A Computationally-Assisted Procedure for Discovering Poetic Organization within Oral Tradition

David Meyer
david.meyer@yahoo.com

Abstract

A procedure is described which is capable of detecting poetic organization within transcribed oral tradition text. It consists of two components: An automated process which generates recurring n-gram patterns, and a verification process of manual review. Applied to a corpus of Tahitian and Mangarevan oral tradition, it has led to discovery of a variety of uses of meter and parallelism; ranging from the universally common to the unanticipated. The procedure should be generalizable to the study of other of the world's oral poetries, having been designed to identify a wide range of organizational possibilities.

1 Introduction

Our knowledge of the many ways by which oral tradition may be organized poetically derives from an uneven study of mostly European, Middle-Eastern, and Asian traditions. On a positive note, descriptions of the oral poetry of Indo-European languages have been sufficient to spawn the field of comparative-historical poetics (see Watkins 1995). Unfortunately, much less effort has been applied to the remainder of the world's oral traditions, which tend to fade away well before their languages die off. In a homogenizing era, unless these vulnerable data are collected and their varied means of poetic organization discovered, much of what could have been learned with regard to oral poetics universally might be forsaken.

When venturing into the study of an undescribed poetic tradition, a purely manual approach is generally insufficient. The investigative path is likely to be lined with wide cognitive gaps from researcher prejudice as to what might be recognized as poetic.

The procedure described here attempts to remedy potential bias by informing the researcher of instances of parallelism which may not have been otherwise detected. The procedure consists of two components: An automated process which generates recurring n-gram patterns,

and a verification process of manual review. Manual verification is recommended given that a tradition may employ many different organizational methods, but the corpora which contain them are often small.

Some of the examples used below are drawn from application of the procedure to two sources of Polynesian oral tradition: A 50,000 word corpus of early 19th century Tahitian material representing multiple genres, and a 10,000 word corpus of early 20th century Mangarevan songs and chants. Treatment of the complete Tahitian corpus was successful at the discovery of two varieties of counting meter (one of which may be unique to Tahiti), complex patterns of meter and sound parallelism, and many uses of syntactic and semantic parallelism (see Meyer 2011 and 2013). Analysis of the Mangarevan data is still underway.

Due to space constraints, the automated procedure's functionality has only been summarized here. It is hoped that enough information will have been provided for the computational linguist reader to be successful at his or her own implementation.

2 Description of the Procedure

As mentioned, the procedure consists of an automated process which generates recurring n-gram patterns, followed by a verification process of manual review.

With regard to former, computationally-generated candidates consist of recurring n-grams of linguistic features,[1] any of which could potentially have application to poetic composi-

---

[1] Among the oral poetries of the world, a wide range of linguistic features have been found organized in uncountable creative ways. With respect to the phoneme, for instance, poetic organization may be of the phoneme itself, of a class of phonemes (e.g. as in an assonant pattern where only the vowels are significant), of a phonemic feature (e.g. a pattern of contrasting +*acute* and -*acute*), etc. The term *linguistic feature* here refers to any linguistic information detectable at the level of phoneme, syllable, word, or line that could serve to form a pattern of poetic meter or parallelism.

tion. The n-grams are sorted and counted, and then presented – in their original context – in multiple interactive reports as preparation for manual review.

The automated component initially attempts to accommodate any linguistic feature a poet may wish to employ. After an initial round of manual analysis, however, it is desirable to pare the feature set down to just those which demonstrate some degree of promise; in order to lighten the load of the overall endeavor.[2] The list in table 1, for example, contains the reduced linguistic feature set which was ultimately selected for treatment of the Polynesian data.

Table 1. Final set of linguistic features treated with regard to the Tahitian and Mangarevan data.

Relevant to meter

Primary word stress count
Primary and secondary word stress count
Word mora count
Word syllable count

Relevant to sound and syntactic parallelism

At the level of the phoneme, one or a series of:
Phoneme
Consonant
Vowel

At the level of the syllable, one or a series of:
Syllable form
Syllable-initial phoneme
Syllable onset
Syllable rhyme

At the level of the word, one or a series of:
Word form
Word consonants
Word vowels
Word lemma
Word part-of-speech
Word-initial syllable
Word-final syllable
Word-initial syllable onset
Word-final syllable onset
Word-initial syllable rhyme
Word-final syllable rhyme
Word-initial phoneme
Word-final phoneme

It may also be necessary to re-apply the procedure were it discovered during manual review

that the oral tradition specialist's poetic use of linguistic features differs from that of general language.[3]

In its implementation, the automated process need not be restricted to observation of a single feature in isolation (single-feature pattern detection), but should attempt to be sufficiently expansive so as to detect an oral poet's efforts to coordinate more than one feature (multi-feature pattern detection). It should also be capable of detecting patterns of inverted parallelism. Line,[4] word, and syllable boundaries may or may not be significant, and therefore all possibilities for boundary inclusion into a pattern should be permitted.

Concerning the raw output of candidate pattern generation, it was found during manual review of the Polynesian data that:

1. Some patterns suggested poetic organization, however the majority – around 90% – held little or no interest.

2. Some patterns pointed to a larger, more comprehensive pattern.

3. Some patterns pointed to a pattern that might be detected better at a different level of analysis.

Patterns demonstrating some degree of promise were filtered through the following criteria:[5]

1. Similar types of pattern should be either non-existent or significantly less frequent in prose.

---

[2] See also the discussion of combinatorial explosion in 2.4 below.

[3] For example, in the treatment of the Tahitian and Mangarevan passages which will be presented below, long and short vowels have been conflated, as it was discovered early on in manual analysis that patterns could be extended, or those near to each other joined, by permitting such an abstraction. It was also discovered that the Tahitian and Mangarevan diphthong /ae/ is poetically equivalent to /ai/, and the Tahitian /ao/ to /au/. Poetic equivalence of /ae/ to /ai/ has been similarly observed by Jacob Love to apply to Samoan rhyme (Love 1991:88). Finally, the glottal stop phoneme /ʔ/ was determined to serve no role in Tahitian poetic function.

[4] A tradition's poetic line must be established before line boundary may be included as a pattern element. Nigel Fabb suggests that the concept of line is a poetic universal (Fabb 2009:54-55). It generally represents a syntactic structure with a specific metrical count, although for some traditions it may be non-metrical, bounded by some indicator such as a pause or lengthened vowel. Its identification, perhaps through trial and error, should be accomplished early on in the analysis.

[5] These criteria were empirically motivated mostly from analysis of the Polynesian data, and so may evolve after the described process has found application to a wider variety of traditions.

2. A pattern should occur multiple times in the same text. A longer pattern need only occur twice in the same text.

3. The placement of the majority of a pattern's occurrences should appear intentional; for example, as when found principally in the same segment of a text, or when placement suggested some higher degree of coordination.

4. Similar types of pattern should be found in at least two other texts of the same genre, in order to filter out patterns which might be unrepresentative of the poetic tradition, or which were perhaps merely the result of chance occurrence.

Poetic intent might subsequently be asserted if either of the following were satisfied:

1. The candidate pattern was found to match any method of poetic organization documented for other of the world's poetic traditions.

2. For promising pattern types unspecified in the literature, a pattern might be esteemed to self-justify as poetic were it found to be sufficiently complex or repetitive so as to eliminate the likelihood of chance.

The following sections will discuss single-feature pattern detection, multi-feature pattern detection, and detection of inverted parallelism. Examples will be provided of application of the procedure to a passage from a familiar English children's poem, and to extracts from several of the transcribed Tahitian and Mangarevan oral texts.

2.1 Single-Feature Pattern Detection

In single-feature pattern detection, only one linguistic feature is analyzed at a time. As with the other detection methods, the possibility exists of poetic intent whenever an n-gram token recurs.

The first four lines of the well-known children's poem *Mary had a little lamb* will serve to initially demonstrate this type of analysis. The passage in (1) has been tagged for three word-level linguistic features: *IPA word form*, *simple part-of-speech*, and *word syllable count*.

(1) Passage from *Mary had a little lamb* tagged for word form, simple part-of-speech, and word syllable count

1. *Mary   had   a   little   lamb*
   mɛɹɪ   hæd   ə   lɪdəl   læm
   NOUN   VERB   FUNC   MODIF   NOUN
   2   1   1   2   1

2. *whose   fleece   was   white   as   snow*
   huːz   fliːs   wəz   waɪt   æz   snoʊ
   FUNC   NOUN   VERB   MODIF FUNC NOUN
   1   1   1   1   1   1

3. *and   everywhere   that   Mary   went*
   ænd   ɛvɹiwɛɹ   ðæt   mɛɹɪ   wɛnt
   FUNC   NOUN   FUNC   NOUN   VERB
   1   3   1   2   1

4. *her   lamb   was   sure   to   go*
   hɚ   læm   wəz   ʃɚ   tu   goʊ
   FUNC   NOUN   VERB   MODIF FUNC VERB
   1   1   1   1   1   1

The list of bi-gram word form tokens from this passage would begin:

mɛɹɪ-hæd
hæd-ə
etc.

The list of 4-gram simple part-of-speech tokens would begin:

NOUN-VERB-FUNC-MODIF
VERB-FUNC-MODIF-NOUN
etc.

From a tally of matching simple part-of-speech bigrams, we note in (2) below four occurrences of *NOUN-VERB*.

(2) Some bigram repetition in the *Mary had a little lamb* passage
Level of analysis: Word
Linguistic feature: Simple part-of-speech
Boundary relevance: Line boundary is significant.
Minimum pattern occurrences = 4

1. *Mary   had   a   little   lamb*
   NOUN   VERB FUNC   MODIF   NOUN

2. *whose   fleece   was   white   as   snow*
   FUNC   **NOUN   VERB** MODIF FUNC NOUN

3. *and   everywhere   that   Mary   went*
   FUNC   NOUN   FUNC   **NOUN   VERB**

4. *her   lamb   was   sure   to   go*
   FUNC   **NOUN   VERB** MODIF FUNC VERB

With prior knowledge that English is an SVO language, however, the *NOUN-VERB* pattern

candidate is dismissed during manual review as being common as well to English prose.[6]

In (3) below, we find repetition of the word syllable count 11-gram: *1-2-1-/-1-1-1-1-1-1-/*,[7] corresponding to *a little lamb | whose fleece was white as snow |*, and *that Mary went | her lamb was sure to go |*.

(3) 11-gram repetition in the *Mary had a little lamb* passage
Level of analysis: Word
Linguistic feature: Word syllable count
Boundary relevance: Line boundary is significant.
Minimum pattern occurrences = 2

1. *Mary had a little lamb*
   2   1   **1 2**   1

2. *whose fleece was white as snow*
   1   1   1   1   1   1

3. *and everywhere that Mary went*
   1   3   **1   2**   1

4. *her lamb was sure to go*
   1   1   1   1   1   1

It may be that parallelism of such a long pattern is metrically significant, although this would be difficult to confirm given just one recurrence. It should be reiterated that while patterns which emerge out of a single text are not always conclusively poetic, when compared with similar pattern organization in other texts, poetic intent often becomes clear.

In (4), we turn to analysis at the syllable level. Here, we discover the apparent end-rhyming bigram */oʊ/-/* of *snow |*, and *go |*.

(4) Some bigram repetition in the *Mary had a little lamb* passage
Level of analysis: Syllable
Linguistic feature: Syllable rhyme
Boundary relevance: Line boundary is significant.
Minimum pattern occurrences = 2

1. *Mary had a little lamb*
   ɛ ɪ   æd   ə   ɪ əl   æm

2. *whose fleece was white as snow*
   uːz   iːs   əz   aɪt   æz **oʊ**

3. *and   everywhere that Mary went*
   ænd   ɛv i ɛɹ   æt   ɛ ɪ   ɛnt

4. *her lamb was sure to go*
   ɚ   æm   əz   ɚ   u   **oʊ**

With prior knowledge that end-rhyme on alternating lines is common to English, French, and several other poetic traditions, we conclude that the intent here is poetic.

In (5), we encounter a passage of a Mangarevan song[8] which consists of a repeated syntactic frame, with the four nouns *vai*, *kukau*, *aʔi,* and *inaina* and the two adjectives *rito* and *ka* serving as its variable elements. We observe end-rhyme in lines 1 and 5 with the syllable rhyme pattern ***a-i/*** (in bold) corresponding to the nouns *vai* and *aʔi*, and note that ***a-i*** as a bigram is also contained within the name of the song's subject, the young woman Tai-tinaku-toro. We additionally observe assonant matching between the syllable rhyme bigram ***a-u*** (in bold underlined) of the noun *ku.ka.u* and the syllables *na.ku* of the woman's name. Finally, we note a match between the syllable rhyme bigram ***ɪ-ᴀ*** (in bold small caps) of the noun *i.na~i.na* and the syllables *ti.na* of the woman's name.

(5) Some bi- and tri-gram repetition in an extract of a Mangarevan song (Buck 1938:170)
Level of analysis: Syllable
Linguistic feature: Syllable rhyme
Boundary relevance: Line boundary is significant.
Minimum pattern occurrences = 2

1. *ko   te   vai*
   o   e   **a.i**
   EXIST   the   fresh.water
   The water

2. *e   rito   nei*
   e   i.o   e.i
   IPFV   clear here
   that is clear here,

3. *ko   te   kukau   ia*
   o   e   u.**a.u**   i.a
   EXIST   the   bath   ANAPH
   it is the bath

4. *o*　　　　　　*tai-tinaku-toro*
　 o　　　　　　　**a.i I.<u>A</u>.u**　o.o
　 INALIEN.WEAK　Tai-tinaku-toro
　 of　　　　　　　Tai-tinaku-toro.

5. *ko*　*te*　*aʔi*
　 o　　e　　**a.i**
　 EXIST　the　fire
　 The fire

6. *e*　*ka*　*nei*
　 e　　a　　e.i
　 IPFV　lit　here
　 that is lit here,

7. *ko*　*te*　*ina~ina*　　*ia*
　 o　　e　　**I.A I.A**　i.a
　 EXIST　the　drying.agent　ANAPH
　 it is the drying agent

8. *o*　　　　　　*tai-tinaku-toro*
　 o　　　　　　　**a.i I.<u>A</u>.u**　o.o
　 INALIEN.WEAK　Tai-tinaku-toro
　 of　　　　　　　Tai-tinaku-toro.

If similar use of assonance were discovered in several other texts of the same genre, such should warrant a claim that *assonant matching between a syntactic frame's variable elements and the poem's theme* is a method of Mangarevan poetic organization.

2.2　Multi-Feature Pattern Detection

In multi-feature analysis, n-gram patterns are comprised of cross-level linguistic feature information. This is motivated by a desire to be sufficiently expansive so as to detect a poet's efforts to coordinate more than one feature.[9]

In the *Mary had a little lamb* passage, the addition of a bit of manual semantic tagging reveals the following multi-feature tri-gram:

Semantics: lamb-part -
Word form: wəz -
Part-of-speech: MODIF

The tri-gram token is provided in context in (6):

(6) Some multi-feature trigram repetition in the *Mary had a little lamb* passage
Level of analysis: Word
Linguistic features: Word form, simple part-of-speech, and "Mary-part" and "lamb-part" semantic tagging
Boundary relevance: All boundaries are ignored.
Minimum pattern occurrences = 2

1. Mary　had　a　little　lamb
　 mɛɹɪ　hæd　ə　lɪdəl　læm
　 NOUN　VERB　FUNC　MODIF　NOUN
　 Mary-part　　　　　　　　　lamb-part

2. whose fleece was white as snow
　 huːz　fliːs　**wəz**　waɪt　æz　snoʊ
　 FUNC NOUN　VERB　**MODIF**　FUNC NOUN
　 lamb-part

3. and　everywhere　that　Mary　went
　 ænd　ɛvɹiwɛɹ　　ðæt　mɛɹɪ　wɛnt
　 FUNC NOUN　　　FUNC　NOUN　VERB
　 　　　　　　　　　　　Mary-part

4. her　lamb　was　sure　to　go
　 hɚ　læm　**wəz**　ʃɚ　tu　goʊ
　 FUNC NOUN　VERB　**MODIF**　FUNC VERB
　 lamb-part

Whether or not the recurrence of this tri-gram might be interpreted as poetic, it should be recognized that it would not have been detected by single-feature analysis.

From the Tahitian corpus, we find an 11-gram multi-feature token which combines information relevant to word form, syllable count, and word vowel:

Line boundary -
Word form: e -
Word form: noho -
Line boundary -
Syllable count: 1 -
Syllable count: 2 -
Line boundary -
Word form: i -
Word form: te -
Word vowels: a-o-a -
Line boundary

---

[9] Multi-feature detection was originally inspired by the bag of trees approach used by Data-Oriented Parsing, which permits assembling syntactic patterns from different levels of tree structure (see Bod 1998).

This token appears initially in lines 1 through 3 and then repeats in lines 4 through 6 of (7) below:

(7) Some multi-feature 11-gram repetition in an extract of "Warning by messengers of the paʹi-atua service" (Henry 1928:158-159)
Level of analysis: Word
Linguistic features: Word form, word vowel, syllable count
Boundary relevance: Line and word boundaries are significant.
Minimum pattern occurrences = 2

1. *e*   **noho**
   e   o o
   1   2
   IPFV   sit
   Sit

2. *i*   *niʔa*
   i   i a
   1   **2**
   at   above
   on

3. *i*   *te*   *maːhora*
   i   e   **a o a**
   1   1   3
   at   the   yard
   the yard,

4. *e*   **noho**
   e   o o
   1   2
   IPFV   sit
   Sit,

5. *e*   *ʔupu*
   e   u u
   1   **2**
   IPFV   recite.a.prayer
   recite

6. *i*   *te*   *ʔahoːʔa*
   i   e   **a o a**
   1   1   3
   DIROBJ   the   brush.clearing.prayer
   the brush clearing prayer.

It might be best to re-interpret this complex n-gram as simply providing evidence of two overlapping methods of organization: A *3-3-5* pattern of syllabic counting meter alongside an *a-o-a* pattern of end-rhyme. During manual review, an attempt should always be made to re-analyse candidates into more generalizable patterns.

From the Mangarevan material, we find a 15-gram multi-feature token which combines information relevant to word form, syllable form, syllable onset, and syllable rhyme:

Line boundary -
Word_form: ena -
Word_form: ʔana -
Line boundary -
Word_form: i -
Syllable_onset: t -
Syllable rhyme: a -
Syllable_form: vae -
Syllable_rhyme: e -
Syllable_rhyme: u -
Syllable_rhyme: a -
Line boundary -
Word_form: te -
Word_form: u -
Line boundary

In (8) below, this 15-gram comprises lines 1 through 3, and then repeats in lines 4 through 6:[10]

(8) Some multi-feature 15-gram repetition in an extract of a *rogorogo* chant (Buck 1938:114)
Levels of analysis: Word and syllable
Linguistic features: Word form, syllable form, syllable onset, syllable rhyme
Boundary relevance: Line, word, and syllable boundaries are significant.
Minimum pattern occurrences = 2

1. *ena*   *ʔana*
   e.na   ʔa.na
     n   ʔ n
   e.a   a.a
   DEM.PROX.2   now
   See there now

2. *i*   *tua-vai-heua*
   i   tu.a-**vai**-he.u.a
                      t   v h
   i   u.**a** a.i **e.u.a**
   at   Tua-vai-heua
   at Tua-vai-heua is

3. *te*   *u*
   te   u
   t
    e   u
   the   milk
   the milk.

97

…

4. *ena*   *ʔana*
   e.na   ʔa.na
     n   ʔ n
   e.a   a.a
   DEM.PROX.2   now
   See there now

5. *i* *te* *vavae* *ʔenua*
   i   e   va.**vae** ʔe.nu.a
     t   v v   ʔ n
   i   e   **a**.ae   **e.u.a**
   at the leg   land
   coming over land is

6. *te*   *u*
   te   u
   t
    e   u
   the   milk
   the milk.

The repeated word forms within this 15-gram perhaps serve to bracket the sound parallelism which occurs between its variable elements.

It should be noted that, with regard to the Polynesian data, the discovery of poetic organization was generally achievable through single-feature analysis. Patterns only detectable through multi-feature analysis were uncommon.

2.3 Inverted Parallelism

In some poetic traditions, patterns of linguistic features are not always repeated as is, but rather by means of an inverted ordering. An example is *chiasmus*, which is an inversion of repeated semantic elements; very common to the Ancient Hebrew of the Old Testament.

Automated detection of inverted parallelism is accomplished by a simply comparing the linguistic feature n-grams of a given document with the n-grams generated from a reverse ordering of those features. As before, matching n-grams are sorted and counted, and then presented within the context of the non-reversed material.

In the Tahitian example given in (9) below, we find the 7-gram pattern of syllabic counting meter *6-4-5-3-3-3-4* which is followed, after a *5* count, by its inverted match *4-3-3-3-5-4-6*.

(9) Inverted 7-gram repetition in an extract of "The genealogies of the gods" (Henry 1928:355-359)
Level of analysis: Word
Linguistic features: Line syllable count
Boundary relevance: Line and word boundaries are significant.
Minimum pattern occurrences = 2

6 1. *e*   *atua*   *anaʔe*
           1   3   3(2)[11]
           EXIST   god   all
           They were all gods,

4 2. *te*   *tahuʔa*
           1   3
           EXIST   artisan
           the artisans

5 3. *ʔe* *te* *ʔa:rere*
           1   1   3
           and   the messenger
           and the messengers.

3 4. *e*   *mana*
           1   2
           EXIST   power
           Power

3 5. *to:*   *ra:tou*
           1   2
           INALIEN.NEUT   3.PL
           was theirs

3 6. *i*   *te* *po:*
           1   1   1
           in   the   Realm.of.Darkness
           in the Realm of Darkness

4 7. *ʔe* *i*   *te* *ao*
           1   1   1   1
           and   in   the   world
           and in the world.

5 8. *rahu-a*   *mai*   *ra*
           3   1   1
           conjure-PASS hither there
           He was conjured forth,

4 9. *te*   *atua*
           1   3
           the   god
           the god

_____

[11] The syllable count for *anaʔe* in this context has been reduced to 2, as the last /a/ of the preceding word *atua* and the first /a/ of *anaʔe* merge to form a single long [a:], having a syllable count of 1.

3 10. *ʔo*   *raʔa*
        1   2
        PROP Raʻa
        Raʻa.

3 11. *e*   *moʔa*
        1   2
        EXIST sacredness
        Sacredness,

3 12. *e*   *mana*
        1   2
        EXIST power
        power,

5 13. *e*   *hana~hana*
        1   4
        EXIST glory
        glory,

4 14. *e*   *mau*   *riri*
        1   1   2
        EXIST to.hold anger
        the ability to hold anger

6 15. *to:*   *raʔa* *atua*
        1   2   3
        INALIEN.NEUT Raʻa god
        belonged to the god Raʻa.

Due to the detection as well of many other patterns of inverted meter in the corpus, *inversion of the patterns which govern syllabic counting meter* was deemed to self-justify, under criteria mentioned above, as a method of Tahitian poetic organization.

2.4   Concerning Combinatorial Explosion

Inherent to the automated process is a combinatorial explosion of n-grams – particularly true with regard to multi-feature analysis. The total number of single- and multi-feature n-gram tokens generated for a given text may be determined as described in figure 1.

The number of word-level n-grams generated from a typical 1,000 word text, after restricting analysis to 10 layers of linguistic feature tagging, is the quite large 6.82×10^{501}. By reducing the number of tagged layers to four and maximum *n* to 10, however, the final count diminishes to a much more tractable 1.39 billion. It should be mentioned that foregoing multi-level analysis would permit maximum *n* to be set much higher.

It follows that a reduction in the interaction of linguistic features for a given pass would result in some patterns being missed by the automated

Figure 1. Calculation for all single- and multi-feature n-gram tokens of a text.[12]

Given:

C = The count of all single- and multi-feature n-gram tokens which might be generated from a text at a given level of analysis (e.g. *word level*, *syllable level*).

E = The number of linguistic elements in the text (e.g. in the passage from *Mary had a little lamb*, we analysed at the word level where there are 22 words and 5 instances of *line boundary*, for a total of 27 word-level elements).

N = The current n-gram *n* number.

Max N = The *n* number of the largest desired n-gram. For an n-gram token to be able to occur at least twice, and thereby potentially demonstrate a pattern, *max n* should not exceed half the total number of linguistic elements (e.g. for word-level analysis of the passage from *Mary had a little lamb*, it would not be useful for *n* to be larger than 13).

F = The number of tagged linguistic features (e.g. the passage from *Mary had a little lamb* in (1) is tagged for three features).

$$ C = \sum_{N=1}^{Max\ N} (E - (N - 1)) \cdot F^{N} $$

process. Therefore, a certain degree of trial and error must be pursued in order to determine which combinations of four features at a time yield the best candidates. Furthermore, with a maximum *n* of just 10, it may become necessary to stitch together – either manually or through an automated process – adjacent and overlapping patterns.

3   Conclusion

Alongside grammatical description, dictionary compilation, language pedagogy, and the other efforts typically undertaken to assist in the preservation of an endangered language, it is important that documentation of a language community's oral tradition, and the poetics commonly embedded therein, be awarded full consideration as well. Both of these serve a significant role in language identity, and the realm of verbal

---

[12] To arrive at the count needed to include analysis of inverted parallelism as well, simply double the *C* result.

art is enriched through their study. Relevant to the level of detail required for such research, John Miles Foley asserts that "We must give the idiosyncratic aspects of each tradition their due, for only when we perceive sameness against the background of rigorously examined individualized traits can we claim a true comparison of oral traditions" (Foley 1981:275).

The procedure which has been described here is admittedly labor-intensive; especially with regard to its manual component. However, it is probably necessary that it be so in order to succeed at documenting the majority of a poetic tradition's *individualized traits*. Relevant to the Tahitian material, the procedure was successful at the detection of a syllabic counting meter based upon word stress (see Meyer 2013:88-105). Such was previously unattested among world poetries, and with its discovery our understanding of what is universally possible for meter became expanded.

References

Rens Bod. 1998. *Beyond grammar: An experience-based theory of language*. Center for the Study of Language and Information, Stanford, CA.

Peter H. Buck. 1938. *Ethnology of Mangareva*. Bernice P. Bishop Museum Bulletin 157. Bishop Museum Press, Honolulu, HI.

Nigel Fabb. 2009. Symmetric and asymmetric relations, and the aesthetics of form in poetic language. *The European English Messenger,* 18(1):50–59.

John Miles Foley. 1981. Tradition-dependent and -independent features in oral literature: A comparative view of the formula. In John Miles Foley, editor, *Oral traditional literature, a festschrift for Albert Bates Lord*, pages 263–281. Slavica Publishers, Columbus, OH.

Teuira Henry. 1928. *Ancient Tahiti*. Bernice P. Bishop Museum Bulletin 48. Bishop Museum Press, Honolulu, HI.

Jacob Wainwright Love. 1991. *Sāmoan variations: Essays on the nature of traditional oral arts.* Garland Publishing, New York, NY.

David Meyer. 2011. A computationally-assisted analysis of early Tahitian oral poetry. PhD Dissertation, Department of Linguistics, University of Edinburgh.

David Meyer. 2013. *Early Tahitian poetics*. De Gruyter Mouton, Boston, MA.

Calvert Watkins. 1995. *How to kill a dragon: Aspects of Indo-European poetics*. Oxford University Press, New York, NY.