

The Context of Polynesian Reduplication

J. C. Finney

This paper seeks universal and AN-wide principles of diachronic change in the process and outcome of reduplication in natural languages as spoken. Beginning with analysis of reduplication in Polynesian languages, it proposes a model for PPN in which partial reduplication is used for plurality agreement on verbs and full reduplication used for aspect. It puts it in context with basic principles; with other language groups; and with PN orthography.

The term Reduplication (hereinafter Rdpc) is well established, though 'duplication' is more accurate. Most Rdpc in AN and Indo-European puts the Copy left of the Base. Proto-PN Rdpc had some obtuse features, but in general it seems to have had a fairly transparent and logically consistent system as shown in Figure 1. Every syllable was a CV mora. Most morphemes and many words were CVCV, composed of two syllables/moras. Much the same was true of Japanese at the time when the 50-syllable kana spelling came into being. There, too, every syllable was a CV mora; and many major morphemes and words were CVCV, composed of two syllables/moras. In recent centuries both Japanese and PN languages (or dialects) have lost some transparency (including transparency of Rdpc), arising in part from phonological change and affecting morphology and syntax. The similar courses took place despite the fundamental difference that Japanese had OV while PN (like most AN languages) had VO order. Further studies comparing PN and Japanese could be enlightening.

Bimoraic syllables containing long vowels and diphthongs are made in PN by word-formation, combining morphemes into longer words, and from other diachronic changes. The creation of these long syllables (in PN and elsewhere) challenges the simple Rdpc system. Where the system had been reduplicating one or two syllables which were also moras, speakers must now choose between reduplicating one or two moras, or reduplicating one or two bimoraic syllables. In defining what unit serves as the Base for Rdpc, scholars have varied in choosing word, morpheme, foot, syllable or mora; and have had to speak of taking the left syllable (or mora, or CVC sequence) of the Base as the part that is used ("template") for the Copy. We shall note examples in Samoan. All this is conceptualized most easily for languages such as PPN and 50-syllable Japanese in which every mora/syllable was CV and every template CV or CVCV.

In Polynesian languages today, double consonants arise only from lenition after Rdpc; while double vowels arise both from Rdpc and from other situations. I'll show an example of incipient change in syllable stress rules that I observed in which

practice varied not only among individual speakers but even among different occasions for the same speaker.

Both the unit question and the stress question involve differences in optimalities. So, though Optimality Theory developed in synchronic generative grammar, and for phonology, we have occasion to mention it in the diachronic account. Criteria are shown for distinguishing the ancestral from the innovative choice.

Linguistic issues create political issues on orthography. Once a spelling system has been established, however unsatisfactory (English as well as Hawaiian, Samoan and Kapingamaringi), powerful resistances operate against changing it.

An incident is told in which a bright eleven-year-old Polynesian child made a linguistically more sophisticated judgment than was made by the teachers.

1. Definitions

Ext (also called Fqv for 'frequentative') is an 'extensive', a form for prolonged, continuing or repeated action. It can also be called Exp for 'expansive'. The pre-existing form is called the Base. The part that is added in Rdpc is called the Copy. In the canonical case in which the Base is $C_1V_1C_2V_2$, the Rdpc is called full if the Copy is also CVCV. Full Rdpc can be somewhat disguised if the Base or the Copy is altered by subsequent phonological change (examples in Japanese and elsewhere). Rdpc may also be called full if the Base was CVCVC or perhaps CVCCVC (as many PAN words were) and the final consonant is not used in the Rdpc process. Partial Rdpc is defined as copying only the left syllable or mora of the Base, making the Copy CV from a canonical Base. The term "canonical" means the usual, normal, customary, unmarked form. Intermediate Rdpc, falling between Full and Partial Rdpc, is the less-common production of $C_1V_1C_2$ as the Copy. In some cases it may turn out that Intermediate Rdpc may have come about from elision after Full Rdpc, but Partial Rdpc is so common that it must normally occur in a single step. "Leftward" or "right-to-left" Rdpc is the case in which the Copy is left of the Base. It prevails in Austronesian and Indo-European. "Rightward" or "left to right" Rdpc, is the case in which the Copy is right of the Base. It prevails areally at least from India through Turkey. With Partial Rdpc the direction is clear, as the shorter form must be the Copy. Fully faithful Full Rdpc gives no clue: Leftward and Rightward Rdpc would look the same. If Partial or Intermediate Rdpc occurs in the same language, the Full Rdpc can be taken to have occurred in the same direction as the Partial or Intermediate Rdpc. Or we may be able to tell the direction from the reconstructed ancestor or from cognates in close-kin tongues. Or, knowing that in older Japanese all stops were voiceless, we might say that in "tokidoki" the right half is the one that has changed and must be the Copy. But that would be invalid, because at the time the Rdpc took place the Copy was faithfully "tokitoki", and the positional change to voiced stops came at a later time.

Let's clarify the present approach by applying it to findings made in another approach, that of Burquest & Steven (1994), which dealt with Roma, a Central MP language. They began by saying: "In his foundational paper, Marantz (1982) proposes an insightful framework to account for patterns of reduplication. Full reduplication, in which the entire morpheme is copied, is well known and rather easily treated; Marantz's major contribution is providing a means to account for patterns of partial reduplication through the use of a CV template, a principle of stem-copying, and the general types of association conventions familiar in autosegmental phonology... We first demonstrate that Marantz's proposal fails to account for patterns of reduplication in some languages of Maluku; we then consider the effectiveness of a more recent proposal (McCarthy & Prince 1990) to provide a more accurate account." They define a simple syllable as (V)V(C) and a complex one as (C)CCV(C) [neither of these is simple to me] and say that in Roma (with defined exceptions) a complex syllable occurs only initially. As an example of the inadequacy of Marantz's approach they cite the following example:

- | | | | |
|----|-------------------|------------------|---------------|
| 1. | | 3sg non-durative | 3sg durative |
| | krumat "to slice" | na-krumat | na-krum-rumat |

They assert that under Marantz's approach the wrong form will result if the stem itself is what is copied over the template, no matter whether the association is left to right (expected in prefixes) as in (a) below, or right to left (the marked case for prefixes) as in (b) below. [I agree with them and Marantz is treating "kr" as CVC.]

- | | | | | |
|----|----|--------|--------|------------|
| 2. | a. | krumat | krumat | *krkrumat |
| | | CVC + | CCVCVC | |
| | b. | krumat | krumat | *matkrumat |
| | | CVC + | CCVCVC | |

Instead, they propose that if the prefix is attached and the string syllabised, the following reduplicative process taking account of syllable structure provides the organization of segments that yields the attested result:

- | | | | | |
|----|---------|------------|------------|----------------|
| 3. | kru.mat | na kru.mat | nak.ru.mat | nak.rum.ru.mat |
|----|---------|------------|------------|----------------|

They note that their approach requires that both affixation of subject prefixes and syllabification precede reduplication (and that the "template" is CVC). They also note that the same correct result follows if syllabification is taken as persistent.

As their intention is to give a synchronic generative path and mine is to give a diachronic natural language path, my derivation is not in competition with theirs. In any AN language (except where PAN medial CC is shown to have persisted) I'll take CCV as having suppressed a vowel, and as having been CVCV at the time the RdpC took place. So for CMP Roma I'll take both "na" and "k" as prefixes, i.e. as

separate morphemes that are not involved in the reduplication. (My approach also accounts for their Arabic examples on their page 94, even though in that case I have to deal with examples cited from Burquest & Steven, who cited them from McCarthy & Prince, who cited them from Marantz, who cited them from another source.) When we take "rumat" as the Base, the simple application of Intermediate (CVC) reduplication gives the correct result immediately. As I don't know the language, I can't choose between the two possible historical courses: either "kV" (with a vowel that has since elided) was once a separate morpheme, or reanalysis at some point in time made separate morphemes of what was always ancestrally a "kVrumat" morpheme of three syllables. Clearly we need more research to settle the question of when it works to define the Base as a word or a morpheme or a phonological unit such as a foot or set of two syllables. Problems arise when what was ancestrally two syllables has become one. The best answer for synchronic analysis will not necessarily be the best answer for diachronic analysis. I can't deal with Burquest & Steven's fascinating Arabic example (page 92) in which a syllable in the output gets its vowel quality from one part of the input and its vowel length from another part. It's three citations deep to them and four to me, and I don't know Arabic.

Another use of the my diachronic approach is to reanalyse the problem in Thao as presented by Chang (1997). What she calls Rightward Rdpc is Leftward in our definition, as the Copy is put left of the Base. Since she takes stretches of more than two moras as Bases, she finds it odd that for many words only the right two moras are reduplicated. She gives it a special name, "Reduplication from the Right Edge". My diachronic approach requires that the relevant morphemes were limited to two moras, and so I don't find it anomalous or exceptional in any way, nor as requiring any special name. It's the normal, canonical, unmarked way to do it.

2. A Basic Principle of Human and Animal Behavior: FAGR

Long ago, experimental psychologists, running hungry rats through mazes seeking food, and doing other experiments with other animals, including human beings, noticed that certain behaviors occurred on the achievement of the reinforcement or reward, and those behaviors were called Goal Responses. In time, the Goal Responses came to make their appearance shortly before the achievement, presumably when the subject came to expect that the reward was about to come. Such anticipatory goal responses occurred not in full but in part, to some degree between zero and 100 per cent. So they were called Fractional Anticipatory Goal Responses (FAGR).

Just completing an utterance is mildly rewarding; in part that's because when we express thoughts we can influence other people. So let me present the case that at least two linguistic processes come about through FAGR: umlaut and Rdpc.

Consider the i-umlaut of the vowel "u", in old Germanic languages, English as well as German. Suppose that in the root syllable you have a vowel "u", and in the

following syllable, some syntactic cases in the plural have a vowel "i". By FAGR the individuals begin to move the tongue into "i" position while they are saying the "u" syllable. The result may be that in such words the "u" takes some of the characteristics of the "i". The lips are rounded as for a "u" but the tongue has moved into the "i" position. Thus the umlaut "u" phonology. If the umlaut conditions are found for more syntactic cases in the plural than in the singular, people take advantage of the difference to generalize and make the umlaut vowel the main or only marker of the plural. English has relics of the process in words such as "mouse/mice" and "foot/feet". One of my children, hearing us mention termites, asked, "What does a termouse look like?"

For Rdpc the FAGR process works a bit differently. Some fraction of the Base comes to be expressed as a Copy, spoken earlier than the Base.

An important point is that FAGR puts the Copy left of the Base because it consists of anticipation. We remember the past and regard it as determined and unchangeable, but we expect or anticipate the future and regard it as incompletely determined and subject to our changing it.

All this is in accord with a basic principle of physics, the Second Law of Thermodynamics. Nevertheless, it is not controlling. There is at least one broad area of the world in which most Rdpc puts the Copy to the right of the Base. It includes Dravidian, and the modern descendants of Sanskrit, as well as most of the Semitic and Turkic languages. Evidently an areal effect can defeat the Anticipation principle. And in Vietnamese the Copy is apparently sometimes left and sometimes right of the Base, in circumstances that have not been stated.

3. Comments on Gathering Evidence

Polynesian (henceforth PN) is a good choice for comparative and diachronic studies within AN. Despite its phonological simplification and its wavering between accusative and ergative syntax, it is a conservative language group with much of its lexicon and syntax very recognizably cognate with PAN. It keeps the PAN format of bisyllabic morphemes and words. (That's in contrast to nearly all Micronesian languages, for example, which, by lenition of most major words to monosyllables and by grammaticalization, have lost ready comparability with PAN and Formosan languages.) Nanumanga and its neighbor Niutao are among the most conservative tongues within PN. I used Noricks's (1981) Tuvalu dictionary, supplemented by Besnier's (1981) and by my own field notes. A paradigm of full and partial reduplication was discovered. Dictionaries and grammars of other PN languages including Samoan and Tongan tend to confirm the paradigm. Findings by others in other major branches of AN were studied, and reasonable hypotheses about PAN Rdp are possible.

Even with the best efforts, words in the language will be missed. In Nanumanga in 1971 I had determined to get a full list of the usual major morphemes, which were canonically CVCV. The possible vowels were five (a, e, i,

o, u). The consonant slots could be filled with any of twelve elements, which were zero and eleven consonants (p, t, k, f, s, m, n, ng, v, l). So I prepared a list of the 3600 possible forms (12 x 5 x 12 x 5) and, over several months, presented them all to my chief informant. Even so, I didn't get the whole list, as I found later that he had given zero answers for tabu words, and had failed to list many homonyms.

4. The Austronesian Ca- Rdpc

The most interesting recent finding is what Blust (1997) described as Ca-reduplication and ascribed to PAN. In form it is a Partial Rdpc in which the vowel *a* is substituted for any other vowel. It had three uses: (1) Numerals. For the numbers one through ten, and the interrogative "how many?" it forms an alternate set of numbers, such as *lalima*, a doublet for *lima*, 'five'; (2) Verbs, to express future or contemplated aspect in Puyuma and in Mayrinax Atayal, and in other languages (two in Eastern Indonesia) some other functions including durative and reiterative aspect and plurality of subject; and (3) Formation of instrument nouns. As we see, the list includes the uses common in Polynesian, both Extensive aspects and agreement of verbs with plural subjects, and shows them to have been expressible by Ca- Rdpc in PAN; and it adds other functions also attributable to Ca- Rdpc in PAN. The discovery raises interesting questions about the relation of Ca- rdpc to full and partial Rdpc in PAN. It may even re-open my willingness to consider the Samoan *ta*-reduplications as possible survivors of PAN Ca- Rdpc through a convoluted history of reanalyses. Blust notes the use of vowel-lengthening Partial Rdpc to express contemplated (future) aspect in Tagalog, and surmises that it may have arisen from or substituted for Ca- Rdpc.

5. The Polynesian Paradigm

I began by examining the Rdpc of Tuvalu because I had spent the greater part of a year, 1970-1971, doing fieldwork in linguistics and cultural anthropology there. In the present study I began by examining Tuvalu dictionaries by Noricks (1981) and by Besnier (1981) and then broadened the corpus to a Samoan dictionary (Milner) and grammar (Mosel & Hovdhaugen) and Churchward's grammar (1953) and dictionary (1959) of Tongan (Churchward, 1959), among other sources including my own notes. In Polynesian, Rdpc occurs essentially in verbs (including adjectives, stative verbs), not much in nouns. Most verbs and almost all nouns are invariant for number. The plurals of verbs "go" and "come" commonly have a different lexical item (unrelated phonologically to the singular). In general, in PN, partial Rdpc makes a plural, and full Rdpc makes an Exp. What came closest to a recurrent paradigm is shown in Table 1. The data for the lexical item used in that table are in Noricks (1981:136).

Tongan also pluralizes some verbs (and nouns) by Partial Rdpc with or without left syllable lengthening, and thus shows that those processes had taken place in Proto-Polynesian.

The concept advanced here is that the forms are created in the order shown in Table 1, from left to right and from top to bottom. Arrows show the proposed directions of historical change in the forms as actually spoken. The sequence of derivation in a synchronic generative grammar need not be the same. The reduplicating verbs in a given PN language may show part or all of this paradigm. Some words such as *tipi* ("cut") show only the Exp (full Rdpc) in Tuvalu, but only the partial reduplication in Samoan. Others such as "tele" ("run") show only the pluralizer (Partial Rdpc) in both Tuvalu and Samoan. All this is compatible with a full regular paradigm in a common ancestral language.

This account requires that when a word consists of more than one morpheme, each morpheme remain a morphological and syntactic unit that can enter into processes. In the case of the singular Fqv form *piupiu*, its right morpheme *piu* functions as an entity that subsequently undergoes partial reduplication to form its plural, $\{ \{piu\} \{pi-piu\} \}$. This account asserts that the formation of the plural Ext form arises from the singular Ext and not from the plural basic verb. It makes sense to do word-formation first and then inflection. In this account the second step is to take the Base and copy its left syllable to the position immediately left of the Base (ignoring that there is already a full Copy left of that point). No "infixing" theory is needed. The alternative explanation (inflection to plural and then word formation to Ext aspect) would require step 2 to take the Base and copy it in full to the left of the compound made by the pluralizing partial Copy plus the Base. That doesn't seem to be the natural order either diachronically or synchronically.

It is true, however, that the forms and their usage vary from one lexical word to another, even within a given PN language. In some words the form shown in the table for Ext (full Rdpc) is given as a plural and not as Ext. There is a tendency to follow the table. Noricks's shows 'pake' with its frequentative as pakeke, and the latter has its plural as pakekeke, exactly by the table; but for the simple word itself no plural is listed, and so, presumably its plural is the same as the simple singular, as is true for great numbers of verbs in Tuvalu and Samoan. In some words the form with partial reduplication, when usable as a singular, displaces the simple singular form and takes the form with full reduplication as its Ext.

Other questions arise, too. One possibility is that the Polynesian Paradigm of Table 1 held true and applied universally in PPN. If so, we'd have to investigate the hypothesis that it came down from Proto-Oceanic, from Proto-Eastern Malayo-Polynesian, from Proto MP, and possibly from PAN. On the other hand, to the extent that it holds true for a group in Polynesian, or for Polynesian as a whole, it is possible that the paradigm developed its applications by progressive generalization, analogously to the way the plural meaning of umlaut developed from generalizations in West Germanic that in the beginning did not signify plurality at all. Was the correspondence between phonological form and syntactic/semantic meaning one that has grown or has weakened over time?

A point needing mention is that the Base morpheme can occur in compound words, having another morpheme at its left or its right or both. The normal Rdpc is to duplicate the Base where it is and leave the other morphemes unchanged where they lie at the far left or far right. In occasional exceptions, however, the other morpheme (or one of the other two morphemes) is treated as the meaningful morpheme and is the one subjected to Rdpc. In one odd example in Samoan, listed by Mosel & Hovdhaugen (op. cit., 218-219) "reduplication is applied across morpheme boundaries" as follows:

4. fe o a'i
 reciprocal prefix verb "go about" SUFFIX
- > feofea'I "use to go about"

"Fe o " is treated as if it were a morpheme. Perhaps that's a sign that it is becoming a morpheme. Despite its three moras, the fact that it is only two syllables helps to license its use as a "template".

A final comment on Polynesian is that the Copy shows vowel changes that seem anomalous; the partial Copy in some cases lengthens the vowel, and in other cases substitutes the vowel a in what looks like Ca- Rdpc. Both occurrences create a resemblance to what is found in other branches of AN.

Finally, I read the section on Rdpc in Samoan Reference Grammar (Mosel & Hovdhaugen, 1992, 218-239). It is delightful that they had found nearly all the same facts that I did, drew many of the same conclusions, and they expressed their findings outstandingly well.

6. Irregular Rdpc in Polynesian with apparent Ca- Rdpc or with change in vowel length

Some Rdpc in Samoan look like Ca- Rdpc. Some examples (Milner 1966), all plural, are:

- | | | | | |
|----|-------|------------|--------|-------------------------------|
| 5. | tele | plane (vb) | plural | tatele |
| | tia'i | get rid of | plural | tatia'i |
| | tipi | cut, slice | plural | tatipi. |
| | Tui | stab | plural | tatui (also tutui and fe-tui) |

These all look like the standard plural formation (Table 1) but for the change in vowel in the Copy. None has any of the functions found elsewhere in AN by Blust (1997) for Ca- Rdpc.

Churchward's Tongan dictionary shows "tele" as an intransitive verb "just to touch or graze in going past; "tele" as a transitive verb "to peel or pare off with a knife; to plane or shave (wood or whiskers); and "tatele" intransitive verb "to skim

along the surface"; and Churchward marks "tatele" as derived from "tele". It is only with knowledge of the Samoan and Tongan forms and meanings that one can detect the (ancestral) relation of *kele* and *kakele* in Hawaiian (Pukui & Elbert 1971).

It turns out that Samoan has taken two extraneous prefixes and come to use them to show plurality of subjects. One is 'fe', which is more reciprocal and less pluralizing in the neighboring languages. The other is 'ta', whose use looks like Ca- reduplication when the Base verb is *te-*, *ti-*, *to-*, or *tu-*. In Hawaiian, if I may substitute the older 't' for the 'k' spelling, *tamoe* 'recline' and *tamotu* 'cut' are defined much like the same verbs without the *ta* prefix. The conclusion is that probably none of the forms in Polynesian turn out to be Ca- Rdpc, despite their appearance.

As for vowel lengthening, it forms the plural in 24 words listed in Mosel & Hovdhaugen; who note that many of them are very formal and obsolescent. Only three of the 24 are disyllabic:

- | | | | | |
|----|------|-----------|--------|------|
| 6. | ala | awake | plural | ala |
| | sali | scoop out | plural | sali |
| | su'e | look for | plural | su'e |

Besides, vowels in the left syllable are often lengthened, especially in nonpluralising full Rdpc.

7. Polynesian "Ext" as more extended in time and less intense, less abrupt

By our definition, Extended aspect or Ext is a reduplicated form that does not pluralize but changes the amount of the act or process, usually increasing it but sometimes even decreasing it. It may change an act (punctual, perfective) to a process extended in time, and often gentler. Often the Ext denotes a process that extends over longer time but is less abrupt or less intense. Examples in Tuvalu, mostly from Noricks, are:

7. fia "wish" > fiafia "be happy"
fia-fai (literally wish to do, fia = wish, like, fai = do, make) plural is fiafafai, Fqv of fai.
fili "twist, roll" > fifili "braid, weave"
fili "pick, select" > filifili "sort through, pick and choose, discuss, decide"
fola (PPN, Samoan) "spread" > folafola "announce, describe publicly at length"
fofola "stretch out (physically)" > folafola "announce at length"
fua "measure, weigh" > fuafua "consider, examine, judge"
kumi squeeze, twist, kukumi grasp, grab > kumikumi squeeze, massage

mili twist, screw > milimili clean ear with object, rub back & forth between palms of hands
 palu cut (fish), mix, stir; > palupalu mix repeatedly, stir continuously, cut into many small pieces
 pona bump > ponapona be bumpy (as a surface)
 pule rule, be in charge, be the boss > pulepule supervise, manage, sit in discussion
 tala tell > talatala (also: gossip)
 teka go, move along, roll > tekateka amble, move along slowly

8. Samoan reduplicative subject choice marking syntactic change in ergativity

In ergative languages, including Samoan, some properties of the subject belong to the Ergative case, the Agent (a former prepositional phrase), and other properties of the subject belong to the Absolutive case (a former Nominative, used for the Subject of intransitives and the patient of transitives). We have seen that some verbs in Samoan use Partial Rdpc to show agreement in number with the subject. But which is the subject for transitive verbs? The Ergative agent or the Absolutive patient? I plan to discuss the matter in another paper. The short answer is that a syntactic change is in progress. The older generation still makes the verb agree in number with the Absolutive patient. The younger people make the verb agree in number with the Ergative agent, thereby uniting A(gent) and S(ubject) in a syntactic category. Thus Samoan is taking a step that is part of the changing from being both syntactically ergative and morphologically ergative to being only morphologically ergative. Samoan is not necessarily crossing a boundary, as other criteria such as raising and equi deletion generally change at other times (making syntactic ergativity a matter of degree). Manning (1996:2-11) defines those terms more technically in generative grammar. Most "ergative" languages are only morphologically ergative. What may happen (and did in Proto-Central-Eastern Malayo-Polynesian) is that the old Ergative form comes to be the new Nominative and the syntax comes full circle back to Nom-Acc. That may not happen in Samoan, because a two-clause structure is coming into use in place of the old Ergative transitive.

9. Orthographic issues in Polynesian arising from Rdpc: vowels

Tuvalu ("valu" means eight), the former Ellice Islands, consists of eight major islands with a total population less than ten thousand, in the Samoic-Outlier group of Polynesia, near where the equator crosses the date line. A typical major word or morpheme consists of two light syllables, $C_1V_1C_2V_2$. A consonant may be zero. All the Outliers may have been settled from Tuvalu.

Though Polynesian words and morphemes are canonically CVCV, reduplication and its aftermath can produce words with double consonants and double vowels. Consequently, issues of orthography arise in some languages.

Double vowels are identical vowels in succession. In Tuvalu, double consonants arise only from reduplication, while double vowels can arise from reduplication but also from lengthening and from other circumstances in which the final vowel of one morpheme is followed by a morpheme whose initial consonant is zero.

In principle, in Proto-Oceanic, and in Tuvalu,

- I. Every syllable is a single mora consisting of a single consonant onset (which can be zero) and a single short vowel; and
- II. Stress falls on the penultimate syllable, which is also the penultimate mora; and
- III. Stress within a word is affected by feet, with two moras in a foot, and the stress on the left mora of the foot. So, counting from right to left, the even numbered moras will be stressed. If a word has an odd number of syllables, the leftmost syllable is unstressed and is not in a foot. Call this the RLMR (right to left measuring rule). But often the word consists of several morphemes of one or two moras each, with a tendency to keep the stress on the penultimate mora of each morpheme, as if the morpheme were a word. So there may be some conflict of rules, and an opportunity for optimality calculation to apply.

For at least some languages as spoken today, the rules may be modified for certain phonological reasons, occurrences of VV with vowels identical or not.

If the word ends with a pair of identical vowels, *aa*, *ee*, *ii*, *oo*, or *ee*, little change, if any, occurs. Singers know that when a two-beat note occurs at the beginning of a measure, it doesn't violate the rule of stressing the first beat in the measure. Likewise, in Polynesian prose. Stress falls, as expected, on the left of the pair. The pair is perceived as a single long vowel.

If the word ends in VVCV, with VV pair identical, as, for example, Rdpc *iita*, the stress should fall on the second *i*, as it is the penultimate mora of the word. But in practice, people tend to fuse the pair into a single long vowel. For singers, this is like having a single syllable occupy the last quarter note in one measure and the first quarter note in another measure, an abnormal happening. Under those circumstances, Polynesian speakers tend to treat the antepenult and the penult as a foot, a single long syllable, even though that violates the normal RLMR.

In Tuvalu the verb or adjective *ita* (be reluctant, unwilling, resentful) has plural form *iita*, made by partial Rdpc. It is said as *ita*, with a single long vowel, and the stress seems to hold steady or decrease slightly through the combined length. Likewise, the verb / adjective *kaitaua* by the normal partial Rdpc, makes its plural *kaitauua*, and the identical vowels in contact fuse to make a long vowel, which can also be spelt *kaitaua*. Certainly the way to show the reduplication and thereby show

the plurality is to spell the words with double vowels. If spelt with a single vowel, with a macron, the reduplication and plurality are harder to discern; and if spelt with a single vowel without a macron, the plural would be taken for the singular.

Now, let's examine the situation with unlike vowels in contact. In Tuvalu, *ceteris paribus*, there is a preference for stressing the lower vowel: $a > i$; $a > u$; $a > o$; $a > e$; $e > i$; $o > u$; $e > u$; $o > i$. Define a rising diphthong as one in which the left vowel is lower than the right: *ai, au, ae, ao, ei, eu, ou, oi*.

When a word ends in *aiCV, auCV, aoCV, aeCV, eiCV, ouCV, euCV, oiCV*, the rule to stress the penultimate syllable conflicts with the rule that lower vowels attract stress. Hawaiian speakers in 1960-1963 wavered in placing the stress. If the word was, for example, *taina*, some followed the penultimate stress rule and stressed the *i*. That was obviously the traditional practice. Others treated the *ai* as a diphthong and stressed the *a* part. Still others wavered from time to time. Optimality differed from one person to another. But if the higher vowel came first, as in *huaga*, no rules were in conflict, and everyone put the stress on the penultimate.

The mayor of the island was named *Faalo*. The name meant "ceiling". He told me in 1971 that on birth he had been given that name because he was born in the month in which the ceiling of the church was constructed, which must have been in 1905 or 1906. When I arrived on the island the orator at the welcoming speech informed me, "Everyone who lives on the island must be a member of a family. Because you are a high-status visitor, we have declared you the son of the mayor, Faalo."

Once while teaching a selected class of brighter-than-average 11-year-olds in preparation for the notorious "eleven-plus" examinations, I wrote on the board the name of the mayor, spelling it as I had seen it, "Faalo". One boy remarked, "Really, it should be spelled with three A's: Faaalo." I could see at once that he was right. The indigenous word was *faka-alo*, with the causative prefix. It had been supplanted on most islands by the Samoan form *fa'aalo*, with a glottal stop regularly replacing old *k*. Samoans didn't write the glottal stop, and so the Tuvalu people treated it as zero. So the boy was right: historically, and synchronically, too. The word was indeed *faa-alo*, and the name needed three A's. It was said with a long vowel followed a short vowel. The sound volume decreased briefly to mark the separation between the long vowel and the short vowel. The morpheme boundary helped to prevent fusion, in contrast to the double vowels within the phoneme, which had become a long vowel with no dip in loudness. The etic sound was a long syllable followed by two short syllables. And (without any technical vocabulary, not knowing the word "mora") the eleven-year old had recognized four moras in the name. That was his ground for objecting to the three-vowel spelling which many had come to use.

When the colony got its independence, the Tuvalu Language Board (1980), following the wishes of the teachers, voted to abolish all spelling of double consonants and long vowels, and ordered all such words to be written with single consonants and single vowels. Of course, they shortened Faaalo only to Faalo and

not to Falo. It was interesting to note that one bright eleven-year-old was linguistically more sophisticated than the teachers were. But it has long been observed that when indigenous people develop their own writing, they write phonemically, except that they don't distinguish long from short vowels. That was true of the ancient Romans. The ancient Greek alphabet distinguished long from short "o" and long from short "e", but only because the sounds differed in quality.

During the British administration, the only typed or printed writings were done in the colonial capital of Tarawa, in the Gilbert (Micronesian) Islands. Spelling of Tuvalu (Ellice) words was erratic and often followed the Gilbertese. The Tuvalu word *sili* was spelt *tili* because in Gilbertese ancestral *ti* had become phonetically *si* without any change in the spelling. The Tuvalu people varied in spelling the velar nasal as "ng" (from the English practice) and as "g" (from the Samoan practice, as they read the Bible in Samoan).

Being aware that the greeting "Taalofa" (Hawaiian "Aloha") had a long *a* in the left syllable, I made a practice of saying it with double length. Soon I noticed that people would humorously reply: "Taaaaalofa", teasing me. I got the point. My double length was too much. The double vowels were ancestrally double length, but the Tuvalu people in 1970-71 said them with only about 1.5 times the length of a single vowel. So I learned from an unintended experiment.

10. Orthographic issues in Polynesian arising from Rdpc: Consonants

Tuvalu and Samoan use the partial reduplication of some verbs for agreement with plural subjects, as we have seen. The dialects of the islands of Tuvalu and Samoan (some more than others) have tended to change in a sequence: (1) partial reduplication to $C_1V_1C_1V_1C_2V_2$; (2) vowel elision to $C_1C_1V_1C_2V_2$. The result is a new word beginning with a double consonant. An example is the old word *pele*, "deceive, speak falsely, tell a lie." In 1971 a word of this sort would typically be *pepele* in Niutao (where the phonology is conservative). In Nanumanga it was *ppelo* and the people were well aware of the double consonant. In South Tuvalu the double consonant may have partly shortened and the people were less aware that *ppelo* differed from *pele*. Samoan had long had a standardized spelling. But Tuvalu had to choose a spelling system when it achieved independence in steps in the 1970-1980 decade.

Kapingamarangi and Nukuoro are the far northwestern members of the Polynesian Outliers. The Outliers were settled long ago, almost certainly from Tuvalu, by back-migration from East to West. The two (like Samoa and Tuvalu) use partial reduplication to show plural agreement for some verbs (including some adjectives). But, like South Tuvalu, they have fully deleted the vowel in the partially reduplicated verbs. An example, in Kapingamarangi (Lieber & Dikepa) is:

8. *wele* "burned (sing.)" *wwele* "burned (pl.)" from **wewele*

For the three stop consonants, the "long" consonants [pp, tt, kk] are written *p*, *t*, *k*, and the "short" ones [p, t, k] *b*, *d*, *g*. That bears some likeness to the fairly recent change in alphabeticalization of Mandarin Chinese words, in which the simple "p" as in "Peking" is spelt "b" now ("Beijing") and the aspirated p (formerly having a diacritical mark) is spelt with a simple "p" now. That's not exactly the same distinction as that between single and double *p* in Polynesia, but it's close. In Kapingamarangi and Nukuoro all other consonants, even *h*, have the long form written by doubling the short character. An exception is the velar nasal, written as a diagraph *ng*, with the doubled form *mng*.

9. *ngala*, "to lose (sing.)" *nngala*, "to lose (pl.)"

Four systems have been proposed for spelling the double consonants. In evaluating the systems, bear in mind that PN languages have no voiced stops and have no clusters of consonants. There is no i-glide (*y*). PPN had a u-glide (*w*) but it's a bilabial "v" or even an English [v] in most PN tongues today. PN languages have a consonant between "i" and "r". All have at least two of the three nasals and one is "m". Many PN tongues have changed "f" or "s" or both to "h". Some have changed "k" to glottal; and some that have done so have later changed "t" to "k", leaving no "t". All consonants can appear initially, and thereby can be doubled by partial reduplication followed by deletion of the vowel.

The four systems are:

- I. Write only a single consonant, as prescribed now in Tuvalu.
 Drawbacks: The system spells alike words that are still said differently by most speakers and may have significantly different meanings. In my judgment, this is the worst of the four choices, though it was the choice made by the teachers. A previous administrator, born in New Zealand, had declared Vaitupu to be the standard dialect of Tuvalu. The disadvantage may or may not diminish or disappear at some time in the future.
 Advantages: Writing takes less time and takes less space on the paper. And in some areas the system may be the wave of the future.
- II. Write the consonant double.
 Advantages: In the lands where this system is considered, this system is faithful both phonetically and phonemically to the language as spoken by the majority.
 Drawbacks: Takes more time to write, and more space on the paper.
- III. Write the ancestral form CVC including the intervening vowel.
 Drawbacks: In certain parts of Polynesia those vowels have been silent for so many years that some might not understand a speaker who voiced them.
 Advantages: People find it easier to handle a writing system that includes letters that have become silent than to handle a system that has no symbols for phonemes that are still present.

- IV. Write *p/t/k* for the double consonants and *b/d/g* for the single consonants, as in Kapingamarangi and Nukuoro.

Advantages: Each of the six stops has its distinctive letter and none requires a digraph.

Drawbacks: This won't work for anything but "p", "t", and "k". Even Kapingamarangi and Nukuoro have to write dyoubled symbols for *w*, sonorants and fricatives. Also, visitors will speak "b", "d", and "g" as they are spelled.

11. Rdpc Elsewhere in Austronesian

We have noted Chang's (1997) study of Rdpc in Thao, a Formosan language with only a half-dozen living speakers. I had the good fortune to interview two of the speakers in the company of P. J.-k.Li, but did not look for Rdpc..

Tagalog (Schachter & Otanes 1972, Zack, 1994) uses *dupA* (Partial Rdpc with Left Vowel Lengthening) where Blust found PAN Ca- Rdpd for contemplated aspect (future). It uses Full Rdpc for intensifying and for moderation (Zack's "imperfectivity" is our "Ext".) Other papers at this conference give extensive synchronic analysis of Rdpc in AN languages.

12. Rdpc in Indo-European

Proto-Indo-European reduplication has been shown with certainty only in certain verbs, for the formation of perfect tense. It was a partial reduplication, CV, and Wright (1917, 146) reconstructs the vowel as PIE *e. Voyles (1992) devotes his chapter 2 to Indo-European phonological rules with emphasis on those that persisted into Germanic. Calling them "The Synchronic Rules of Indo-European", he says, "The following are both morphosyntactically and phonologically conditioned rules which have left notable traces in Germanic." Among them is Reduplication (15-16) applying to the second and third principal parts of class VII strong verbs. Voyles cites Sanskrit "stha, tasta", Latin "spondeo, sponsondi"; Greek "grapho, gegrapha" and "strategeo, estrategeka", to which Wright (1917, 146) adds "leipo, leloipa". Verbs such as "sleep", "let" and "sow" belonged to that class in Gothic. The extent to which this process may have affected strong verbs in other branches of Germanic is disputed: Wright (1917:149) and Lass (1994:164-166) show that the weak preterite in Gothic dual and plural is formed by adding the verb -ded- which had arisen as the reduplicate perfect of the verb "do". Lass adds, "There are traces of reduplication in OHG, interestingly enough in the verb 'do' itself: infinitive 'tu-on', pret 1 pl 'tat-um'." Neither of them points out that the same form survived independently as the English past tense "did". So I make that proposal here, in case nobody else has done so. Sievers & Cook (1903) consider a great portion of the Old English strong verbs to have arisen from reduplication (as opposed to ablaut) but

list only five as showing reduplication in Old English. Two of them are still in use, "read" and "let" but have lost the reduplication. The Old English present infinitives 'raedan', 'laetan', had preterites 'reord' and 'leort' with "r" for the second "l" (Gothic "lailot"). I propose that "did" is the only reduplicative preterite in English today. Sievers and Cook (*ibid*, 350-359) classify "do" with three other verbs (am/was/be, will, and go) as Proto-Indo-European "-mi" verbs. Note that English "(I) am", German "(ich) bin", and Latin "sum" are the only relics of -mi verbs in those languages. Latin has also some reduplicated presents, notably "si-sto" ('set', perfect "steti", p.p. "statum", Hale & Buck: 1903, 99). MacDonnell (1927: 86) says, "Five verbal formations take reduplication in Sanskrit: the present stem of the third conjugational class, the perfect, one kind of aorist, the desiderative, and the intensive." His examples show that in general the CV of the Base appears unchanged in the Copy, but with exceptions under complex rules. The Copy may well have been fully faithful to the left syllable of the Base at the earliest time.

English is often said to have many reduplicated words. But most of them fail to stand up to scrutiny. The go-go word and the can-can word are borrowed from French. The others that come to mind are repetitions, with change of a vowel or a consonant. The most promising candidate is "hurry-scurry". The Copy seems to be a humorous addition, not made in the normal Rdpc process, and having an unmotivated change in initial consonant. Long after its creation, the "scurry" part came to be used as a word in its own right. An even less promising candidate, "helter-skelter" doesn't even have a Base in the language. The form "walkie-talkie" (an instrument that lets a person walk while talking) takes humorous advantage of a fortuitous rime, and might almost be called a pun. Others are teeter-totter, titter-tatter, skitter-scatter, harum-scarum and shilly-shally (from "shall I?").

No Base: helter-skelter

10. Base Right, Copy Left:
 dilly-dally, harum-scarum, shilly-shally, (tittle-tattle?)
 Both Bases:
 teeter-totter, walkie-talkie, (tittle-tattle?), skitter-scatter (not in dictionary)
 Base Left, Copy Right:
 hurry-scurry

13. Rdpc in Japanese

The common and apparently productive use of RpdC in Japanese involves nouns, among which I encountered the following in an introductory course:

11	hito	person	hitobito	people in general, crowd of people
	[pito]		[pitopito]	
	toki	period of time	tokidoki	from time to time, sometimes, at times
			[tokitoki]	
	tokoro	place, location	tokorodokoro	here and there
			[tokorotokoro]	
	ware	I	wareware	we

Forms in brackets are the older forms of the word, in which Rdpc produced fully faithful copies. Subsequent phonological changes have lessened the faithfulness.

The process of Rdpc is affected by the form of the Base morpheme. Japanese and Polynesian have the same canonical phonological form of the Base morpheme, normally CVCV.

Rdpc is common and potentially productive in both Polynesian and Japanese. For Samoan Polynesian Mosel & Hovdhaugen (op. cit.: 219) say, "Full reduplication is clearly productive and can also be applied to loanwords, cf. e.g. 'afa'afa v. 'take equal shares'. cf. 'afa n./non-erg.v. 'half, halfway' (< Eng. half)."

By the Greenberg (1974) Principle, syntactic word-order, to wit, that of Object (O) and Verb (V) has powerful effects on a language's syntax. Languages with VO order put the hanger (preposition) left of the Noun, while those with OV order put the Noun left of the hanger. As Japanese is OV in contrast to the VO order of PN and most AN, can we expect Japanese to put the Copy after the Base of Rdpc? Our examples are ambiguous, as, unlike partial Rdpc, full Rdpc's Copy is indistinguishable from the Base.

14. Tentative Conclusions and Hypotheses Needing Exploration

- I. Diachronic analysis can account for most reduplication in AN languages. Find the core ancestral root within the word as a CVCV (or ancestrally CVCV) morpheme and use it as the Base. The attested form in living languages will usually be produced by regular application of Full or Partial (occasionally Intermediate) Rdpc. In test cases this worked for two languages for which reports found difficulties.
- II. The Polynesian Paradigm begins canonically with a CVCV morpheme as the Base. Canonically, the Extensive is formed by full reduplication of the bisyllabic morpheme. The form that (for some verbs) marks plural agreement with the subject is the Partial Rdpc in which the Copy is the left syllable or mora of the Base.
- III. In the AN family and in the Indo-European family, Rdpc seems to be used predominantly on verbs and the Copy is to the left.

- IV. More research is needed to settle the question, when it works to define the Base as a word or a morpheme or a phonological unit such as a foot or set of two syllables. The best answer for synchronic analysis will not necessarily be the best answer for diachronic analysis.
- VI. Both Polynesian and Japanese today are only moderately changed from earlier forms where every syllable was CV and most words were CVCV. So it would be worth while to do diachronic studies of reduplication in Polynesian as compared with Japanese.

References

- Besnier, Niko, 1981. Tuvalu-English Dictionary. Washington: U. S. Peace Corps.
- Blust, R., 1997. Ca- reduplication and Proto-Austronesian grammar. Papers for the Eighth International Conference on Austronesian Linguistics. Taipei: Academia Sinica.
- Burquest, D. A. & L. A. Steven, 1994. Prosodic structure and reduplication: data from some languages of Maluku. Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society 1992, 85-101. Tempe: Arizona State University.
- Chang, L. M., 1997. Thao reduplication. Papers for the Eighth International Conference on Austronesian Linguistics. Taipei: Academia Sinica.
- Churchward, C. M., 1953. Reduplication. Chapter 32 in Churchward, C. M., Tongan Grammar. London: Oxford University Press.
- Greenberg, J. H., 1974. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, edited by J.H. Greenberg, Cambridge: MIT Press.
- Hale, W. G. & C. D. Buck, 1903 (reprinted 1985). A Latin Grammar. Tuscaloosa: University of Alabama Press.
- Lass, R. 1994. Old English: a Historical Linguistic Companion. Cambridge, England: The University Press.
- MacDonnell, A. A., 1927. A Sanskrit Grammar for Students. London: Oxford University Press.
- Manning, Christopher D., 1996. Ergativity. Stanford, California: CSLI.
- Marantz, A. 1982. Re reduplication. *Linguistic Inquiry* 13:435-482.
- McCarthy, J. & A. Prince, 1990. Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory* 8:209-283.
- Milner, G. B., 1996. Samoan Dictionary. London: Oxford University Press.
- Mosel, U. & E. Hovdhaugen, 1992. Samoan Reference Grammar. Oslo: Scandinavian University Press.
- Naylor, P. B., 1986. On the semantics of reduplication. Focal I: Papers from the Fourth International Conference on Austronesian linguistics, edited by P. Geraghty, L. Carrington & S. A. Wurm. *Pacific Linguistics C:93.*, 175-185.

- Noricks, J. S., 1981. A Tuvalu Dictionary. New Haven: Human Relations Area Files.
- Sievers, E. & A. S. Cook, 1903 (Third Edition). Grammar of Old English. Boston: Ginn & Co.
- Stevens, A. M., 1992. Madurese reduplication revisited. Papers from the Annual Meeting of the Southeast Asian Linguistics Society, 357-381. Tempe: Arizona State University.
- Tuvalu Language Board, 1980. Standard orthography policy. Tuvalu News Sheet (Funafuti, 25 June) 109:19.
- Voyles, J. B., 1992. Early Germanic Grammar. San Diego: Academic Press.
- Wright, J., 1917. Grammar of the Gothic Language, Third Edition. Oxford: Clarendon Press.
- Zack, R., 1994. Imperfectivity as a unifying feature of reduplication in Tagalog and Hiligaynon. Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society 1992, 389-398. Tempe: Arizona State University.

Dr. Joseph C. Finney
11561 Spur Road
Monterey, CA 93940-6621
U S A
jcfinney@redshift.com