The standard ALIGN and ANCHOR constraints cannot be used here: since the left/right edge of a segment is the segment itself, they cannot distinguish forms such as [CaCaCi] from [CaCiCi] as both satisfy ANCHOR-LEFT and ANCHOR-RIGHT equally well.

3.1.1 Faroese

In Faroese (Germanic), consonant epenthesis takes place in situations of heterosyllabic vowel hiatus (Anderson 1972a,b). The direction and nature of epenthesis varies in different environments, as described below:

(13) (1) Copy from the left if [i] or [u] precedes:
e.g. [si:j̲ur] 'custom', [hyuw̲ir] 'skins'
otherwise (2) Copy from the right if [i] or [u] follows:
e.g. [so:j̲in] 'boiled', [mæaw̲ur] 'man'
otherwise (3) Do not epenthesize.

The glides [j] and [w] are featurally identical to the vowels [i] and [u], respectively; this is why this is a case of copy epenthesis. They differ only in terms of moraic content – vowels license a mora while glides do not (see the next section for further discussion). However, [j] and [w] are not as similar to non-high vowels. In fact, there are no glide equivalents of non-high vowels in Faroese, a fact which will prove to be significant.

The standard situation in Faroese is for Ë to copy from the left, hence ||COPY-LEFT » COPY-RIGHT||. However, this is blocked when a non-high vowel precedes Ë. This is readily explained in terms of constraints: in such a situation, copying to the left would necessarily incur violations of BË-IDENT-F since (as mentioned above) there are no consonantal equivalents of non-high vowels in Faroese. So, this system can be described as "Copy left, except when that would result in an unfaithful copy, then copy from the right." Since BË-IDENT-F blocks COPY-LEFT, it must outrank it, as shown in the tableaux below:

(14)

1.	/i_u/	BË-IDENT-F	COPY-LEFT	COPY-RIGHT
☞	$i_1 j_1 u$			x
	$i w_1 u_1$		x!	
2.	/o_i/	BË-IDENT-F	COPY-LEFT	COPY-RIGHT
	$o_1 w_1 i$	x!		x
☞	$o j_1 i_1$		x	

Example (1) shows the usual situation of left-copying. Example (2) shows how a change in direction of copying can be compelled by BË-faithfulness. Note that we ignore the case where Ë does not correspond to anything at all (see the next section for discussion).

However, there are two responses to blocking of leftward copying. The response shown in example (2) is to change direction. However, this is only possible if doing so would not violate BË-IDENT-F. In an environment such as /e_o/, copying from either direction would incur faithfulness violations. In such situations, epenthesis simply does not take place.

The constraint that triggers epenthesis in Faroese is ONSET "Syllables must have onsets." Since epenthesis is blocked just in case it is not possible to copy faithfully, BË-IDENT-F must outrank ONSET. This is illustrated in the tableau below:

(15)

/oa/	BE-IDENT-F	ONSET
(a) $o_1\underline{w}_1a$	x!	
(b) $o\underline{w}_1a_1$	x!	
☞ (c) oa		x

This tableau shows that it is better to ignore the pressure to epenthesize imposed by ONSET than to produce an unfaithful copy – both (a) and (b) fatally violate BE-IDENT-F, leaving the non-epenthesized form (c).

Faroese is much like CIM in that a condition triggers a hierarchy of responses: it is best to copy from the left, then to copy from the right, then to not epenthesize at all.

3.1.2 Typology

Faroese shows that faithfulness constraints can block the usual direction of copying. However, the most common reason why the usual direction of copying is prevented is that in some environments there is simply nothing to copy. For example, epenthetic vowels in Hawaiian (Polynesian) loanwords are copied from the left, but when epenthesis is initial there is nothing appropriate to copy, so direction reverses:

(16) Hawaiian (Kitto 1997)

(i) Left-Copying:	[ʔen<u>e</u>kinia] 'engineer'	
(ii) Initial Right-Copying:	[p<u>e</u>lekania] 'Brittania'	[p<u>a</u>lani] 'brandy'

The morphological status of what is copied can also figure in blocking. For example, epenthetic vowels copy to their right in Winnebago (e.g. [hip<u>e</u>res] 'know' – Miner 1992:34). However, they only copy root material; when only affix segments are to the right of E, a [ə] is inserted: e.g. /wanĩk/ 'bird' + RED_{Af} 'little' + /ra/ 'the' → [wanĩk<u>ə</u>nĩk<u>ə</u>ra] 'the little bird', *[wanĩk<u>ĩ</u>nĩk<u>a</u>ra]. A similar situation is found in Wolof (Ka 1994).

This type of blocking can be accounted for by invoking a root-specific faithfulness constraint: COPY-ROOT "If E corresponds to B, then B is in a root." (see Urbanczyk 1996 for an almost identical constraint on the Base-Reduplicant dimension). With high-ranked COPY-ROOT, copying can be blocked, allowing the low-ranked markedness constraints M(V→ə) to emerge.

This brings us to consider the responses to blocking. Faroese illustrated two responses: (1) reversing direction and (2) failing to epenthesize. The former response is also found in Hawaian (as shown above), and in Tamil (Dravidian) consonant epenthesis (Wiltshire 1998). The latter response is also found in consonant epenthesis in Dutch (Booij 1995) and Karo Batak (Woolams 1996) and for vowel epenthesis in Damescene Arabic (Semitic) (McCarthy 1979).

A third response is found in Wolof and Winnebago, as mentioned above. When copying in one direction is impossible, the epenthetic vowel takes on default features. In Winnebago, for example, when copying from the right is blocked, E is realised as [ə]. A

similar situation is found in Malay: epenthetic consonants are copies of vowels to their left, but when this is blocked a default [ʔ] is inserted.

To account for this option, the constraint BE-CORR "E must correspond to something" must be invoked. In cases where this constraint is violated, E does not correspond to any segment. So, BE-faithfulness constraints are irrelevant in determining E's quality; hence its featural content is entirely due to markedness constraints. In cases where the response is to default to an unmarked value, then, it is not BE-identity that is blocked, but rather BE-CORR. As an illustration, COPY-ROOT blocks affix copying in Winnebago (as discussed above). With COPY-ROOT outranking BE-CORR, it is better not to correspond at all rather than copying non-root material:

(17)

1. wanĩk + ra	COPY-ROOT	COPY-LEFT	BE-CORR
(a) wanĩk$\underline{a}_1$ra$_1$	x!		
(b) wanĩ$_1$k$\tilde{\underline{i}}_1$ra		x!	
☞ (c) wanĩk$\underline{ə}$ra			x

The empirical aim of this section was to show that the Base of epenthesis can vary, not only cross-linguistically but within languages as well. The theoretical aim was to show that both the causes of and responses to blocking of the usual direction of copying can be explained by BE-correspondence constraints and their interaction with other constraints. However, direction is only one dimension of variation; the following section explores similar variation in Base-E proximity.

3.2 Dynamic Proximity

The constraint that requires proximity is BE-ADJACENCY "B and E must be adjacent." Violations of BE-ADJACENCY are calculated gradiently, with more violations the further B and E are apart. If this constraint is ranked high enough, an epenthetic element and its Base must be adjacent. Such a case is found in Southeastern Pomo (Hokan), where pre-tonic consonant clusters are broken up by vowel epenthesis (Moshinsky 1974). The inserted vowel agrees in place and roundness with adjacent consonants:[8]

(18) E → i / after front consonants: e.g. [čij̲ačjakin] 'green'
→ u / after labials: e.g. [mu̲wata] 'talk!'
→ u / after back consonants: e.g. [xu̲wan] 'dance house'

In other languages, BE-ADJACENCY is violated. For example, in Bardi (and many other cases of vowel-copy epenthesis) E copies the nearest vowel, ignoring intervening consonants. The reason for this is clear: if E copied anything except for a vowel, it would

[8] This is simplifying the situation somewhat. In some situations, the epenthetic vowel agrees with the following vowel: e.g. [ʔe̲ke] 'to catch'. The exact environments that trigger such agreement are not entirely clear in Moshinsky's data, although the general trends described above are evident.

incur violations of BE-IDENT-F. This situation is analogous to Faroese: faithfulness constraints again act as blockers.

(19)

	/an/	BE-IDENT-F	BE-ADJACENCY
	$an_1\underline{a}_1$	x!	
☞	$a_1n\underline{a}_1$		x

Variation in what is copied can be ascribed to the fact that there are individual IDENT constraints for each feature. For example, BE-IDENT-[high] requires E and its base to agree in terms of [high] alone.

Feature-specific IDENT constraints can be used to explain the minimal variation found in Winnebago and Awtuw. In Winnebago E copies the nearest **vowel** (Miner 1992), while in Awtuw (Papua New Guinea) the nearest **vocoid** is copied (Feldman 1986):

(20) Winnebago: E copies nearest vowel to its right: [bo:p$\underline{u}$nũs]] 'hit at random', [hip$\underline{e}$res] 'know'
ignoring intervening glides: [ho$\int\underline{a}$waza] 'you are ill', *[ho$\int\underline{u}$waza]
[k$\underline{e}$we] *no gloss*, *[k$\underline{u}$we]
Awtuw: (i) E copies nearest glide: [k$\underline{i}$njel] 'cry!', *[k$\underline{e}$njel]
else (ii) nearest vowel: [k$\underline{o}$nor] 'copulate!', [k$\underline{u}$mpuya] 'hit!'

It cannot be the case that every BE-IDENT-F constraint outranks BE-ADJACENCY in Awtuw otherwise /knjel/ would be realised as *[k$\underline{e}$njel]; [k$\underline{i}$njel] would be ill-formed since the epenthetic [i] is not a perfect copy of [j] – [i] bears a mora while [j] does not, so violating BE-IDENT-μ.

Even so, this case is easily explained: it is simply a matter of ranking the relevant BE-IDENT-F constraints low. In this case, BE-IDENT-μ must be ranked below BE-ADJACENCY. This is shown in the tableau below, where BE-IDENT-F' stands for every BE-IDENT constraint except for BE-IDENT-μ:

(21)

	/knjel/	BE-IDENT-F'	BE-ADJACENCY	BE-IDENT-μ
	(a) $k\underline{i}_1n_1$jel	x!		
	(b) $k\underline{e}_1nje_1$l		x x!	
☞	(c) $k\underline{i}nj_1$el		x	x

Candidate (a) satisfies BE-ADJACENCY, but at the cost of violating BE-IDENT-F' since [n] and the necessarily vocalic epenthetic element disagree in several features. While candidate (b) satisfies all BE-IDENT constraints, it violates BE-ADJACENCY twice since two segments intervene between E and its Base. These violations are crucial, compared with candidate (c). With (c), only one segment intervenes between E and its Base. While

[i] is not perfectly identical to [j], the constraint that is sensitive to their difference – BE-IDENT-μ – is ranked low enough to be irrelevant.

So, requirements on BE identity can affect Base proximity in a gradient manner. This approach predicts a variety of other systems. For example, an epenthetic vowel would copy the nearest **sonorant** if BE-IDENT(sonorant) outranked BE-ADJACENCY and all other conflicting BE-IDENT-F constraints were ranked low (i.e. those referring to the features [vocoid], [approximant], [continuant], etc.). Also, it is predicted that blocking of adjacency should induce the same array of responses found in blocking of direction (e.g. if an adjacent segment can't be copied, then don't copy at all). We leave the determination of the validity of these predictions to future research.

3½ Processes and Base-E identity

The aim of this section is to examine the interaction between phonological processes and BE-identity.

Phonological processes that apply generally in a language may also apply to epenthetic vowels. For example, as shown in section 3.1, epenthetic consonants in Faroese are copies of preceding high vowels. However, after non-front vowels, [w] is realised as [v]: /uwa/ → [uva]. In other languages, a phonological process applies only to epenthetic segments. An example is found in Cook Islands Maori: epenthetic vowels are usually copies, but labial attraction forces them to be [u] after labial consonants: e.g. [ra:mu] 'ram', *[ra:ma], [terepu] 'trip', [ʔaopu] 'hop'.

The difference between these two sorts can be expressed by the ranking of Input-Output faithfulness constraints with respect to the trigger of the phonological process P. If P outranks IO-FAITH, P will apply to both underlying and epenthetic segments. With the opposite ranking, P will only apply to epenthetic segments.

An analogous case is found in Balangao reduplication (McCarthy & Prince 1995:39). Codas are permitted in general, but are banned in reduplicants. This can be explained by the ranking ||IO-MAX » NOCODA » BR-MAX||.

Phonological processes can also 'misapply'. For example, they may apply to epenthetic segments even when those segments are not in the right environment. An example is found in Winnebago nasalisation (Miner 1992). Winnebago vowels are nasalised after nasal consonants: e.g. [wamãnũke] 'thief'. Epenthetic elements copy these vowels, nasality and all: e.g. /knak/ → [kãnãk] 'marry'. Note that the epenthetic vowel is nasalised despite the fact that it is not preceded by a nasal consonant. Such a situation results from the requirements of BE-IDENT-F. With BE-IDENT-F outranking any constraint against nasal vowels in general (e.g. $*V^{nasal}$), it is free to copy the Base with impunity[9]:

[9] IO-IDENT outranks $*V^{NASAL}$ as there are lexically nasal vowels in Winnebago.

(22)

/knak/	BE-IDENT-[nasal]	*NV^{oral}	IO-IDENT	*V^{NASAL}
(a) kanak		x!		
(b) kanãk	x!		x	x
☞ (c) kãnãk			x	x x

There are analogous cases of overapplication in reduplication. In fact, Madurese presents an almost identical case: segments following a nasal segment are nasalised, and this nasality is reflected in the reduplicant even when it is not in the right environment: [ỹãt-nẽỹãt] 'intentions', *[yat-nẽỹãt] (McCarthy & Prince 1995:§4.1).

Cases of underapplication in epenthesis – when a process that applies in the language generally fails to apply to epenthetic segments – also occur. One of the most common types of epenthetic underapplication is where stress avoids epenthetic vowels (Alderete 1995). There are also cases involving segmental processes:

(23) Sekani: ə→e/_s e.g. /nə-sə-kat/ → [nezkat]
cf /s-d-yhõ/ → [əsdyhõ], *[esdyhõ] (Hargus 1988)
Harari: t → č/_i e.g. /tikäft-i/ → tikäfči
cf /tikäft/ → [tikäfti], *[tikafči] (Kenstowicz 1981)

In Sekani, [ə] is realised as [e] in s-conjugation verbs, except when the [ə] is epenthetic. The Harari (Semitic) case is slightly different: in this situation the epenthetic vowel [i] fails to **trigger** a process.

As in reduplication, such cases do not fall out naturally from permuting the ranking of BE-FAITH, IO-FAITH and the triggering constraint P. Instead, a constraint that effectively prevents the process from applying in epenthetic contexts alone is needed. For the stress-avoidance cases, a positional faithfulness constraint such as IO-DEP-σ́ "If *x* is in a stressed syllable then *x* has an input correspondent" can be employed (Alderete 1995, Beckman 1998). With IO-DEP-σ́ outranking the stress placement constraints, stress will avoid epenthetic vowels.

In the Harari case, Kenstowicz (1981) invokes a constraint on paradigm distinctiveness. He notes that if the masculine form /tikäft/ was realised as [tikäfči] it would be identical to the feminine form [tikäfči]. Hence, failure to apply palatalisation in this context is not due to epenthetic vowels, but to a coincidental requirement on morphological distinctiveness.

The final situation to be discussed here is 'Alternative application': when epenthetic segments and non-epenthetic segments undergo different processes to satisfy the same constraint. An example of this is discussed in section 4, so only a brief summary will be presented here. In Selayarese, main stressed syllables must be bimoraic. To achieve this goal, non-epenthetic vowels lengthen: e.g. /sahala/ → [sahá:la] 'sea cucumber'. However, epenthetic vowels do not lengthen: instead, a glottal stop is inserted or the following consonant is geminated: e.g. /sahal + mu/ → [sahala?mu] "your (fam.) profit", *[sahala:mu].

In this case, the change in both underlying and epenthetic vowels is triggered by the same constraint: *σ_{μ} "Monomoraic stressed syllables are prohibited." However, the usual response – lengthening – is blocked for epenthetic vowels and the less preferred response of coda-insertion emerges. The details of this mode of application are discussed in section 4. For the moment, it is enough to note that the essential part of this ranking is that the blocking constraint B outranks the constraints responsible for the responses to the triggering constraint.

To summarize, three general situations can be described in the interaction of phonological processes with epenthetic elements. A constraint T can trigger a process in E – normal application. However, T can be blocked from applying to E (underapplication), and in other cases T's effects on E's Base can be seen in E even when E is not in the correct environment (overapplication). The final situation is when T applies to both E and underlying segments, but the responses to T are different in both cases (Alternative application).

Of course, there are a number of other predictions. For example, back-copying is predicted – where a process applies to E and its effects are reflected in its correspondent. As yet, we have found no entirely convincing cases of back-copying (see McCarthy & Prince 1995 for discussion of back copying in reduplication).

4.0 Alternatives

The aim of the preceding sections has been to explore some of the basic consequences of using correspondence to account for epenthetic quality. However, there are a number of alternatives to correspondence. In particular, feature sharing approaches, as in autosegmental theory, can be used to explain many of the same facts.[10] The purpose of this section is to examine the differences between the correspondence approach and feature sharing theories. [11]

In feature sharing theories, variation in epenthetic quality can be explained as differences in feature sharing: if E shares all its features with a nearby segment, copy epenthesis results; if E does not share any features, it will take on default values.[12] In many cases, the correspondence approach and feature spreading theories produce the

[10] By 'feature sharing theory', we refer to theories that allow a single feature to belong to more than one segment (most notably autosegmental phonology). Our aim in this section is not to consider particular theories, but rather identify properties that feature sharing theories generally agree on, and compare them to the correspondence approach.

[11] It was suggested to us at AFLA VI that copy epenthesis could be a type of reduplication. While our proposal employs essentially the same mechanism as used in reduplication, there is a difference: reduplication signals a morphological change, while epenthesis is triggered by entirely phonotactic considerations. If copy epenthesis were reduplication, this would be tantamount to suggesting that reduplicative morphemes could be inserted to satisfy phonotactic constraints. If so, it is a mystery why epenthetic elements do not have the same variation in size as reduplicants: epenthesis always inserts a single segment while reduplicants are usually larger. More concretely, if copy epenthesis is reduplication, why are VCV sequences never inserted to break up consonant clusters?

[12] In the case of vowel copy epenthesis, different theories allow different sets of features to spread across consonants (compare Clements 1985, Sagey 1986 with Hume 1992, Clements & Hume 1992). For discussion relevant to copy epenthesis, see Gafos & Lombardi (1997).

same results. However, there are differences. Due to space limitations, only one major difference will be discussed here: the treatment of suprasegmentals.[13]

Segmental properties such as stress (headedness), tone, and duration (moraic content) are commonly assumed to be associated with prosodic units above the root node (e.g. the mora or syllable) (see Hayes 1995:49ff for stress, Odden 1995:448ff for tone, and Broselow 1995 for duration). For feature-sharing theories, the suprasegmental nature of these properties means that they cannot be shared in copy epenthesis otherwise ill-formed representations result. This is particularly evidence for moraic content since it is not a unitary feature, hence not even a sharable unit. In comparison, correspondence theory allows reference to such properties.

A relevant case is found in Selayarese (Mithun & Basri 1985), where agreement in moraic content between Ɛ and its Base is necessary. In this language, syllables have the form (C)V(C). The one exception is the main-stressed syllable, which must be bimoraic. This is usually achieved by lengthening. However, when the stressed vowel is epenthetic, no lengthening takes place; instead, a glottal stop is inserted (or the following onset is geminated). This is shown in the following minimal pairs:

(24) (1) Underlying vowel lengthens:
/sahala+mu/ → [sahalá:mu] 'your (fam.) sea cucumber'
(2) Epenthetic vowel does not lengthen:
/sahal+mu/ → [sahala̲ʔmu], *[sahala̲:mu] 'your (fam.) profit'

This case is similar to the contextually-conditioned variation in epenthetic quality found in section 2. A hierarchy of repairs are triggered by a constraint requiring bimoraic main-stressed syllables (i.e. *$\acute{\sigma}_{\mu}$ "no monomoraic stressed syllables"). The dispreferred repair is the creation of a coda, while the preferred repair is vowel-lengthening. This can be put in OT terms by ranking NOCODA "Syllables must not have codas" above *V:, which bans long vowels.

As shown in section 2, in order to block the effect of NOCODA, a specific constraint must outrank it: ||ℂ » NOCODA » *V:||. Something can already be deduced about the identity of ℂ: it must refer to the epenthetic nature of the stressed vowel since this is the only property that distinguishes it from the other vowel types. This means that ℂ cannot be a markedness constraint since such constraints refer only to output configurations, and not to the relation between output and input segments. So, ℂ must be a correspondence constraint.

With BƐ-correspondence the solution is straightforward: epenthetic vowels cannot be lengthened because this would make them unfaithful to their Base in terms of moraic content (Ɛ's base is always monomoraic in Selayarese). In other words, BƐ-IDENT-μ "Ɛ and its Base must agree in terms of moraic content" outranks NOCODA. The constraints are shown in the following tableau:

[13] See Kitto & de Lacy (1999) for further discussion.

(25)

(1)	/sahal+mu/	*σ$_\mu$	BE-IDENT-μ	NOCODA
	saha$_1$lá$_1$mu	x!		
	saha$_1$lá:$_1$mu		x!	
☞	saha$_1$lá$_1$ʔmu			x
(2)	/sahala+mu/	*σ$_\mu$	BE-IDENT-μ	NOCODA
	sahalámu	x!		
☞	sahalá:mu			
	sahaláʔmu			x!

As shown in example (1), lengthening of the stressed vowel is blocked by BE-IDENT-μ since such a change would introduce disparity between the epenthetic segment and its correspondent. Example (2) shows the effect on an underlying vowel: moraic identity between epenthetic segments and their correspondents is irrelevant here, so lengthening is permitted.

This solution has one further effect: when epenthetic vowels are final, stress appears on the antepenult, not the penult (e.g. /sahal/ → [sá:hala] 'profit'). This is explained straightforwardly under the present approach. The epenthetic vowel corresponds to the penult, but must agree with it in terms of moraic duration. However, if the penult were stressed it would be forced to lengthen, so violating BE-IDENT-μ. The strategy taken to avoid this is simply to retract the stress from the penult – in [sá:ha$_1$la$_1$], stressed vowels can lengthen and BE-IDENT-μ can be satisfied, at the cost of the penultimate position of stress.

In summary, it seems that agreement in terms of suprasegmental features between epenthetic segments and their Bases is possible, lending credence to the correspondence approach.[14]

The Correspondence and feature-sharing approaches do seem to diverge on another point. Feature-sharing theories restrict sharing to local configurations. This means that consonants cannot share features with non-adjacent segments without also sharing them with segments inbetween, and vowels may share features with either adjacent segments or the nearest vowel, skipping intervening consonants .[15] Correspondence, on the other hand, does not seem to be as inherently limited – long-distance correspondence seems à priori possible. Nevertheless, we will argue that in some cases non-adjacent copying between consonants can be prohibited entirely, as can long-distance copying between vowels. In other situations, the reason why long-distance copying does not take place is

[14] A case involving identity of tone is found in Sekani (Hargus 1988: 133). Certain prefixes assign low tone to immediately preceding vowels. This does not happen to immediately preceding epenthetic segments, though. This can be explained as the imposition of tonal identity between E and its correspondent – the prefix vowel. A case of stress retraction in Tahitian (Bickmore 1995) can be explained as a requirement that E and its Base be identical in terms of headedness (~stress).

[15] Some feature sharing theories (e.g. Hume 1992, Clements & Hume 1995) attempt to derive this asymmetry by specific feature geometries and the line-crossing prohibition. Recent approaches (e.g. Gafos 1996, 1998) require all feature sharing to be strictly local, and the asymmetry results from different incompatibilities between vowel and consonant features.

part of a larger problem – why markedness constraints cannot force violations of adjacency constraints.

As shown in section 3, violations of adjacency take place when a constraint C outranks and conflicts with BE-ADJACENCY. C can be either a faithfulness constraint or a markedness constraint. Cases where faithfulness constraints trigger violations of adjacency will be discussed first.

Faithfulness constraints can force violations of adjacency. This is achieved by employing an IDENT constraint that refers to a feature that E and an adjacent segment inherently disagree on. For example, vowels have moras while consonants do not, so BE-IDENT-μ can be used to force a vowel to correspond to another vowel, ignoring intervening consonants (as shown in section 3). Put in this way, the problem of long-distance correspondence reduces to identifying consonant features that are inherently incompatible with vowels, and vice-versa. We propose that there is an asymmetry here: vowels have a feature that is incompatible with consonants, but no consonantal feature is incompatible with vowels. In other words, BE-IDENT-*f* can never be violated if a consonant corresponds to a vowel, but it can be violated if a vowel corresponds to a consonant.

What exactly is the incompatible feature? Notice the difference between the vowel /u/ and the consonant /w/: these are featurally identical except for the fact that the former has a mora, while the latter does not. So, the asymmetry, we suggest, results from the following constraint:

(26) EB-IDENT-μ "If E has a mora then B has a mora."

Crucially, there are no constraints of the form "If B has a mora then E has a mora" or "If E does not have a mora, then B does not have a mora."

Since consonants have no mora, an epenthetic consonant can correspond to a vowel without violating EB-IDENT-μ (since it fails to satisfy the antecedent of the constraint "If E has a mora").[16] So, in terms of faithfulness, epenthetic consonants will always correspond to an adjacent segment – whether it is a vowel or consonant is irrelevant, since both segment types are compatible. Long-distance correspondence is prohibited because it would cause gratuitous violations of BE-ADJACENCY. This is shown in the following tableau:

(27)

/kua/	BE-IDENT-*f*	BE-ADJACENCY
☞ $ku_1\underline{w}_1a$		
$k_1u\underline{k}_1a$		x!

[16] Of course, we are assuming that epenthetic consonants always end up without moras. This may not be the case with some epenthetic coda consonants. In such cases, EB-IDENT-F could cause the coda consonant to copy the nearest vowel, skipping intervening consonants. Such cases are rare though, and so may be reasonably ignored at this point in time.

Because consonant features are compatible with both vowels and consonants, both incur the same faithfulness violations. This allows BE-ADJACENCY to emerge, ruling out copying between non-adjacent segments.

This can be compared with vowel epenthesis. As pointed out above, BE-IDENT-μ is violated when an epenthetic vowel corresponds to a consonant. This creates a conflict with BE-ADJACENCY, so allowing non-adjacent vowel correspondence:

(28)

	/up/	BE-IDENT-μ	BE-ADJACENCY
☞	$up_1\underline{i}_1$	x!	
	$u_1p\underline{u}_1$		x

In summary, faithfulness constraints alone cannot force violations of BE-ADJACENCY if the epenthetic segment is a consonant, but can if it is a vowel. This accounts for part of the consonant-vowel correspondence asymmetry.

However, it does seem à priori possible for **markedness constraints** to force long-distance copying between consonants. For example, suppose that a markedness constraint *C/+cont "Consonants must be stops" outranks BE-ADJACENCY. This would force E to be a stop, and since vowels are not stops, the best sort of segment for E to correspond to would be a consonant, so forcing long-distance consonant correspondence:

(29)

	/kua/	*C/+cont	BE-IDENT-*f*	BE-ADJACENCY
☞	(a) $ku_1\underline{w}_1a$	x!		
	(b) $ku_1\underline{p}_1a$		x!	
	(c) $k_1u\underline{k}_1a$			x

Candidate (a) fully satisfies BE-IDENT-*f*, but fatally violates *C/+cont since [w] is [+continuant]. While candidate (b) satisfies *C/+cont, it violates BE-IDENT-*f* since [p] and [w] disagree in terms of continuancy. Candidate (c), on the other hand, satisfies *C/+cont and satisfies BE-IDENT-*f*. BE-ADJACENCY is violated, but this is inconsequential due to its low rank.

Unlike the cases where faithfulness alone triggers violations of BE-ADJACENCY, feature incompatibility does not seem to offer an obvious solution to this problem. In fact, we have no good explanation for this state of affairs. However, we suggest that this problem is not solely an issue for BE-correspondence. It is symptomatic of a larger generalisation: markedness constraints cannot participate in identifying the Base. Evidence for this generalisation can be found both in epenthesis and reduplication.

For epenthesis, markedness constraints cannot force copying to change direction. We found no cases where epenthetic vowels or consonants changed direction in order to copy a less marked segment. For example, there were no cases where an epenthetic vowel that usually copied from the left copied from the right just so it could copy an [i]:

e.g. [at̲iki], *[at̲aki], *cf* [it̲ika]. All cases of direction reversal are triggered by faithfulness constraints (see section 3.1).[17]

Markedness constraints also cannot force long-distance consonant copying in reduplication. In virtually all cases (that we know of), reduplicant consonants correspond to the nearest available non-reduplicant consonant.[18] More concretely, we know of no example where the reduplicant copies the nearest least marked consonant: e.g. [ʔa͟͟kaʔa] *cf* [ʔaʔaka], [ta͟͟kata].[19] So, the question is not why long distance consonant correspondence is not permitted in epenthesis, but rather why it is banned in **every** situation, including an area where a correspondence approach seems necessary – reduplication.

In summary, markedness constraints cannot participate in determining the Base of either epenthesis or reduplication. Although we do not have a principled reason for this restriction, it shows that locality considerations cannot be used as an objection against the correspondence approach to epenthesis alone.[20] Rather, the problem is an issue for correspondence theory in general.[21]

As a final comment, we note that feature sharing is not necessarily incompatible with correspondence theory. However, if both are allowed, this results in a great deal of redundancy. In order to fully eliminate this redundancy in favour of correspondence, feature sharing must be prohibited. In effect, this returns to a conception of segments as sets of features, much as in SPE. 'Feature sharing' would then be the situation where two segments in correspondence agree in the value of a feature. Of course, this takes us far beyond the scope of the present paper, and affects areas such as vowel harmony and assimilation. We merely note that the logical endpoint of the proposals in this paper is the elimination of feature-sharing in phonological representations.

4.1 Breaking

Breaking is another approach to explaining copy epenthesis (McCarthy 1997). Breaking describes the situation where a single underlying segment has two output correspondents, as in diphthongization (e.g. /a/ → [ei]).

Breaking differs from the current approach in terms of the array of correspondence relations it produces. In breaking, the input segment corresponds to each of its output segments, but crucially the output segments do not correspond to each other. The

---

[17] This generalisation is also a potential problem for feature sharing theories. If markedness is evaluated at the segmental level, /akti/ → [ak̲ati] will violate *V^{-high} more than [ak̲iti], so predicting markedness-driven direction reversal.

[18] The absolute nearest consonant may not be available as this would cause multiple correspondence. So, in [pata͟͟pata], the reduplicant *t* does not correspond to the Base *p* because this would result in violations of BR-UNIFORMITY (McCarthy & Prince 1995).

[19] A case where markedness seems to drive long-distance vowel copying is found in Nakanai (Carlson 1997).

[20] A line of inquiry is that faithfulness constraints can be evaluated separately from markedness constraints. For proposals regarding this separation, see McCarthy (1999).

[21] One brute force solution would be to adopt a partially-static definition of the Base: the Base for consonants must be an adjacent segment, while the Base for vowels must be an adjacent segment or a nearby vowel. Of course, this is really a recapitulation of the problem, but it suggests the possibility that full Base-dynamism is incorrect, and that at least a partially static component of the Base-identification mechanism is needed.

diagram below shows correspondence relations in breaking in (1), and those of the present theory in (2):

(30) (i) Breaking: Input: α (ii) Input: α

The fact that there is no direct link between output segments in breaking predicts that any changes to one output segment will not be reflected in the other. In comparison, in the present approach the epenthetic β agrees with its output correspondent γ; α is irrelevant.

Cases of overapplication show that Breaking is an inadequate solution. An example is found in Winnebago, where the vowel [e] ablauts to [a] before certain suffixes: e.g. /mã:čé/ 'cut a piece off' + /ire/ {3rd person} → [mã:čaire], *[mã:čeire]. In cases of copy epenthesis, the ablauted vowel [a] is copied, not the underlying [e]: e.g. /kre/ 'leave returning' + ire {imperative} → [karaire], *[keraire]. Breaking cannot account for this situation since the quality of the epenthetic vowel crucially relies on something to which it bears no relation.

5.0 Concluding Remarks

The proposal we have advanced herein can be stated straightforwardly: epenthetic segments can be related to other output segments via Correspondence. This approach is theoretically minimal: we do not invoke any new mechanisms for evaluating or relating constraints, and the constraints we utilise are trivial extensions of constraint families that have been shown to be necessary elsewhere (McCarthy & Prince 1995). Even so, the correspondence theory of epenthesis enjoys broad empirical coverage, explaining the variation in epenthetic quality found both cross-linguistically and language-internally. In addition, it lends insight into the causes of contextual variation, as well as the repairs employed in such cases.

A great deal of attention was given to the Base of epenthesis – the segments that stand in correspondence with epenthetic elements. We showed that the Base cannot be identified by a 'static' mechanism, but is instead 'dynamic' – the location of the Base can change in different environments. In effect, the identification of the Base reduces to constraint interaction. This opens up the possibility that the the Base of reduplication is similarly determined (*cf* McCarthy & Prince 1994).

Our proposals lead to a more general question: Why should correspondence relations be restricted to holding between an **epenthetic** element and another segment? Could they hold between **any** pair of output segments? After all, if copy epenthesis is due to correspondence, then it seems plausible that smaller scale copying – as found in assimilation and vowel harmony – will also be amenable to a correspondence-based explanation.[22]

[22] We note that some recent work – Krämer 1999 – employs correspondence in explaining vowel harmony.

References

Akinlabi, A. (1993). Underspecification and the phonology of Yoruba /r/. *Linguistic Inquiry* **24,** 139-160.

Alderete, J. (1995). Faithfulness to prosodic heads. Manuscript, Amherst, MA: University of Massachusetts at Amherst.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Sue Urbanczyk. (in press) Reduplication with fixed segmentism. To appear in *Linguistic Inquiry*. [Also Rutgers Optimality Archive #226.]

Anderson, S.R. (1972a). *U*-Umlaut and skaldic verse. In *A festschrift for Morris Halle*, edited by S.Anderson and P.Kiparsky. New York: Holt, Rinehart and Winston.

Anderson, S.R. (1972b). The vowel system of Faroese and the Faroese *Vershärfung*. In *Contributions to generative phonology*, edited by M.Brame. Austin: University of Austin press.

Anderson, S. R. (1974). *The organization of phonology*. New York: Academic Press.

Aronoff, Mark, Azhar Arsyad, Hassan Basri & Ellen Broselow (1987). Tier configuration in Makassarese reduplication. In *CLS 23: Parasession on autosegmental and metrical phonology*, edited by A.Bosch, E.Schiller, and B.Need. Chicago: Chicago Linguistic Society.

Beckman, J. N. (1998). *Positional faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst.

Bickmore, Lee S. (1995). Refining and formalizing the Tahitian stress placement algorithm. *Oceanic Linguistics* **34**, 410-442.

Booij, G. E. (1995). *The phonology of Dutch*. Oxford: Oxford University Press.

Broselow, E. (1995). Skeletal positions and moras. In *The handbook of phonological theory*, edited by J.Goldsmith, 175-205. Cambridge, MA: Blackwell.

Carlson, K. (1997). Sonority and reduplication in Nakanai and Nuxalk (Bella Coola). Manuscript, University of Massachusetts, Amherst. Also Rutgers Optimality Archive #237.

Chafe, W. (1968). The ordering of phonological rules. *International Journal of American Linguistics* **24**, 115-136.

Clements, G. N. (1985). The geometry of phonological features. *Phonology* **2**, 225-252.

Clements, G. N. and E. V. Hume (1995). The internal organization of speech sounds. In *The handbook of phonological theory*, edited by J. A. Goldsmith, 245-306. Cambridge, MA, and Oxford, UK: Blackwell.

Cowan, Hendrick. 1965. Grammar of the Sentani language. Koninklijk Institut voor Taal- Land- en Volkenkunde (KITLV) 47.

Elenbaas, N. (1999). *A unified account of binary and ternary Stress: Considerations from Sentani and Finnish*. OTS: LOT.

Feldman, Harry (1986). *A grammar of Awtuw*. Pacific Linguistics B 94.

Frajzyngier, Zygmunt & Robert Koops (1989). Double-epenthesis and N-class in Chadic. In *Current progress in Chadic linguistics*, edited by Zygmunt Frajzyngier, 233-250. Philadelphia: John Benjamins.

Gafos, A. (1996). *The articulatory basis of locality in phonology*. Doctoral dissertation, Johns Hopkins University.

Gafos, A. (1998). Eliminating long distance consonantal spreading. *Natural Language and Linguistic Theory* **16**, 223-278.

Gafos, A. & L. Lombardi (1997). Consonant transparency and vowel echo. *NELS* 29.
Hargus, S. (1988). *The lexical phonology of Sekani*. New York: Garland Publications.
Harris, J. (1983). *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge, MA: MIT Press.
Hayes, B. (1995). *Metrical stress theory: Principles and case studies.* Chicago: Chicago University Press.
Hume, E. (1992). *Front vowels, coronal consonants, and their interaction in nonlinear phonology*. Doctoral dissertation: Cornell University.
Ka, O. (1994). *Wolof phonology and morphology*. London: University Press of America.
Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Doctoral dissertation, MIT.
Kenstowicz, M. (1981). Functional explanations in generative phonology. In *Phonology in the 1980s*, edited by Didier L.Goyvaerts, 431-444. Ghent: E.Story-Scientia.
Kettunen, L. (1940). Suomen murteet IIIa. Murrekartasto. Suomalaisen Kirjallisuuden Seura. [Finnish dialects IIIa. Dialect atlas. Finnish Literature Society.]
Kitto, Catherine (1997). *Epenthesis and Polynesian loanwords*. MA Thesis: University of Auckland.
Krämer, M. (1999). A correspondence approach to vowel harmony and disharmony. Manuscript: Heinrich-Heine-Universität Düsseldorf.
Krause, S. (1980). *Topics in Chukchee phonology and morphology*. Doctoral dissertation.
Martinez-Gil, F. (1997). Word-final epenthesis in Galician. In *Issues in the phonology and morphology of the major Iberian languages*, edited by Alfonso Morales-Front (ed.). Washington: Georgetown University Press.
McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology*. Doctoral dissertation: MIT.
McCarthy, J. (1982). Prosodic templates, morphemic templates, and morphemic tiers. In *The structure of phonological representations: volume 1*, edited by H. van der Hulst and N. Smith, 191-223. Dordrecht, Foris.
McCarthy, J. (1997). Segmental phonological processes in O.T. Class Lecture Notes. [Available on CD-ROM from the University of Massachusetts].
McCarthy, J. (1999). Sympathy, cumulativity, and the Duke of York gambit. Manuscript, University of Massachusetts, Amherst. Also Rutgers Optimality Archive #315.
McCarthy, J. J. and A. S. Prince (1993). Generalized alignment. In *Yearbook of morphology***,** edited by G. Booij and J. van Marle, 79-153.
McCarthy, J. J. and A. S. Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of NELS 24*, edited by M. Gonzàlez, 333-379. Amherst, MA: GLSA.
McCarthy, J. J. and A. S. Prince (1995). Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics UMOP 18*, edited by J. Beckman, S. Urbanczyk and L. Walsh Dickey, 249-384. Amherst, MA: GLSA.
Metcalfe, C.D. (1975). *Bardi verb morphology*. Pacific Linguistics B30.
Michelson, K. (1988). *A comparative study of Lake Iroquoian Accent*. Dordrecht: Kluwer.

Miner, K. (1992). Winnebago accent: The rest of the data. In *Indiana University Linguistics Club twenty-fifth anniversary volume*, 28-53. Bloomington, Ind.: Indiana University Linguistics Club.

Mithun, M. and H.Basri. (1985). The phonology of Selayarese. *Oceanic Linguistics* 25:210-254.

Moshinsky, J. (1974). *A grammar of southeastern Pomo*. University of California Publications in Linguistics 72.

Myers, Scott (1991). Persistent rules. *Linguistic Inquiry* **22**, 315-44.

Ní Chiosáin, M. (1995). Barra Gaelic vowel copy and (non-)constituent spreading. *WCCFL* **13**, 3-15.

Odden, D. (1995). Tone: African languages. In *The handbook of phonological theory*, edited by J.Goldsmith, 444-475. Cambridge, MA: Blackwell.

Payne, D. L. (1981). *The phonology and morphology of Axininca Campa*. Dallas, TX: Summer Institute of Linguistics.

Rehg, K. and D. Sohl (1981). *Ponapean reference grammar*. Honolulu, University of Hawaii Press.

Roberts, John R. (1987). *Amele.* London: Croom Helm.

Sagey, E. (1986). *The representation of features and relations in nonlinear phonology*. Doctoral dissertation. Cambridge, MA: MIT.

Sapir, E. (1922). The Takelma language of southwestern Oregon. *Bulletin of American Ethology* **40**(2): 1–296.

Shaw, P. (1980). *Theoretical issues in Dakota phonology and morphology*. New York: Garland Publications.

Steriade, D. (1990). Gestures and autosegments: Comments on Browman and Goldstein's paper. In *Papers in laboratory phonology*, edited by J. Kingston and M. Beckman. Cambridge: Cambridge University Press.

Steriade, D. (1995). Underspecification and markedness. In *The handbook of phonological theory*, edited by J. Goldsmith, 114-174. Cambridge, MA, Blackwell.

Urbanczyk, S. (1996). Patterns of reduplication in Lushootseed. Doctoral dissertation: University of Massachusetts, Amherst.

Wilbur, R. (1974). *The Phonology of reduplication*. Doctoral dissertation: University of Illinois.

Wiltshire, Caroline (1998). Extending ALIGN constraints to new domains. *Linguistics* **36**, 423-468.

Woolams, Geoff. (1996). *A Grammar of Karo Batak, Sumatra*. Pacific Linguistics C-130.

Yar-Shater, E. (1969). *A grammar of southern Tati dialects.* Columbia University Publications in Near and Middle East Studies B3. The Hague: Mouton.

Zou, K. (1991). Assimilation as spreading in Kolami. In *1990 Mid-America Conference Papers*, edited by Frances Ingemann, 461-475.

Paul de Lacy
Linguistics Department
South College
University of Massachusetts
Amherst, MA 01003
U.S.A.

<delacy@linguist.umass.edu>
http://www-unix.oit.umass.edu/~delacy

Catherine Kitto
Linguistics Department
130 St.George St
Room 6076
Toronto M5S 3H1
Ontario
Canada

<c.kitto@chass.utoronto.ca>