The Phonetic Nature of Niuean Vowel Length*

Nicholas Rolle
University of Toronto

In this paper, I argue that in Niuean, long vowels are underlying sequences of two qualitatively identical vowels, extending an analysis laid out in Taumoefolau (2002) on the related Tongic language Tongan. She argues that a long vowel will result predictably when stress falls on the first element of a double monophthong vowel sequence. Thus vowel length in these languages is a phonetic, not phonological, phenomenon. This account for vowel length is worked into an Optimality Theory (OT) framework, which crucially constrains RH-CONTOUR and ALL-FT-RIGHT above a constraint against double articulation, DIST(INCT). In addition, I provide a review of the literature concerning vowel length across Polynesian languages, and conclude that this analysis cannot be extended to all members of the Polynesian language family.

1. Introduction

Across Polynesian languages, the presence of long vowels is uncontroversial, though whether these are only surface long vowels (i.e. phonetic) or underlying long vowels (i.e. phonemic) remains a contentious topic. In this paper, I outline an analysis of Tongan long vowels laid out in Taumoefolau (2002), and later supported by Anderson & Otsuka (2006), and extend this analysis to the related Tongic language Niuean. It is argued in Taumoefolau (2002) that Tongan long vowels are underlying identical vowel-vowel sequences (/V_iV_i/), whose surface manifestation is predictably a long vowel if stress falls on the first vowel of this sequence, and a double articulated vowel if stress falls on the second. This analysis is laid out in detail in section 3, preceded by a section describing some basic facts of Tongan phonology, and a review of the previous analyses of Tongan vowel length. In section 4, I extend this analysis of vowel length to Niuean, providing additional evidence of their underlying /V_iV_i/ form from a dictionary survey of permissible CV templates, and evidence from nativized loanwords. I then incorporate this analysis of Niuean long vowels into an Optimality Theory (OT) framework. In doing so, I provide an explanation of the allophonic distribution of long vowels and double articulated vowels through simultaneous stress placement and phonetic realization, based on the ordering of violable constraints (RH-CONTOUR, ALL-FT-RIGHT >> DIST¹). In section 5, I discuss vowel length across the Polynesian language family, attempting to extend this analysis of Niuean and Tongan long vowels to their less related language family members. The analysis extends well to Rennellese-Bellonese of the Samoic branch, though cannot be extended to Hawaiian and Tahitian of the Eastern branch or Tokelauan of the Samoic branch, whose long vowels have “phonemized”. I attempt to find a precise point in the historical development of the Polynesian languages of this “phonemization”, though this investigation is confounded by the presence of free

---

* Many thanks go to Keren Rice, Diane Massam, and Donna Starks for their constructive criticism, and also to Chiara Frigeni, Catherine Macdonald, Sara Mackenzie, Amarpaul Dhillon, Isaac Gould, Ewan Dunbar, Michelle St-Amour, and Derek Denis.

¹ DIST = DISTINCT. This is a coined constraint, described in section 5, against two segments of the same quality being adjacent.
variation between long vowels and double articulated vowels in certain words in Maori, Samoan, and even Tongan. It is the hope of this paper to provide insight into the role of prosodic processes such as stress in the formation of surface realizations of underlying phonemic units, and further, to provide a valuable case study of the genesis of contrastive vowel length in natural language.

2. Previous Analyses of Tongan Vowel Length

This section outlines some basic facts of Tongan phonology, and lays out the previous analyses on Tongan vowel length.

Tongan contains the following qualitatively distinct consonants and vowels: /p t k ʔ f v s h l m n ŋ a e i o u/\(^2\). Codas are not permitted in the language, and there is a one-to-one relationship between morae and syllables underlyingly\(^3\), a fact which will become important in section 3. Any monomoraic words (e.g. ki ‘to’) are grammatical words, and cannot stand on their own (i.e. they must be phonologically incorporated into an adjacent word); there are no lexical monosyllabic words (Taumoefolau 2002:344). Thus, the basic minimal word is of the form (C)V(C)V.

Prior to Taumoefolau (2002), many authors have provided or adopted accounts of vowel length in Tongan. Churchward’s (1953) reference grammar of Tongan distinguishes three “degrees of vowel length”: normal length <a>, the long vowel <ä> (now written as <ā>), and the double vowel <aa> (Taumoefolau 2002:346). This last grapheme is best described as a double articulated vowel, in which both vowels are pronounced separately.

(1) Tongan three-way distinction:

a. Normal: \(<\text{mana}>\) ‘supernatural’

b. Long: \(<\text{mā}>\) ‘bread’

c. Double: \(<\text{maama}>\) ‘light’

Churchward (1953) proposes that long vowels and double articulated vowels are separate phonemes, entailing that their distribution is not predictable, nor complementary.

Despite the appearance of long vowels, many authors have proposed that vowel length is not a true contrastive distinction in Tongan. Tongan long vowels therefore have been treated in two ways: as separate phonemes, and thus underlying long vowels (2a), or as only surface long vowels, and thus are phonetic in nature (2b). In the illustration below, ‘V:’ indicates a long vowel, and ‘V/V’ indicates a sequence of two short identical vowels.

(2) Tongan Vowel Length

a. Phonemic Distinction:

\[
\begin{array}{ccc}
/V/ & /V:/ & /V/V/ \\
[V] & [V:] & [VV] \\
\end{array}
\]

b. Allophonic Distribution:

\[
\begin{array}{ccc}
/V/ & /V:/ & /V/V/ \\
[V] & [V:] & [VV] \\
\end{array}
\]

The analysis in (2b) has been adopted by the majority of authors working on Tongan since Churchward (1953). Biggs (1971:469) claims that in Polynesian languages in general, “contrastive vowel length” should best be thought of as “geminate clustering”. With specific reference to Tongan, Feldman (1978) states that despite the fact that Tongan distinguishes long vowels and double

\(^2\) Churchward (1953), among other places.

\(^3\) That is, syllables may not branch. They may do so only within a fusion process, described later, which is a phonetic phenomenon.
articulated vowels in its orthography, these two may still be thought of as two realizations of the same phonemic sequence, /\text{ViVi}/:

The Tongan orthography is phonemic in all respect save one...: it distinguishes between a sequence of two like vowels with the stress on the second, and one where the stress distributes itself over both; VV and [Vː], respectively.

In addition, De Chene (1979:44) claims that in Tongan, “stress should in fact be described as falling on the penultimate mora, with a later rule distributing it over a long vowel”, though does not discuss what this “later rule” would be.

Poser (1985) also supports analyzing Tongan as not having underlying contrastive vowel length, noting that if two identical vowels are adjacent, whether they “form a single syllable” (i.e. a long vowel) or “belong to two different syllables” (i.e. a double articulated vowel-vowel sequence) will be predictable (Poser 1985:266). He calls this formation of a single syllable “syllable fusion”, and notes that one of the conditions of this fusion is that “stress must not fall on the second of the two vowels”, a stipulation discussed in depth in Taumoefolau (2002). Also, Schütz (2001) states that Churchward did not realize that there was no phonemic contrast between long vowels and double vowels, and further notes that long vowels behave exactly as diphthongs in that under “certain conditions, (they may) function as the nucleus of a single syllable” (Schütz 2001:320). Schütz cites Clark (1974:103), as well, in claiming that long vowels are best treated as vowel-vowel sequences (Schütz 2001:311). Thus, within the literature on Tongan vowel length since Churchward (1953), there has been general consensus that vowel length is phonetic, and not phonemic. Taumoefolau (2002) develops this theory into a comprehensive analysis, laid out in section 3 below.

3. Taumoefolau’s (2002) Analysis of Tongan Vowel Length

Taumoefolau (2002) proposes an analysis of Tongan surface long vowels as underlying sequences of two identical adjacent vowels. She builds upon a large amount of literature, laid out above, which runs counter to the original claim made in Churchward (1953), which distinguished three degrees of vowel length: normal (/\text{V}/), long (/\text{Vi}/), and double (/\text{ViVi}/) (see example 1). Taumoefolau concludes that this three-way distinction is quite unusual for a natural phonological contrasting system, rejecting the phonemic distinction between “long vowel” and “double vowel”. She claims that these are “merely different phonetic realisations of two consecutive like vowels”, which coincide with “two different kinds of stress groups” (Taumoefolau 2002:346). Thus, what accounts for the distinction is not a contrastive length distinction, but rather where stress groups (i.e. feet) are formed.

Vowel length, then, is an incidental phenomenon, falling out from the stress algorithm of the language. Double articulated vowels surface when stress is placed on the second segment of the like vowel sequence, whereas long vowels surface when stress is placed on the first segment of this sequence. This is illustrated in examples (3a–b) below.
(3) **Stress Placement**

a. long vowel:

/\text{CV}_1\text{V}_2/

/\text{maa}/

footing  (ma.a)
stress  (má.a)

[má:]  ‘bread’

b. double articulated vowel:

/\text{CV}_1\text{V}_2\text{CV}_3/

/\text{maama}/

footing  (ma.a.ma)
stress  (maá.a.ma)

[maáma]  ‘light’

Thus, above, we see two words, *ma:* /\text{maa}/ ‘bread’ and *maama* /\text{maama}/ ‘light’. These words are first parsed into stress groups, and then stress is assigned. The stress algorithm of Tongan is stated as falling on a stress group, which Taumoefolau calls “a sequence of two or three syllables at a time which sustains a single stress on the penultimate syllable” (Taumoefolau 2002:341). In both (3a) and (b), stress falls on the underlying penultimate syllable, though the phonetic results are different. Though both words contain the same /\text{aa}/ sequence, when stress falls on the first segment of the sequence, a long vowel results. Thus, this analysis of long vowels complements the theory of “syllable fusion” proposed in (Poser 1985:266), which states that when the first segment of an identical vowel vowel sequence is stressed, the two separate syllables “fuse” into a long vowel. For longer words, such as five to eight syllable words, this situation is complicated by the fact that they may be split up into multiple different stress groups. For example, a five syllable word may be split up into two different stress groups, either a 2 syllable foot followed by a 3 syllable foot, or vice versa (i.e. 3-2). Within the stress group, however, it is always the penultimate syllable which is stressed, and most often, stress groups are determined by morphosyntactic boundaries (Taumoefolau 2002:345). Taumoefolau provides an example of the derived word *maamaloa* ‘brilliant’, illustrated below.

(4)  

/\text{CV}_1\text{V}_2\text{CV}_3\text{VV}_4/  

/\text{maamaloa}/  

footing  (ma.a.ma)(lo.a)  
stress  (maá.a.ma)(lo.a)  

[maámalóa]  ‘brilliant’

Taumoefolau (2002) simplifies the stress algorithm of the language. A previous analysis about stress placement in Tongan stated that stress falls on the penultimate syllable of the word, with

---

4 Studies in stress interaction across morphological and phrasal borders are not discussed at present, though they certainly are important to vowel length, especially with certain monosyllabic suffixational clitics. These clitics cause a reassignment of stress and thus alter length: /\text{maa}/ [má:] ‘bread’ + /\text{ni}/ [ni] ‘this’ (enclitic) = /\text{maani}/ [maáni] ‘this bread’.
the stipulation that if the final syllable is long, stress falls there, overriding the penultimate stressing rule. Under her analysis, however, this may be conflated as simply a difference of stress placement, which she claims is “ultimately a difference in stress group type”, which gives “the impression that ‘long vowels’ and ‘double vowels’ are distinct” (Taumoefolau 2002:346). Thus, stress does not fall on the long vowel per se, but rather helps to create the long vowel in the first place. This is an important reversal, implying that stress applies to (or can apply to) the abstract phonemic representation of a form, and not necessarily to its final phonetic realization. Stress is therefore an active agent in creating surface forms from underlying forms, interacting in a much more connected way with other phonological processes.

A final note, it falls out from this analysis that there is a one-to-one relationship between the mora and the syllable. Each syllable is mono-moraic before the stress rules are applied; a syllable may only be bimoraic if there is “syllable fusion” after stress is assigned. This is important because it indicates that in Tongan, there is no underlying heavy/light syllable distinction, thus allowing stress to freely fall on the penultimate syllable of the stress group (i.e. foot). Taumoefolau (2002) is supported in this regard by Anderson and Otsuka (2006), who argue, from syllable theory and acoustic measurements, that long vowels in Tongan should be considered a sequence of two syllables, and that the syllable and the mora are coextensive (Anderson and Otsuka 2006:41). Further, Macdonald (2007:59) argues that syllables are either “redundant, or they are irrelevant to stress assignment”; thus the only prosodic unit necessary is the mora. She also cites, though does not adopt, an OT analysis in Prince & Smolensky (2004), which argues that stress assignment and syllabification occur simultaneously (Macdonald 2007:61). This theory does not require a separate syllable fusion process to occur post-stress, as the input is evaluated at only one level. I will return to an OT framework in section 4.2; I now extend the analysis in Taumoefolau (2002) to the related Tongic language Niuean.

4. Niuean Vowel Length

The analysis of Tongan vowel length laid out in Taumoefolau (2002) can be extended to the Polynesian language Niuean. Tongan and Niuean are genetically closely related, being the only two languages of the Tongic branch of the Polynesian family. Niuean contains the following segmental inventory: /p t k f v s h l m n ŋ a e i o u/ (Sperlich 1997:2). The only segment which Tongan has retained which Niuean has not is the glottal stop /ʔ/. Sperlich (1997:9) states that the common structure of syllables is (C)V(V), and that root words are typically mono-, di-, or trisyllabic. There are no consonant clusters, and codas are not permitted. Sperlich maintains that Niuean contains long and short vowels, though he does not indicate the phonemic or phonetic status of this distinction. I apply two diagnostics to argue that long vowels are two underlying vowels, extending Taumoefolau (2002), and thus arguing that vowel length in Niuean is phonetic in nature. These diagnostics involve a dictionary survey of permissible CV shapes of words and long vowel distribution, and a survey of nativized loan words.

4.1 CV Template Constraints

If there is a true phonemic distinction between long and short vowels in Niuean, there should not be complementary distribution of double vowels and long vowels. This is the exact opposite of what is found. Taumoefolau’s analysis predicts that a long vowel can never result from stress falling on the second element of a V_iV_i sequence. It also predicts that stress will never fall on the first element of a V_iV_i sequence and produce a double articulated vowel sequence. Stress in Niuean “almost invariably” falls on the penultimate syllable of a word (Sperlich 1997:10), though Sperlich
makes the exception that long vowels in any position will attract stress. If we assume that long
vowels are underlyingly two syllables (coextensive with the mora), then a long vowel will only result
if stress falls on the first of these two. Therefore, when stress is assigned onto the penultimate
syllable of a so-called measure group, or foot, the syllable shape of the word will affect whether a
long vowel or a double vowel results. Words of the phonetic shape [CVCV_i:] and [CVi:CV] are not
expected to exist in Niuean, because it would indicate that an unnatural stress algorithm produced
them; words of the shape [CVCV_i:] are expected. This is illustrated in (5) below.

(5) a. /CV_i:CV/  
footing (CV_i:CV)  
stress (CV_i:CV)  
 -ok[CV_i:CV]  
fusion *-[CV_i:CV]  
b. /CVCV_i:/  
footing (CV.CV_i:)  
stress (CV.CV_i:)  
 -ok[CVCV_i:]  
fusion *-[CVCV_i:]  

Thus, it follows from Taumoeefolau (2002) that if a word has the underlying form /CV_i:CV/, the
vowel sequence will always form a double articulated vowel, because stress will invariably fall on
the penultimate syllable. If a word has the underlying form /CVCV_i:/, the vowel sequence will
become long, due to penultimate stress falling on the first vowel. Based on a dictionary survey of
Sperlich (1997), this is precisely what is found. There are no words which have the form [CVi:CV] or
[CVCV_i:CV], though there are many words which have the form [CVCV_i:] and [CVV_i:CV].
Derivations of the words iina ‘to frighten someone’, fita: ‘to have already done something’, la:ko: ‘to
babble’, and pe:kau ‘wing’ are provided below:

(6) a. /V_i:CV/  c. /CV_i:CV,CV/  
footing (i.i.na)  footing (la.a)(ko.o)  
stress (i.i.na)  stress (là.a)(kó.o)  
  ok[iina]  fusion  ok[là:ko:]  
  *[i:na]  *-[làakóo]  
b. /CVCV_i:/  d. /CVV_i:CV/  
footing (fi.ta.a)  footing (pe.e)(ka.u)  
stress (fi.tà.a)  stress (pé.e)(ká.u)  
  ok[fitá:]  fusion  ok[pè:kaú]  
  *[fitáa]  *[pèekáu]  

NICHOLAS ROLLE
Thus, a dictionary survey provides us with evidence that Taumoefolau’s analysis is able to be extended to Niuean. The exact forms which are predicted to be impossible phonetic realizations of words by her analysis are absent in Niuean.\footnote{Further evidence for long vowels as /V\textsubscript{i}V\textsubscript{i}/ sequences comes from the fact that when words with long vowels or double vowels undergo derivation or affixation, their form may change. For example, Sperlich (1997:264) notes that with \textit{peehi} ‘to press, push down’, the “rearticulated ee in \textit{peehi} changes to a long e: when suffixed”. A study of derived and suffixed forms is outside of the scope of this analysis, though at present provide no contradiction to the present analysis. For more, see footnote 4.}

4.2 Niuean Nativized Loanwords

Further evidence that Niuean does not possess a phonemic length contrast comes from English loan words that have been nativized in Niuean. This diagnostic is based on Taumoefolau (2002:347) and Schütz (2001:314), both of whom use loanwords as evidence for their analyses. Schütz states succinctly that loanwords are not “aberrations”, but rather “simply fall into the available slots” of a language, and thus bring “the phonological rules of the language into sharper focus” (Schütz 2001:314). Because these loan words have been nativized, they must fit into the natural stress algorithm of Niuean. If vowel length is reliant on stress placement, then it is predicted that only certain long vowels from English may be preserved.

The majority of British English varieties and New Zealand English, from which English words are borrowed, can be described as being r-less, without /\textipa{ɹ}/ in a coda position. In order to preserve the mora which has consequently been left vacant, the preceding vowel is lengthened. Thus, /\textipa{f\textipa{ɑ}ɹ}/ ‘far’ becomes \[\textipa{f\textipa{ɑ}:}\]. When words with a long vowel are adopted into the Niuean lexicon, the perceived length is preserved when it is possible to be created from the stress algorithm of the language, i.e. when the first vowel of the underlying identical vowel-vowel sequence is able to be stressed. This is illustrated below:

\begin{tabular}{ll}
(7) & a. English & Niuean \\
& “cartoon” [k\textipa{a}:tun] & /kaatuni/ \\
& & (k\textipa{a}.a)(tu.ni) footing \\
& & (k\textipa{a}.a)(tu.ni)\footnote{Though each foot receives a stress, primary stress (word stress) falls on the final, rightmost foot. Further, an epenthetic \textipa{i} is attached as Niuean does not permit codas} stress \\
& & [k\textipa{a}:t\textipa{u}n\textipa{i}] \\

& b. English & Niuean \\
& “jar” [d\textipa{ʒ\textipa{a}:}] & /siaa/ \\
& & (si.\textipa{a}.a) footing \\
& & (si.\textipa{a}.a) stress \\
& & [si\textipa{a}:] \\
\end{tabular}

Thus, we see in the examples in (7) that the perceived long vowel is preserved if the segments may be grouped into stress groups which will allow stress to fall on the first vowel of a /V\textsubscript{i}V\textsubscript{i}/ sequence. In certain nativized loan words, it is impossible to create stress groups which will allow a long vowel to result. The result is a double articulated vowel.\footnote{There are other cases where the long vowel of the English word becomes a single vowel in Niuean, though this is not studied at present.}
If Niuean had a true vowel length contrast, then a long vowel should be able to be preserved in any position. Above, the long vowels from “marker” and “farm” are not preserved; if they were preserved, it would violate the penultimate stressing algorithm of the language, indicated with the starred examples above. A nativized loan word must fall within the phonological structure of Niuean, and it is only certain long vowels from loan words that may be preserved, dependent upon where they are grouped in stress assignment. This gives evidence that long vowels are phonetic in nature in Niuean, arising from an interaction between stress and identical vowel sequences.

5. Niuean Vowel Length within OT

Taumoefolau’s analysis elegantly accounts for the distribution of long vowels in Tongan and, as I have argued, Niuean through an uncomplicated penultimate-stressing strategy. In this section, I propose an Optimality Theory (OT) analysis to account for the distribution of long vowels [Vː] and double articulated vowels [ViVi]. An OT account is especially attractive as it eliminates the entire “timing” issue of syllabification and stress. Under OT, all possible candidates are evaluated simultaneously at a single level (Kager 1999:175). This will eliminate any need to posit a post stress process of “syllable fusion”, as proposed in Poser (1985). In order to determine which candidate will be selected, we must establish a series of ranked constraints. This provides an explanation for the allophonic distribution.

The issue of allophonic vowel length can best be thought of as a resolution of a marked form. A sequence such as [ViVi] (the double articulated monophthong) is more marked and than [Vː], a phenomenon which can be subsumed within the OT constraint DIST(INCT). This constraint expresses that segments in a sequence must be qualitatively distinct from segments which they are immediately adjacent to. Thus, the double articulation of a vowel means two segments pronounced in such a sequence are not distinct from one another; coalescence of such a sequence is less marked than double articulation. The reasons behind this involve acoustic properties of perception and production, not discussed here (for a thorough analysis of vowel hiatus within OT, see Casali 1998).

---

8 There are further ways in which a long vowel may theoretically be preserved which I do not discuss. None of these alternatives provides a counter example to the analysis being proposed, however. Taumoefolau (2002:347) describes the multiple ways in which a Tongan speaker adopts the bimoraic diphthong [aj] in “MacGyver” to fit within the phonological rules of their grammar: [mákaíva], [makáivá:], and [makéva].
We must contend with the fact, though, that we do find double articulated vowel sequences in Niuean and Tongan. As a result, it is proposed that this constraint DIST interacts with the stress constraints of the language which are ranked higher than it in order to ensure the proper output. Since these constraints are ranked higher, it means that candidates which violate these stress constraints are eliminated earlier than those who do not. Crucially, we need DIST ranked above the constraints RH-CONTOUR and ALL-FT-RIGHT (RH-CONTOUR, ALL-FT-RIGHT >> DIST). RH-CONTOUR states that a foot must end in a strong-weak contour at the moraic level. This means that the peak of the foot, being that mora which is stressed and therefore strong, must be followed by one weak mora. This constraint is proposed by Kager (1999:174) based on cross-linguistic preferences; he provides the following diagram to illustrate the constraint. [Note that the star above a mora indicates that it is stressed, and therefore the peak of the foot.]

![Diagram illustrating stress constraints](image)

We see in (9b) and (10b) that those feet whose strong mora are followed by two weak morae or no weak morae, violate RH-CONTOUR. Therefore, ranking this constraint highly in the grammar eliminates the possibility for words of the form [CV;CV]: a long vowel followed by a short vowel matches the forms in (9b.i), which violates RH-CONTOUR. Thus, a double articulated vowel surfaces here, despite violating DIST. Further, it is necessary to account for the fact that stress is computed from the ridge edge. This can be accomplished with ALL-FT-RIGHT, which states that every foot stands at the right edge of the prosodic word, built leftwards (Kager 1999). ALL-FT-RIGHT guarantees that the rightmost syllable is placed within a stress group.

We require the following list of constraints in order to eliminate other ungrammatical candidates; this list includes other stress based and non-stress based constraints, adopted from (Kager 1999).

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIST</td>
<td>segments must be qualitatively distinct from segments which they are immediately adjacent to</td>
</tr>
<tr>
<td>RH-CONTOUR</td>
<td>a foot must end in a strong-weak contour at the moraic level</td>
</tr>
<tr>
<td>ALL-FT-RIGHT</td>
<td>every foot stands at the right edge of the prosodic word</td>
</tr>
<tr>
<td>RIGHTMOST</td>
<td>the head foot is the rightmost in a PrWd</td>
</tr>
<tr>
<td>FT-BIN^9</td>
<td>feet are binary under moraic or syllabic analysis</td>
</tr>
<tr>
<td>MAX-µ-IO</td>
<td>input morae must be present in the output</td>
</tr>
<tr>
<td>DEP-IO</td>
<td>output segments have input correspondents</td>
</tr>
<tr>
<td>WSP</td>
<td>Heavy Syllables are stressed</td>
</tr>
<tr>
<td>RhITYPE=T</td>
<td>Feet have initial prominence (i.e. GRWD=PRWD: a grammatical word must be a</td>
</tr>
</tbody>
</table>

---

^9 In adopting an OT account of Niuean long vowel and double vowel distribution, I posit that all stress groups are bisyllabic, against Taumoefolau (2002:344), which allows the existence of trisyllabic feet. The leftmost syllable is left unfooted. This is of course complicated by the presence of 5 to 8 syllable words, discussed earlier. It is tentatively speculated that morphological boundaries are very powerful, allowing syllables to remain unfooted at the left edge of their boundary, as well as that boundary of the word. This, crucially, happens only at the left edge.
generates trochees not iambs) proodic word (i.e. must be stressed, footed)

|MORA=SYLL: a mora is coextensive with the syllable (i.e. each syllable is monomoraic)|

I provide tableaux for the Niuean words *fita*: ‘to have already done something’ and *iina* ‘to frighten someone’ (repeated from examples (6a-b)), and the Tongan words *ma*: ‘bread’ and *maama* ‘light’ (repeated from examples (3a-b)), showing the different surface patterns.

<table>
<thead>
<tr>
<th>T1. /fitaa/</th>
<th>GrWd</th>
<th>RhType</th>
<th>WSP</th>
<th>Dep-IO</th>
<th>Max-μ-IO</th>
<th>Ft-Bin</th>
<th>Right-most</th>
<th>All-Ft-Right</th>
<th>Rh-Contour</th>
<th>Dist</th>
<th>Mora=Syll</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. fi.(tá.a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (fi.ta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. fi.(tá:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. fi.(tá.ta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. (fi.tá:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. fi.ta.a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. (fi.ta:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T2. /iina/</th>
<th>GrWd</th>
<th>RhType</th>
<th>WSP</th>
<th>Dep-IO</th>
<th>Max-μ-IO</th>
<th>Ft-Bin</th>
<th>Right-most</th>
<th>All-Ft-Right</th>
<th>Rh-Contour</th>
<th>Dist</th>
<th>Mora=Syll</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. i.(í.na)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (í.na)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (í:.na)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. i.(tí.na)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. (í:.na)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. i.i.na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. (í:.ní)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. (í:.ńa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T3. /maa/</th>
<th>GrWd</th>
<th>RhType</th>
<th>WSP</th>
<th>Dep-IO</th>
<th>Max-μ-IO</th>
<th>Ft-Bin</th>
<th>Right-most</th>
<th>All-Ft-Right</th>
<th>Rh-Contour</th>
<th>Dist</th>
<th>Mora=Syll</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (má.a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (má)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (má:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. (má.ta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. (má:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. ma.a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T4. /maama/</th>
<th>GrWd</th>
<th>RhType</th>
<th>WSP</th>
<th>Dep-IO</th>
<th>Max-μ-IO</th>
<th>Ft-Bin</th>
<th>Right-most</th>
<th>All-Ft-Right</th>
<th>Rh-Contour</th>
<th>Dist</th>
<th>Mora=Syll</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ma.(á.ma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (má.ma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (má:.ma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ma.(ta.ma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. (má:.ma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. ma.a.ma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. (má:.ńa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. (ma:.ńa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The output of the form /fitaa/ is [fita:], despite it incurring a violation of MORA=SYLL. All other candidates violate a highly ranked constraint, such as [fitaa], which violates DIST. In T2, the optimal

---

10 An OT account therefore explains why a syllable is coextensive with a mora in all instances except long vowels (and possibly diphthongs).
output for /iina/ is [iina]. The candidate [(i:.na)]\textsuperscript{11} violates RH-CONTOUR, as a heavy syllable \( i: \) is followed by a light syllable na within the same foot (Kager 1999:173-174)\textsuperscript{12}. If we posit that \( i: \) forms its own foot to the exclusion of na, [(i:.na)], then we violate ALL-FT-RIGHT, as we would then have an unfooted syllable at the right edge of the word, for which we have no additional proof of in Niuean or Tongan. If the two Tongic languages contain the above constraint ranking, then it would accurately predict when a double articulated vowel will surface and when a long vowel will.

6. Vowel Length across Polynesian Languages

In this section, I provide a short overview of what previous Polynesianist material has stated with regard to vowel length, followed by an attempt to extend the analysis in Taumoefolau (2002) to other languages of the Polynesian family. It is found that the analysis can be extended to Rennellese-Bellonese [Samoic], though it cannot be extended to Hawaiian [Eastern], Tahitian [Eastern], and Tokelauan [Samoic]; it is uncertain whether it can be extended to Maori [Eastern] and Samoan [Samoic], due to the presence of free variation between long vowels and double articulated vowels in certain forms.

6.1 Background on Polynesian Vowel Length

The Polynesian branch of the Oceanic language family [Austronesian] is a group of languages spoken in the Pacific Ocean as north as Hawaii (Hawaiian) and as east as Easter Island (Rapanui). It is estimated that the earliest Proto-Polynesians settled near Samoa and Tonga sometime in the second millennium B.C.E. (Krupa 1973:14). The Polynesian languages split not long after this initial settlement, branching into the Proto-Tongic and Proto-Nuclear Polynesian sub-groups. Proto-Nuclear further split into Proto-Samoic and Proto-Eastern; these two sub-branches include the bulk of the family’s languages\textsuperscript{13}.

Most Polynesian languages have been analyzed as having contrastive vowel length, meaning, reiterated from above, that the occurrence of both long vowels and short vowels is unpredictable, and each acts as a phonemic unit able to combine with other phonemic units to form lexical items. This analysis has been put forward, or adopted, by Sperlich (1997) for Niuean [Tongic] and Churchward (1953) for Tongan [Tongic], as discussed above, and by Elbert & Pukui (1979) for Hawaiian [Eastern], Bickmore (1995) for Tahitian [Eastern], Moyse-Faurie (1997) for Futunien [Samoic], and Cablitz (2006) for Marquesan [Eastern]. It is a general convention of Polynesian languages to mark long vowels with a macron, as in the name of the language Māori (IPA: [mɑ:ɔɾi])\textsuperscript{14}. No author, though, has ever claimed that all Polynesian languages have a true short/long vowel contrast, and most within the field have ignored whether long vowels are merely surface long vowels, or underlying contrastive phonemic units.

There have been many authors who have analyzed long vowels as a phonetic sequence of two like vowels. Most have focused on Tongan, as described at length in sections (2–3), though not all. Biggs (1971:469) makes a general claim about Polynesian languages that contrastive vowel length is

\textsuperscript{11} I include stress group brackets when it is necessary, though always include them for completeness sake within OT tableaux.

\textsuperscript{12} Kager (1999) cites Hayes (1995) on quantitative asymmetry between iambs and trochees, where ‘uneven’ iambs are preferred to ‘even’ ones, though ‘even’ trochees are preferred to ‘uneven’ ones. Thus a stressed light-unstressed light syllable sequence is preferred to a stressed heavy-unstressed light one ( (\( LL \)) > (\( HL \)) ).

\textsuperscript{13} There are further more specific sub-branches, not discussed here.

\textsuperscript{14} This macron usage is not as widespread in informal writing, though.
more accurately interpreted as “geminate clustering”, though he makes no reference to stress.\(^{15}\) Krupa (1973) states that long vowels across Polynesia are “preferably treated as sequences of two identical vowels” which would “make the prediction of stress simpler” (Krupa 1973:52). Earlier, Krupa (1968:27) provided quite a bit of evidence in support for analysing Maori [Eastern] long vowels as “sequences of two phonemes”, arguing based mostly on historical and statistical (i.e. frequency) data. In addition, with resepect to Samoan [Samoic], Hovdhaugen (1990:95) states that “long vowels in Samoan, as in most Polynesian languages, are usually interpreted phonemically as a sequence of two consecutive like vowels”. Hovdhaugen (1990:96) also cites Pawley (1966), who assumes that in Samoan “long vowels [are] phonologically a sequence of two identical short vowels”, and adds that,

> If long vowels are interpreted as a sequence of two short vowels (i.e. \(V:=VV\)), then stress rules can be simplified to assigning stress to the penultimate vowel of the word.

Moreover, a short note in Elbert (1975) on Rennel-Bellonese [Samoic] makes casual reference to analyzing long vowels as allophonically long (i.e. only on the surface). Thus, we see that the issue of vowel length remains a contentious topic within the Polynesianist literature. If vowel length is phonemic in a Polynesian language, we must determine whether this is an inherited property from an earlier proto-language, or a divergence restricted to the individual language.

There is evidence to suggest that the proto-Polynesian language did not contain a phonemic vowel length contrast. Proto-Tongic is a direct descendent of proto-Polynesian (Clark 1976). I have argued that the only two languages within the Tongic branch, Niuean and Tongan, do not display phonemic vowel length contrast, therefore I assume that proto-Tongic did not display a length contrast either. Given that proto-Tongic descends directly from proto-Polynesian, it is unlikely that proto-Polynesian contained a vowel length contrast, as it would imply that a less marked form (a long vowel) developed into a more marked form (a double articulated vowel). Further, Anderson & Otsuka (2006:40) cite Hayes (1995) in claiming that rules of foot construction may not split syllables. Thus it is theoretically impossible under this assumption for a long vowel to be split by any rule of foot construction into two short vowels. Therefore, it is unlikely that proto-Polynesian had long vowels, as this would imply that they split into double vowels in certain environments in proto-Tongic.

There is much evidence to suggest that many long vowels and double vowels in Niuean, in fact, are derived from a \(CV,CVi\) sequence in which the intervocalic consonant has been lost. Examples (11a–b) illustrate this below, taken from Sperlich 1997:

\[(11) \quad \text{a. } *\text{sege} > [\text{he:}] \quad \text{b. } *\text{maqala} > [\text{maala}]\]

‘locust like insect’ ‘bush garden’

Further, Proto-Polynesian came from Proto-Oceanic which Lynch (2000) claims as not having a phonemic vowel length distinction. Also, Fijian [Central Pacific: Fiji] is the closest relative to the Polynesian family of the Central Pacific branch of the Oceanic Languages, and does contain contrastive vowel length (Schütz 1999; Scott 1948)\(^{16}\). Given that, proto-Polynesian did not contain contrastive vowel length, if a daughter language does contain such a contrast (as claimed by the

---

\(^{15}\) N. C. Scott (1948), as cited in Schütz (1999:142), treats long vowels as geminates in Fijian, as well, a language closely related to the Polynesian Family.

\(^{16}\) Schütz (1999) claims that “although long and short elements contrast in penultimate position in Hawaiian, obligatory vowel and diphthong-shortening in this position in Fijian neutralizes any such contrasts” and also that “Fijian allows no vowel-length contrast without a change in accent as well”.

12
authors outlined above), then it is likely to have developed in the Samoic and/or Eastern Polynesian sub-branches. This issue will be explored as I attempt to extend Taumoefolau’s analysis to other Polynesian languages.

6.2 Successful Extension

Taumoefolau’s (2002) analysis can be extended to account for the presence of long vowels in the Samoic language Rennellese-Bellonese. Elbert (1975) states that this language has both single and double vowels, where the double vowel is allophonically long when the regular penultimate stress falls on the first vowel of the sequence, but is “rearticulated” when it falls on the second vowel (Elbert 1975:xvi). Two examples are provided below:

(12) a. /CV,V,CVCV/
    /haahine/
    footing (ha.a)(hi.ne)
    stress (hà.a)(hí.ne) *(ha.à)(hi.ne)
    [hà:híne]  *[haàhíne]
    ‘woman’

    b. /CV,V,CVCV/
    /ghaagho/
    footing gha.(a.gho)
    stress gha.(á.gho) *(ghá.a.gho)
    [ghaágho]  *[ghá:gho]
    ‘to make twine’

Thus, when penultimate stress falls on the first vowel of a double vowel sequence (12a), they merge into a long vowel, and when it falls on the second (12b), they surface as a double articulated vowel. In (12a), it would be impossible to produce a double articulated vowel, as it would mean that there was no penultimate stress of a foot (stress group); in (12b), a long vowel is impossible, because it would imply that the antepenultimate vowel is irregularly stressed. Rennellese-Bellonese long vowels are accounted for easily within Taumoefolau’s (2002) analysis, though the related languages Tahitian, Maori, Hawaiian, Samoan, and Tokelauan cannot.

6.3 Unsuccessful Extension

Long vowels in Tahitian and Hawaiian of the Eastern branch and Tokelauan of the Samoic branch cannot be accounted for with Taumoefolau’s analysis. This is due to the fact that the sequence [CVːCV] is found in these languages. In section 4.1, I showed that in Tongan and Niuean, there is complementary distribution between long vowels and double articulated vowels. This distribution is determined by the stress algorithm of the language, illustrated by example (13) below, repeated from example (5) above.

---

17 This can also be interpreted as (gha.a.gho) with no repercussions to the analysis. See fn 9 for more on this topic.
However, in Tahitian, Hawaiian, and Tokelauan, the form [CV:CV] which is predicted to not occur according to Taumoefolau’s (2002) analysis, does in fact occur. This is attested by the forms in (14).

(14) a. Tahitian (Bickmore 1995):
   má:ha ‘satisfied’, vá:hi ‘place’

b. Hawaiian (Elbert & Pukui 1979:35)\textsuperscript{18}:
   ká:ne ‘man’

c. Tokelauan (Office of Tokelau Affairs 1986):
   fá:o ‘seize, snatch’

Taumoefolau’s analysis predicts that these forms would surface as maáha, vaáhi, kaáne, and faáo respectively, stressing the second vowel of the double vowel, and thus prohibiting coalescence into a single long vowel. Thus, because we do find a long vowel here, we cannot extend Taumoefolau’s analysis to these languages.

It is proposed that in these languages, there has been a process of phonemicization, i.e. the formation of new abstract mental sound units. Unlike Niuean, Tongan, and Rennellese-Bellonese, double vowels in Hawaiian, Tahitian, and Tokelauan have merged into a long vowel in all positions. The phonemicization of two short vowels into a single long vowel is speculated to have occurred due to the fact that its variant (V_iV_i) did not surface in any pronounced form. Thus, because there was no learnable variation, the previous allophonic variant became the exclusive phonemic form. This is evidenced further by the fact that there are no double articulated vowels found in these languages\textsuperscript{19}. This process of phonemicization is illustrated below, using the Tahitian word vá:hi ‘place’ and the Tokelauan word fá:o ‘seize, snatch’.

\textsuperscript{18} However, they state, interestingly, that the relative frequency of CV:CV sequences is lower than CVCV, CVCV:, and CV:CV: sequences.

\textsuperscript{19} The exception to this is adjacent vowels across morphological boundaries which are outside of the scope of this paper. An example from Pukui & Elbert’s (1986) Hawaiian Dictionary is ka’aaahi /kaʔa=ahi/ ‘train’, lit. ‘wagon-fire’.
THE PHONETIC NATURE OF NIUEAN VOWEL LENGTH

(15) a. /vaahi/

footing (va.a.hi)
fusion (va:hi)
stress (va:.hi)  phonemicization

[va:.hi]  Æ  /va:hi/

footing (va:hi)
stress (va:hi)  

b. /faao/

footing (fa.a.o)
fusion (fa:o)
stress (fa:.o)  phonemicization

[fa:.o]  Æ  /fa:o/

footing (fa:.o)
stress (fa:.o)

Thus, I propose that before stress is assigned, all double vowels merged into a long vowel. What triggered this merger pre-stress, I argue, is a reordering of constraints, constraints found in (T1–4) above. Niuean and Tongan crucially rank RH-CONTOUR and ALL-FT-RIGHT above DIST in order to eliminate candidates which violate RH-CONTOUR and ALL-FT-RIGHT, and allow certain candidates to incur a violation of DIST. However, because we never find double articulated vowels in Tahitian, Hawaiian, and Tokelauan, we must propose that DIST moved higher up in the constraint ranking, eliminating those candidates having qualitatively non-distinct segments adjacent to each other (and thus violating DIST). This diachronic constraint reranking is illustrated below with /kaane/,
the historical form of the Hawaiian word ka:ne ‘man’.

<table>
<thead>
<tr>
<th>T5. /kaane/ (Ha: ‘man’)</th>
<th>GrWd =PrWd</th>
<th>RhType =T</th>
<th>WSP</th>
<th>Dep-IO</th>
<th>Max-μ-IO</th>
<th>Ft-Bin</th>
<th>Right-most</th>
<th>All-FT-RIGHT</th>
<th>Dist</th>
<th>RH-CONTOUR</th>
<th>Mora =Syll</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ka.(a.ne)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (ká.ne)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (ká:.ne)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ka.(ta.ne)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. (ká:.ne)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. ka.a.ne</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. (ká:.nê)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. (ka:.nê)</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here, the term /kaane/ surfaced as [ká:.ne] due to DIST ranked above RH-CONTOUR, eliminating the form [kaáne]. Though the form [ká:ne] ends in two weak morae, and violates RH-CONTOUR, it incurred its violation after all other possible candidates are eliminated. This constraint reranking thus explains why we find CVi:CV forms in these languages, though do not find them in Tongan, Niuean, and Rennellese-Bellonese. Further, due to the process of phonemicization of long vowels described above, this form [ka:ne] has underlyingly become /ka:ne/, with stress attracted to the long vowel, rather than creating it. This phonemicization, of course, has further repercussions on the stress system

---

20 This form /kaane/, under the proposed phonemicization scenario, would synchronically be /ka:ne/. I choose the older form in order to illustrate the constraint reranking which must have occurred.
of these languages, an issue I do not explore at present. In these languages, a long vowel is stressed, otherwise the penultimate element is stressed (Bickmore 1995; Elbert & Pukui 1979; Office of Tokelau Affairs 1986). Stress is always attracted to a long vowel due to the highly ranked WSP constraint, stating that heavy syllables must be stressed. A more thorough analysis of stress in these languages is left to future investigation; in general, vowel length has become phonemic in Hawaiian, Tahitian, and Tokelauan. This vowel length variation between Polynesian languages accounts for the great variety of analyses put forward by different authors.

The precise point in the proto-Nuclear Polynesian timeline in which phonemicization occurred is hard to determine due to the fact that in many languages, there is free variation within the same form between a double articulated vowel and a long vowel. This issue is investigated in the section below.

6.4 Variation

It is difficult to conclude where the exact point of phonemicization of vowel length occurred, for example, whether it was in proto-Eastern Polynesian, proto-Nuclear Polynesian, or in the proto-Samoic branch. This is due to documented free variation between long vowels and double articulated vowels in all three branches of the family, attested in Maori [Eastern], Samoan [Samoic], and even Tongan [Tongic].

In Maori, sequences of double vowels may be realized for the same word as a double articulated vowel, a long vowel, or two vowels separated by a glottal stop. This is documented by Bauer (1993: 534), who notes with respect to this free variation that forms containing long vowels have “a variety of pronunciations” which depend on factors such as “formality, speed, and emphasis”. All of the following realizations may be heard:

(16) a. whakaara \(\Rightarrow [\text{f} \text{k} \text{e} \text{r} \text{a} \text{r}] \sim [\text{f} \text{k} \text{e} \text{r} \text{e} \text{r}] \sim [\text{f} \text{k} \text{a} \text{r}]\) ‘arouse, motivate’

b. iriiri \(\Rightarrow [\text{i} \text{r} \text{i} \text{r} \text{i}] \sim [\text{i} \text{r} \text{i} \text{r}] \sim [\text{i} \text{r} \text{i} \text{i}]\) ‘baptise, christen’

A realization as a long vowel is noted as “the norm in casual fast speech”, indicating that realizations where both vowels in the sequence are pronounced separately are used in careful and formal speech. Note above that in Maori, long vowels and diphthongs are stressed; in the absence of these, the first vowel which is not more than 4 vowels from the end of the word is stressed (Williams 1996). Thus, we see that in Maori, stress does not in fact create long vowels, but rather long vowels attract stress; this further complicates extending a phonetic analysis of vowel length to this language.

Furthermore, in Samoan, Hovdhaugen (1990:97, fn 6) states that long vowels of certain roots are pronounced as a sequence of two identical short vowels when they are spoken in a very formal style. Mosel & Hovdhaugen (1992) provide examples to illustrate this, repeated in (17) below. [Note that stress in Samoan falls in the following order of preference: a long vowel, a diphthong, the penultimate vowel.]

\[\text{Note that stress in Samoan falls in the following order of preference: a long vowel, a diphthong, the penultimate vowel.}\]

---

21 Interestingly, Hovdhaugen also notes in the same footnote that derived words whose stem contains a long vowel do not show this variability in formal speech, providing the following example:

\[\text{tōga } [\text{tō} \text{ŋ} \text{a}] \text{ ‘(of houses) erection, putting up’ } \xleftarrow{\text{tō ‘(of a house) to put up, erect’ + -ga NOMINALIZER}} *\text{[to.o.ŋa]}\]
(17) a. /taane/ → [taáne] ~ [tá:ne] ‘husband’
b. /tooŋa/ → [toóŋa] ~ [tó:ŋa] ‘fine mat’
c. /feoloolo/ → [feoloólo] ~ [feoló:lo]

A minority of adult speakers (consisting mostly of men) always pronounce the above forms as double articulated vowels at any level of formality, as well (Mosel & Hovdhaugen 1992:30). Thus there are instances of free variation in both the Samoic branch and the Eastern branch of the Polynesian language family.

Further, even Taumoefolau claims that her analysis applies only to informal speech, and not a formal, conservative register. She writes that “an example of Churchward’s ‘long vowel’ is written as má ‘bread’” which in careful speech is pronounced as [má-a] (Taumoefolau 2002:346). This indicates that two identical adjacent vowels are not merged into a long vowel when pronounced in careful speech. These three documented cases of variation, dependent upon sociolinguistic factors, complicate the issue significantly of constraint ranking and reranking. If the same underlying form may have two surface forms which are determined by sociolinguistic factors (such as formality), then how can one posit a constraint ranking, which is able to account for this variation to the exclusion of other variant forms? In other words, how can we develop an analysis which allows a form to surface as two different manifestations depending on context? This issue is not addressed at present, though does touch upon a serious theoretical concern for phonological analysis of natural language systems.

Further, with regard to our current study, the issue of the exact location of constraint reranking and subsequent phonemicization of vowel length cannot at the moment be pinpointed. Due to the presence of the free variation, we are only able to determine for the particular language whether or not there is phonetic or phonemic length contrast, or whether double vowels and long vowels vary with respect to the factors mentioned above. The process of phonemicization can perhaps be thought of as a change in progress in languages such as Maori and Samoan, whereas this process has been completed in languages such as Hawaiian, Tahitian, and Tokelauan. This investigation, however, rests entirely upon secondary research, and without primary research into formality variation in Polynesian languages, the historical development of vowel length across Polynesian languages remains incomplete.

7. Conclusion and Further Issues

This paper has sought to provide an overview of vowel length across Polynesian languages. I have summarized previous analyses of vowel length in Tongan, and have argued that an analysis proposed by Taumoefolau (2002) for Tongan can be extended to Niuean. Niuean long vowels therefore are two identical adjacent single vowels underlyingly which surface as a long vowel if stress is placed on the first element, and surface as a double articulated vowel if stress is placed on

---

22 Anttila and Cho (1998) attempt to incorporate variable phonological features within an OT framework. They borrow heavily from formal mathematical set theory, and put forward an OT grammar model in which a language may be partially ordered, i.e. a certain constraint may not necessarily be ranked with respect to a neighboring constraint. In this case, the constraints DIST and RH-CONTOUR are not “connected” (i.e. ranked with respect to one another), allowing for the variation that is seen. This paper, and the ideas within, were brought to my attention by Derek Denis (pers. comm.), and may act as a starting point to untangling any variation issues.
the second element. Thus, vowel length in Niuean and Tongan is a phonetic, not phonological, phenomenon. I have provided evidence for extending this analysis from the lack of certain CV templates in Niuean, namely *[CV:CV] and *[CVCV:V], and from restricted adaptation of English loan words. I accounted for this complementary distribution between double articulated vowels and long vowels within an OT framework, providing a set of constraints and a series of tableaux which crucially order the constraints RH-CONTOUR and ALL-FT-RIGHT above the constraint DIST, which states that all adjacent segments must be distinct from each other. I tentatively extend this analysis to Rennellese-Bellonese [Samoic], though it cannot extend to the Polynesian languages Tahitian [Eastern], Hawaiian [Eastern], and Tokelauan [Samoic]. I propose that these languages do contain contrastive vowel length, which resulted from a constraint reranking (DIST > RH-CONTOUR), with a subsequent phonemicization of vowel length. The exact location of this phonemicization within the historical development of the Polynesian family cannot be determined at present due to documented free variation between double vowels and long vowels. This free variation is found in Maori, Samoan, and Tongan, representing all three branches of the family. Several further issues were left unresolved. Which forms can be in free variation within an individual language? Those with stress on the first element, those with stress only on the second, or both? In which languages is this variation allowed over morphosyntactic borders? Also, in languages which have fully phonemicized vowel length (e.g. Hawaiian), is there possibly undocumented length variation in very formal situations? What is the status of vowel length, and its interaction with stress and footing, in the dozens of other Polynesian languages? These issues require additional secondary and primary investigation in order to draw conclusions. It is the hope that this paper will provide a starting point for describing vowel length for the individual Polynesian language, and the Polynesian family as a whole, while making theoretical contributions to the study of contrastive and non-contrastive sound units.

References