<u>Expressive Alliteration in Mon and Khmer</u>
Christian DiCanio
dicanio@berkeley.edu
University at California, Berkeley

I.   <u>Introduction:</u>

One need not look very deeply into Mon-Khmer languages to discover a plethora of expressions which are used to convey specific manners of motion, sound, feeling, and gesture. These types of descriptive words are often called "expressives" in the Southeast Asian linguistic literature (Diffloth, 1976). In fact, in some languages, like Semai (Aslian, Mon-Khmer; ibid), there is ample morphological, syntactic, and semantic evidence to treat these words as a separate lexical category. There also tends to be phonological evidence for the existence of expressives as a class; they tend to involve some degree of phonological repetition.[1] This is certainly the case for expressives in Khmer which are "par excellence" reduplicative compounds (Jacobs, 1992). While expressives are not limited in this way to repetitive structures, the reverse is indeed true; phonological repetition will almost always involve some type of "expressive" meaning in Mon-Khmer languages. The pattern of repetition used in expressive compounds in Mon-Khmer languages is almost exclusively reduplication with vowel ablaut, or reduplication with rime alternation. Both patterns maintain the same onset between the reduplicant and the base. Hence, I am calling all such patterns "alliterative."

It is usually the case that repetition is used in more descriptive speech or as a poetic device (Jacobs, 1979) in languages of the world. The use of such English expressions as "dilly-dally," "zig-zag," or "pish-posh," is *limited* to more "decorated" speech. However, for speakers of Mon-Khmer languages, phonological repetition is a salient use of language which appears in everything from regular day-to-day speech to formal prose (ibid, Haiman & Ourn, 2001, Bauer, 1989). Whereas English has a relative handful of such words, Mon-Khmer languages have hundreds or thousands. Any serious investigation into the grammar of these languages will include an investigation of these words. Throughout the Mon-Khmer family, this process of alliteration (or reduplication) is either morphologically productive or lexicalized. In contrast to alliteration in Indo-European languages which is marginal in the grammar, alliteration is a common way that phonological repetition takes place in Mon-Khmer languages.

This paper is a critical examination of the phonology, morphology, and semantics of alliterative compounds in Mon and Khmer. I will focus exclusively on compounds which have vowel alternations, which are the most common types of expressives in these languages. In a sense, this paper is exploring the many ways in which one could organize the data on alliterative compounds; i.e. by their morphology, semantics, or phonology. I have built on earlier findings by Schiller (1999) and will provide evidence in favor of the claim made by Haiman and Ourn (2001) on the morphological shape of these compounds. Specifically, alliterative compounds are unique in Mon and Khmer because they fit a 2/4 syllable prosodic template with a "weak-strong" pattern. In addition to other morphological characteristics, alliterating or ablauting compounds compounds in Mon and Khmer are related to the grammatical category of "elaborate expressions" (Matisoff, 1973) in these languages, and often denote repetitive or mimetic

---

[1] While the category "expressive" includes both onomatopoeic and other reduplicative words, it is not limited to these types.

semantics, similar to reduplication. All these factors make alliterative compounds a grammatically interesting lexical category that deserves thorough investigation.

In Section II, I review previous research done on expressives in Mon and Khmer and I summarize the method I used in organizing the data from both languages. In Sections III and IV, I analyze Mon and Khmer, respectively. In Section V, I will expound on the similarities and differences between alliterative compounds in both languages and discuss the implications that these observations have for the rest of the Mon-Khmer family. In Section VI, I conclude with some remarks on directions for future research.

II.   Background

2.1 The Phonology of Mon-Khmer Languages

Since the data in this paper involves rime-alternation, it is important to give a proper definition of the term "rime." Ancient and modern phonological theory divides syllables into two components, an onset and a rime, the latter of which includes the vowel and final consonant(s) of the syllable. There is a very old grammatical tradition in the languages of the "Sinosphere" (Matisoff, 2003) to view these parts of the syllable as primary rather than individual phonemes, which are given primary treatment in many modern (and non-Sinospheric) grammars. Thus, languages in East and Southeast Asia are often described as having an inventory of possible "initials" (Chinese 'shēngmǔ') and an inventory of possible "rimes" (Chinese 'yùnmǔ'). There are many reasons to think of such units as primary in these languages. In Mon-Khmer languages for instance, the presence of voice-register on a word was historically determined by the onset consonant type (Ferlus, 1979). The rimes in Mon-Khmer words have no historical influence on register. Furthermore, onsets do not affect the stress of a word, but rimes do. In alliterative compounding in Khmer, for instance, the location of stress occurs where we find longer vowels and more sonorant codas; i.e. in a longer rime. In the compound /kɑntrek-kɑntraak/ 'to hang in tatters,' the compound stress occurs on the final word /kɑntraak/, which has a longer vowel than /kɑntrek/. The complexity of the onsets in such words does not affect the stress. These are only a few of the factors which signify the importance of the 'rime' in the phonology of Mon-Khmer languages (Henderson, 1965).

It is also important to provide some explanation of what is being considered a coda in this work. The final consonant in a rime in Khmer and Mon may be a stop, nasal, liquid, or glide. Many Mon-Khmer languages have large vowel inventories which include a large set of diphthongs. In Khmer, many of these diphthongs come from monophthongal vowels which changed into diphthongs due to the voice register distinction (Ferlus, 1979). Thus, Khmer has the following diphthongs: /iə/, /ae/, /ao/, /aə/, /ea/, /ɨə/, /uə/, /oa/, and /ɔə/, all of which may be followed by a coda consonant. There are a number of other diphthongs which end in glides[2] : /aʊ/, /əʊ/, /eeʊ/, /ɨʊ/, /ay/, /əy/, /ey/, /oy/, and /uy/. None of these diphthongs are followed by a coda consonant; i.e. the glide [j] or [ɥ] is the coda itself. I treat the first set of diphthongs like a single vowel, but the second set like a vowel with a coda.

---

[2] The consonant /ʋ/ occurs as [ʊ] in onset position, but as [ɥ] as a coda (Henderson, 1952).

2.2 Reduplicative Compounding in Khmer and Mon

Phonological repetition is realized in a number of ways in Mon-Khmer languages. There is whole-word reduplication, reduplication with some vowel or consonant alternation, reduplication with vowel ablaut, rhyming reduplication, and many types of repetition of more than two words that involve maintaining one word in an alternating position; the so-called elaborate expression A-B-A-C (or A-B-C-B) pattern with four syllables (Jacobs, 1968). The two patterns that are most common in Mon and Khmer are those which involve a change in the vowel. This includes both cases of alliteration, where the initial consonant remains the same and the rime changes across stems, henceforth 'alliterative rime-alternating,' and cases of vowel ablaut where the stems of the compound[3] are identical except for their vowels, henceforth 'alliterative vowel-alternating.' I offer examples of each type in (1) and (2):

		Khmer Alliteration with rime alternation		*Mon Alliteration with rime alternation*
(1)	a. khsək-khsuəl	'in a sobbing or weeping manner'	c. sàŋ-sùi [4]	'to be deliberate in movement, ponderous'
	b. rələəp-rəluəŋ	'shining and wet'	d. bɔn-boa	'to be interlaced'

		Khmer Alliteration with vowel alternation	*Mon Alliteration with vowel alternation*
(2)	a. kɑntrek-kɑntraak	'to hang in tatters'	c. kəmɛt-kəmot 'piece of burning wood'
	b. cheeʋ-chaaʋ	'to be sizzling'	d. kraŋ-krɛŋ 'to wander about'

In (1), we notice an alternation in the rimes between the two stems of the compounds. For example, the word in (1d) has an initial stem ending in a nasal while the final stem has no coda consonant. In (2), we notice an alternation in only the vowels of the alternating stems. For example, in (2a), the final stem /kɑntraak/ differs from the initial stem /kɑntrek/ by its vowel alone. In both (1) and (2), the vowel in the ultimate syllable of the first stem is distinct from the vowel of the ultimate syllable of the final stem. These two patterns occur quite often in both Khmer and Mon and are the focus of this paper.[5]

While a number of works on Khmer mention the pattern of vowel ablaut in reduplication or alliteration, there is no work on Mon exploring this phenomenon. The earliest work on vowel ablaut and alliteration in Khmer is Henderson (1952), who briefly mentions alliterative expressions, stating that they are produced as a succession of two monosyllables, and not as two disyllables. In this way, she implies that words like (1a) or (2b), which consist of monosyllabic stems, are not produced like a sesquisyllable (Matisoff, 1973), where the initial stem is extremely reduced and the coda consonant often deleted. For instance, one Khmer word for 'to give an order' is carefully pronounced (and written) as [ɓɔŋˈkoəp], but is often pronounced as a reduced

---

[3] For the purposes of this paper, I will use the word "compound" to refer to anything that is the result of the adjoining of two bases, bound or free, as the result of reduplication or compounding.
[4] Mon has a register distinction where vowels are phonemically breathy (grave accent) or clear (no accent).
[5] While vowel and rime alternating alliteration occur almost equally in Khmer, alliteration with whole rime alternation is much more frequent than vowel alliteration in Mon.

sesquisyllable [pəˈkoəp], where the initial consonant is devoiced, its vowel reduced, and its coda deleted. Henderson (1952) mentions that only words like this with the shape ***CV***.C(C)V(C) or ***CVN***.C(C)V(C) are pronounced as sesquisyllables[6]. The minor syllables of these words are either a CV syllable or a CVN syllable where the N represents a homorganic nasal. Words with the shape ***CCV(C)***.C(C)V(C) are never pronounced as sesquisyllables, but as a sequence of two major syllables. Words like these are full disyllables. For example, the word [əw.pʊk] 'father' is never pronounced with a reduced initial syllable, even though stress falls on the final syllable.

Sesquisyllabic words are prevalent in Mon-Khmer languages and prosodically different than the alliterative rime and vowel alternating compounds I will look at here, which are never pronounced this way. Thus, the compound [cheeʋ-chaaʋ] 'to be sizzling' in (2b) is never pronounced *[chə-chaaʋ], where the vowel reduces along with coda deletion. Rather, alliterative rime and vowel alternation often results in some prosodic distinction between stems that is salient to the listener based on vowel length, vowel quality, or rime sonority. These compounds have a prosodic shape of either 4 or 2 syllables, where the second syllable of *each* prosodic foot is stressed; giving these compounds an iambic pattern. Furthermore, while sesquisyllabic words tend to consist of the order 'bound + free morpheme' or may be monomorphemic, most alliterative and ablauting compounds vary somewhat in the order of how a free and bound morpheme are adjoined and may in fact consist of two free morphemes. Sesquisyllables, however, never consist of two morphological words.

The majority of the work on alliterative compounds in Khmer is descriptive in nature (Gorgonijev, 1963; Huffman, 1967, Jacobs, 1968, Nacaskul, 1978), but some authors have argued for both phonological tendencies (Schiller, 1999) and semantic tendencies in the data (Haiman and Ourn, 2001). The most extensive discussion is in the work of Judith Jacob (1968, 1979) who gives a detailed account of the processes, both phonologically and semantically. It is important to note that most authors have labelled alliterative rime and vowel alternation as types of reduplication. Jacob (ibid) calls the process shown in (2) "chiming" reduplication, where the vowel in the reduplicant is different from the vowel in the base. She relates these words to the chiming words in English expressions like "ding-dong, tick-tock, or knick-knack." She also mentions the pattern shown in (1) where the rimes in both the reduplicant and the base are different.[7] It is problematic to call these two patterns "reduplicative;" many compounds (as I have called them) result from the adjoining of two free morphemes, one bound and one free morpheme, or two bound stems. The notion of "reduplication" in modern phonological theory assumes that there is a free morpheme from which one derives a reduplicant. The reduplicant is derived and bound to the stem. This concept of reduplication does not fit the data on these compounds in Mon-Khmer languages because the "reduplicated" form often does not contain an independent base or may contain two independent bases that happen to alliterate when compounded. Therefore, I have chosen to call these two processes examples of compounding with "phonological repetition," a term which more neutrally states what they are. This reduplicative-like process of compounding is not uncommon in languages of Southeast Asia (Mortensen, 2005).

---

[6] Jacob (1968) calls these "restricted di-syllables" in contrast to "full disyllables."
[7] Huffman (1967) calls this pattern "pseudo-reduplication," which may convey the problem scholars have had with the term "reduplication" in studying Mon-Khmer.

It is tempting to think of such compounding patterns as phonologically "lawless" because there is no consistent phonological alternation between the compounded stems, bound or free. Furthermore, a bound stem can appear on the left *or* right side of a base to which it compounds. I claim that the processes of vowel and rime alternating alliteration in Mon and Khmer are prosodically rule-governed. In all cases, a prosodic template is upheld in the formation of such compounds in Mon and Khmer. This template may be filled with either "reduplicated" material from the base, or free morphemes available in the lexica of these languages. I will extrapolate on this claim more thoroughly in sections 3.2 and 4.2.

2.3 *Reduplication in Mon-Khmer*

As a family, Mon-Khmer languages contain all types of phonological repetition, including those mentioned above in section 2.2. Regardless of the regularity of the pattern of repetition, vowel ablaut is a common characteristic between the phonological structure of the base and that of the reduplicant. This pattern of ablaut applies when two alliterating free morphemes are concatenated in a compound as well as in cases where a "reduplicant" is 'created' out of a base. Despite this degree of segmental variation in the "reduplicant", there are a number of tendencies in the phonological shape and length of reduplicated rimes. For instance, in Khmer and Mlabri there is a much higher likelihood to find reduced vowels in the reduplicant if it is prefixal (Rischel, 1996)[8]. This pattern seems to correlate not only with the prosodic structure of Mon-Khmer words, which are heavily iambic and sesquisyllabic, but also with the morphological structure, which is prefixing. In some languages this reduction is the result of productive processes of reduplication, while in others it is the result of compounding or the lexicalization of such processes. For example, in some languages, like Semai or Bahnar (Diffloth, 1976; Banker, 1964), a productive process of reduplication results in ablaut. Data are shown in (6) and (7):

Bahnar (Bahnaric)

(3)	a.	’bok	'man'	>>	’bok-’bǎĭ	'not worthy to be called a man, but is'
	b.	koˇleng	'to look at'	>>	koˇleng-koˇlaĭ	'not worth looking at, but look anyway'

----

	c.	sɑ	'to eat'	>>	soˇsɑ	'to eat first, and then do something else immediately.'
	d.	tep	'to sleep'	>>	toˇtep	'to sleep first, and then do something else immediately'

Semai (Senoic; Aslian)

(4)	a.	klcwɛ̃c >>	klcwũc-klcwɛ̃c	'irregular flapping circular movements'
	b.	pradɛk >>	praduk-pradɛk	'noises of scattered small drops of rain falling on dry leaves or the roof'
	c.	mɲuy >>	mɲɛːy-mɲuːy	'people in a crowd raising their heads here

---

[8] I will expound on this tendency in great detail in section IV.

and there.'

We notice in (3) that the Bahnar pattern results in alliteration of the final syllable of the base in (3a) and (3b), which both have a deprecatory meaning. In these forms, the base is on the left and the reduplicant is suffixal. In (3c) and (3d), different semantics are involved and the reduplication is prefixal. In all the examples in (3), we find vowel alternations between the base and the reduplicant. In (4), the reduplicant appears to the left-edge of the base and differs from the base only in its vowel quality. This pattern is called 'antiphonic reduplication' by Diffloth (1976) and closely resembles the pattern we noticed from Khmer in (2). In fact, (4c) resembles a form in Khmer [ŋee-ŋɤɤ] 'nodding from side to side (of the head); confused, in a confused manner' in its repetitive nature, phonological shape, and semantics. Both compounds involve the phoneme /ŋ/, a front-back vowel distinction, and semantically involve the movement of the head. Whereas in Khmer this ablaut pattern is lexicalized, it is productive in Semai. The result of such lexicalization in Mon-Khmer languages is an abundance of words which alliterate and alternate vowels. Indeed, such a process must have existed for Khmer (Schiller, 1999), which has thousands of such compounds.

While the Bahnar reduplication pattern results in the same vowel in every reduplicant, so-called 'fixed segmentism' in much of the current phonological literature (Alderete et al, 1997), the pattern for Semai, Khmer, and Mon results in this "antiphonic" alternation (Diffloth, 1976) that is irregular. Part of the reason I have chosen to call these processes "alliteration with vowel or rime alternation" is that they do *not* involve "fixed vowel segmentism," but rather contain this irregular pattern. Such a pattern is not derivable from a simple phonological re-write rule applicable to all such alliterative compounds.

Recent work on the semantics of these word types is found in Vuori (2000) who offers an excellent overview of repetitive patterns in both East Asian and Southeast Asian languages. The specific ablauting processes shown above are not restricted to Mon-Khmer languages, but are found throughout all of Southeast Asia. They extensively occur throughout the Tibeto-Burman, Tai, and Miao-Yao languages. Of these latter groups, the Tai languages utilize this pattern the most extensively. This may be due to their close contact throughout history with the Mon-Khmer family. Many similar alliterative compounds in Thai have Khmer origin (Varasarin, 1984).

In his dissertation, Vuori examines nominal, adjectival, and verbal repetitive lexemes and phrases in East and Southeast Asian languages. While many of these languages are examined both phonologically and semantically, Vuori offers major insight into the semantic nature of these repetitive forms. First, when a nominal form is repeated, it tends to have distributive, plural, or emphatic meaning. Similarly, it often is used to express "diversity of referents." This pattern is present in Khmer, where the form for 'woman' /srəy/ has a plural or generalized meaning when reduplicated, /srəy-srəy/. Second, when an adjectival form is repeated, it tends to express intensity, continuity, or distributivity. Finally, verbal forms tend to express delimitation or continuity when they are repeated. This sense of continuity and intensity is expressed in Vietnamese reduplication. While Vuori mentions that Mon-Khmer languages contain patterns of alliteration and ablaut throughout the language family, his work on Mon-Khmer is limited to the Katuic and Vietic languages. In Vietnamese, for example, we find cases of alliterative compounding similar to those of Khmer and Mon shown in (2):

Vietnamese Alliteration (Vuori, 2000)

(5) a. nho$^{?}$ 'small' --> nho$^{?}$-nhắn 'slender'

  b. rao 'to announce' --> rêu-rao 'to spread scandal'

  c. láng 'to abound' --> lai-láng 'to overflow'

Vuori mentions that the patterns of alliteration and ablaut in Vietnamese tend to derive a more intensified meaning from the base. In (5b) and (5c), the stem and alliterated compound have related meanings and their only distinction appears to be that of intensity (and specific quality as well in 5b). In (5a) we notice that the free stem is the initial stem in the alliterative compound, which differs from the situation in (5b) and (5c). The irregularity in the location of the base here exemplifies an important point that Vuori makes about Vietnamese; we find morphological variation in both bound and free stems. Vietnamese does, however, favor a suffixal pattern where the first stem of the compound is free and the alliterative stem is final and bound. The variation shown above is typical of Mon-Khmer languages and the propensity for suffixation of stems in repetitive compounding mirrors what I will show for Mon.

2.4   Method

     While many authors writing on phonological repetition in Mon-Khmer give a limited set of examples, I thought that it would be extensively revealing to look at all examples of rime and vowel alternating alliteration in Mon and Khmer. To do this, I collected almost all such words in Headley's (1977) dictionary on Khmer, in total 1,184 compounds. Words were put in a database and organized based on phonological shape, morphological structure, and semantics. For the Mon data, I collected words by using Shorto's (1962) dictionary of Spoken Mon. I found 202 words that have alliteration with rime or vowel alternation. These words were all compiled in a database similar to the Khmer one. The details of how each entry was labelled are given in appendix A.

III.   Mon

3.1   Phonology

     Mon is a Mon-Khmer language belonging to the Monic branch of Mon-Khmer and is spoken in Burma. Similar to Khmer, alliterative compounds are quite frequent in Mon (Bauer, 1989). In this section, I will go into more detail on the phonological patterns found in the data, with respect to register, syllable structure, and vowel quality in alliterative compounds. It is worth noting that alliteration with a change in rime is far more common than alliteration with a change in the vowel in Mon compounds. Out of 202 alliterative compounds, 176 have a rime alternation while only 26 have a vowel alternation. Both patterns are shown in the data in (2).
     A salient feature of Mon-Khmer phonetics and phonology is the presence of voice-register, or phonation-type. The phonation-type distinction in Mon-Khmer arose historically from voiced initials which conditioned breathiness on the following vowel. The voiced initials were then devoiced, merging with the voiceless initials. Breathy phonation then became phonologically contrastive with clear phonation (Ferlus, 1979). As I mentioned in section 2.1, register is normally a characteristic of the phonological word. While many alliterative compounds maintain the same register across stems in Mon, this is not always the case. Yet, all compounds which show a change in register across stems happen to be composed of two free

morphemes[9] which both contribute to the meaning of the compound. In this sense then, these exceptional alliterative compounds are not exactly the result of any productive reduplicative process, but rather the result of the compounding of two free morphemes which happen to alliterate, as in (6).

Mon Register-Crossing Alliteration

	Alliterative Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(6)	a. poa-pò	'football match'	(festival, contest // ball, football)
	b. pao-pàk	'to wrap around'	(to bind a cord or cloth around // to put around)
	c. sὲʔ-seaŋ	'to be withered or dried up'	(to wither, be withered // to wither or dry up)
	d. klɔŋ-klah	'to be clear or untroubled'	(to ring out; resound // to be distinct or clear)
	e. pyɜ̀-pyòk	'to stir up a quarrel; make mischief'	(to incite to enmity // O[10])

The examples in (6a) – (6c) show that compounds which have different register have semantics which are predictable from the stems of the compound. In (6a), for example, the stem /poa/ means 'contest,' while the stem /pò/ means 'football.' These two words combine to create an alliterative coordinate compound whose meaning is transparent. I have included (6e) above to show that compounds with bound stems which maintain the same register do in fact occur in the data.

Only monosyllables and disyllables may alliterate in Mon compounds. When a disyllable alliterates in a compound, it is the final syllable that participates in the rime or vowel alternation. It is significant here that the initial syllable is identical between the stems. The examples in (6) consist solely of monosyllables. I show examples with disyllables in (7).

	Alliterative Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(7)	a. kəlɛt-kəlot[11]	'to be smooth, fluent, or free'	('to be smooth and fluent' // O)
	b. həma-həmòiŋ	'tailor'	('occupation-prefix' / 'seam, hem')
	c. pəre-pərɔŋ	'to pacify'	('to smooth' // O)

In all the data in (7), we see that the initial syllable of the disyllables contain the vowel /ə/. In contrast to Khmer, this vowel may have a morphological function in Mon. While some Khmer words with complex initials insert an anaptyctic vowel, as in /p̆teəh ~ pteəh/ 'house,' such short

---

[9] I have one exception to this rule, the word /ŋaiŋ-ŋòiŋ/ 'buttock,' (buttock/O). In the Mon data, a grave accent marks breathy phonation.

[10] The 'O' means that the stem is bound and has no independent meaning.

[11] The form /kəlɛt/ is reconstructed Old Monic (Dvaravati) *t(r)lɯɯt (Diffloth, 1984) and similar to the PTB form *g-lwat given in Matisoff (2003).

vowels have morphological function in Mon. There is a morphological distinction between initials of the shape /kl/ and those with the anaptyctic vowel /kəl/. No such distinction exists in Khmer (Christian Bauer, p.c.). In the Mon data, there are fewer alliterative compounds consisting of two disyllables than those consisting of two monosyllables. There are 151 compounds made up of a sequence of two alliterating monosyllables but only 46 compounds made up of a sequence of two alliterating disyllables. In other words, 75% of the Mon alliterative compounds consist of two monosyllables, but only 23% consist of disyllables. This tendency in the Mon data contrasts to Khmer which has more disyllabic (and sesquisyllabic) stems in alliterative compounds. However, it demonstrates that alliterative compounds are foot-based in Mon; compounds have either one iambic foot or two iambic feet. Out of 202 compounds, 197 have either 2 syllables or 2 disyllables. Only 5 words are composed of 3 syllables. There is a prosodic structure maintained in alliterative compounds in Mon.

In addition, alliterative compounds in Mon show a wide range of variation in the vowel type that alternates. Stems alternate in both vowel length and vowel quality.

	Alliterative Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(8)	a. sam-sot	'a type of snake'	(O // O)
	b. wǐ-wòa	'impending or potential calamity'	(O // 'malice, enmity')
	c. tòiŋ-tɛ̀ʔ	'to be dazed'	('to be dazed' // 'to be numb')
	d. hɜram-hɜrem	'fine dust'	('dust, powder' // O)
	e. pʰòiŋ-pʰɛ̀aŋ	'to be dazed; to be deprived of one's faculties as a result of shock'	('opium' // 'to be dazed; to be deprived of one's faculties as a result of shock')

In (8a), the vowel of the first stem is /a/ while the vowel of the second stem is /o/. This contrasts with (8d) which has the same rime as (8a), /am/, but followed by the final stem vowel /e/. The vowel alternation in Mon is not predictable. One often finds the same rime or vowel alternating with many other rimes or vowels in an alliterative compound. This same variation is found in the coda types across rimes in these words. For example, in (8e), the rime /òiŋ/ alternates with the rime /ɛ̀aŋ/, but in (8c), it alternates with /ɛ̀ʔ/. The salient characteristic of these types of compounds in Mon is the fact that there is always *ablaut* between words, where the initial stem's vowel is distinct in height or frontness from the final stem's vowel. The same holds true for the coda alternation, where the coda in the initial stem is distinct from the coda in the second stem.

When I examined the Mon data statistically, I determined that the stems do not regularly alternate with any phonological tendency regarding vowel height, backness, or length. In other words, high and low vowels are as likely to appear on the initial stem as they are on the final stem. In most of these compounds, singleton vowels occur in both stems with equal likelihood. I have constructed a chart, shown below, exemplifying the vowel patterns we find in the Mon data.

(9)   *Stem and Vowel Type in Mon Alliterative Compounds*

Initial Stem Vowels						Final Stem Vowels					
Front		Central		Back		Front		Central		Back	
i/ĭ	16			u/ŭ	11	i/ĭ	8			u/ŭ	6
e/ĕ	14	ɜ/ɜ̆	15	o/ŏ	17	e/ĕ	13	ɜ/ɜ̆	8	o/ŏ	18
ɛ/ɛ̆	21	ə	1	ɔ/ɔ̆	23	ɛ/ɛ̆	23	ə	1	ɔ/ɔ̆	20
a/ă	28			ɒ	8	a/ă	39			ɒ	5

In the table in (9), we notice that there is a slightly higher tendency for high vowels to occur in the initial stem of a compound than on the final stem of an alliterative compound. This pattern is not statistically relevant. The presence of more alliterative forms in Mon would allow a conclusion on vowel quality in compound stems to be drawn. Aside from this, there is no other discernible tendency shown in the data above.

On top of having 18 monophthongal vowels, Mon has 21 diphthongal vowels: ɔə, ɔe, ɔi, ao, ea, oa, ai, əi, əe, əɛ, əa, əɒ, əɔ, əu, əo, əui, ui, ɛ̀a, ài, òa, and òi. Since diphthongs change vowel quality, by their nature, I have separated them from the monophthongs. There are two ways to organize the data on diphthongs, by onset vowels and by offset vowels, shown in (10a) and (10b) respectively.

(10)   *Stem and Vowel Diphthong Type in Mon Alliterative Compounds*
        a.   *(organized by onset vowel)*

Initial Stem Vowels		Final Stem Vowels	
Diphthong	**Frequency**	**Diphthong**	**Frequency**
ao, ai, ài	16	ao, ai, ài	20
əi, əe, əɛ, əa, əɒ, əɔ, əu, əo, əui	1	əi, əe, əɛ, əa, əɒ, əɔ, əu, əo, əui	1
ɔə, ɔe, ɔi	6	ɔə, ɔe, ɔi	3
ea, ɛ̀a	7	ea, ɛ̀a	20
oa, òa, òi	9	oa, òa, òi	9
ui	1	ui	3

        b.   *(organized by offset vowel)*

Initial Stem Vowels		Final Stem Vowels	
Diphthong	**Frequency**	**Diphthong**	**Frequency**
ɔi, ai, ài, ui, òi, əui	15	ɔi, ai, ài, ui, òi, əui	22
ɔe, əe	4	ɔe, əe	3
ea, əa, ɛ̀a, oa, òa	13	ea, əa, ɛ̀a, oa, òa	26
ao, əo	8	ao, əo	5
əɛ, əɔ, əɒ, əu	0	əɛ, əɔ, əɒ, əu	0

The data in (10a) show no tendency for certain diphthongs to occur in initial stems over final stems. The data in (10b) show a slight tendency for diphthongs which end in a high vowel glide /i/ and a low vowel /a/ to be more common in final stems than in initial stems. I include this data in order to examine any potential tendency in the vowels in compounds in Mon, despite the fact that I do not have enough data to do any statistical analysis. As a whole, diphthongs occur with more likelihood on the second member of a compound than on the first; 40 compounds exist with initial diphthongs, but 56 exist with final diphthongs. This preference for diphthongs in final stems in Mon is rather slight. This particular pattern is quite different from Khmer, where diphthongs and long vowels are much more common on the second member of a stem than on the first. Khmer also differs from Mon in having a more unequal vowel distribution on initial stems than it does on final stems in alliterative compounds. I explore the pattern in Khmer in section IV.

3.2   Morphology

Mon alliterative compounds show four possible patterns. They may be composed of two free stems, a bound and free stem occurring in the order B-F or F-B, or two bound stems. Compounds which have two bound stems may consist of morphemes which no longer occur outside the compound or morphemes which have no semantic contribution to the meaning of the compound. Such compounds are idiomatic or exocentric.[12] As I mentioned in section 2.3, there is a great deal of variation with respect to bound and free morpheme order in how alliterative compounds are constructed in Mon-Khmer languages. These four morpheme orders in the alliterative compounds of Mon are shown below:

Alliterative Compound	Order	Gloss	Composition (Stem 1 // Stem 2)
(11) a. dɒp-de	B-B	'butterfly'	(O // O)
b. hɔcɜ-hɔcah	B-F	'to run counter, be contrary'	(O // 'to go towards, meet, oppose')
c. klɜ-kloiŋ	F-B	'wolf'	('dog' // O)
d. cao-cɒp	F-F	'to arrive back'	('to turn, go back' // 'to arrive, reach')

In (11a), the alliterated syllables are bound and do not occur outside of this compound. In (11b), the first stem is bound and does not occur outside of this compound. In (11c), the final stem is bound and does not occur outside of this compound. In (11d), both stems are free. Despite the fact that these compounds may be constructed in any order morphologically, there are tendencies in the data, shown in (12).

(12) Morphological Pattern:	Frequency:	Percentage:
Free+Free	78/202	38.6%
Free+Bound	69/202	34.2%

---

[12] I use the term exocentric here in a purely semantic fashion. If both stems of the compound have no discernible semantic relationship with the compound as a whole, it is exocentric.

Bound+Bound	33/202	16.3%
Bound+Free	22/202	10.9%

There are two tendencies in the data in (12). First, Mon prefers to have compounds composed of two free roots over bound roots. There are 78 compounds with the F-F pattern and 31 with the B-B pattern. Alliteration appears to be derived largely from free stems in Mon because there are relatively few completely idiomatic (B/B) forms. Second, alliterative compounds in Mon tend to have bound final stems. Out of 202 words, 69 consist of the F-B pattern while 22 consist of the B-F pattern. Thus, 76% of compounds *with a bound and free stem* will be ordered F-B, but only 24% will be ordered B-F. These tendencies are significant; $\chi^2(3) = 43.9$, $p < 0.00001$. Similar to Vietnamese, there is a preference in Mon for bound stems in compounds to be final, or suffixal.

Alliterative compounds in Mon have three types of semantic composition; endocentric, exocentric, and partially-endocentric. Endocentric compounds occur where both stems transparently combine to create a compound meaning; they are both heads to the compound. The semantics of such compounds are endocentric. An example of this sort is (11d) [cao-cɒp] 'to arrive back' where the first stem means 'to go back' and the second means 'to arrive.' Exocentric compounds occur where neither stem contributes to the *meaning* of the compound.[13] An example of this sort is /kɛŋ-ki/ 'centipede,' where the first stem means 'to stare' and the second means 'to bark at.' Clearly, neither stem is related to 'centipede.' Finally, compounds may be partially-endocentric, where one stem is the semantic head of the compound, but the other is not. An example of this sort is [klàk-klài] 'to seek diligently; to pry into' where the first stem means 'to emit or be overcome by smoke' while the second means 'to look for, search for, seek.' Only the second stem has a meaning related to the whole compound. Therefore, the compound is only partially-endocentric. These types of composition are organized in the table below.

---

[13] Some alliterative compounds are indeed exocentric, but the relationship between both morphemes is culturally and linguistically specific and unclear to the outsider's eye. For instance, the word [tĩ-tὲa] is a kind of tree. The first stem means 'thigh,' while the second has an unrelated meaning 'to run away or stray.' It is difficult to see how 'thigh' relates to 'tree' unless one looks closely at the meaning given for a particular compound. In this case, [tĩ-tὲa] is not just any tree, but "a kind of tree with a long straight trunk and a leafy top, resembling a duck's drumstick." Knowing this particularly specific definition aids us in figuring out two things. First, the second stem of the compound appears to be cognate with the Khmer word for 'duck,' /tiə/, which occurs in the second register, corresponding to the breathy register in Mon. Mon uses a different, unrelated word for 'duck,' but has retained this stem in an alliterative compound. Second, this definition provides us with a knowledge of the shape of the tree which makes the stem meanings of the compound seem endocentric. This is one of the interesting characteristics of alliteration in Mon-Khmer. It is often the case that an alliterative word appears opaque because we, as outsider linguists, have no access to the figurative use of the language. Indeed such problems could be greatly solved if one were to pursue an analysis of the metaphorical and conceptual semantics of Mon-Khmer languages. Finally, many stems that existed as free morphemes in the language may have been frozen in alliterative compounds where they are bound. Perhaps the prevalent use of alliteration in Mon aids in a process of reanalysing the stem to become bound to a particular compound even though it may no longer be used as a free morpheme in the language.

	Compound Type	Morphological Composition	Stem Type
(13)	a. exocentric	Bound + Bound	both non-heads
	b. endocentric	Free + Free	both heads
	c. partially-endocentric	Free + Bound / Bound + Free	free stem = head
			bound stem = non-head

In both exocentric and partially-endocentric compounds, a bound stem may be of two possible types. The first type is a stem with unrelated semantics that must be semantically bound to the compound, i.e. the stem [klàk] in the compound [klàk-klài] shown above. In these cases, the bound root may have no relationship to a freely attested root in the language. The stem [klàk] in the alliterative compound may be accidentally homophonous with the root meaning 'to emit or be overcome by smoke.'[14] There are also compounds which have stems which are not homophonous with free morphemes in the language. In partially-endocentric compounds, some stems are just "filler stems" which are attached to a morphological head. In the compound [cùn-cèa] 'a swing for a baby,' the first stem is unattested outside of the compound, but the second stem means 'net, lattice,' (the material with which a swing is made.) These "filler stems" make the pattern of compounding in Mon resemble a productive pattern of reduplication where a reduplicant is created out of the phonological material of the base. However, the phonological structure of these "filler stems" is irregular. The pattern is impossible to explain through a regular phonological rule that would apply uniformly within the language. I have adopted a broad concept of the term 'bound' in alliterative compounds in Mon by including both "filler stems" and semantic non-heads in this category.

The morphological structure of di-sesquisyllabic alliterative compounds in Mon is morphologically complex. They may consist of four morphemes instead of two. In many cases, the pre-syllable portion may consist of a prefix which has attached to a free or bound stem. For instance, the compound [pəlɒm-pəlaik] 'to destroy' consists of two sesquisyllables that are each bimorphemic. The stem [lɒm] is 'to be damaged' while the stem [làik] is 'to fall down.' The presyllable [pə] is a causative prefix that has attached to both stems. If these alliterative expressions are indeed compounds, we would *not* expect a prefix to apply to both stems separately. Rather, we would expect an output like *[pəlɒm-laik]. Yet, this is not what we find in alliterative compounds in Mon. Additionally, the re-application of an affix occurs on so-called "filler stems." For instance, the compound [pəre-pərɔŋ] 'to pacify' consists of the causative infix [ə], a stem [pre] 'to be smoothed, be appeased' and a filler stem [pərɔŋ] which is unattested (as is [prɔŋ]). This "filler stem" has the same prosodic structure as the initial stem, consisting of a sesquisyllable. The prefix appears to have re-applied to both stems of the compound. What exactly does it mean for a prefix to apply to a filler stem though? Why is a prefix applying to both the stems of a compound? I propose that the re-application of these affixes in di-sesquisyllabic alliterative compounds must be prosodic. In the first example, [pəlɒm-pəlaik], two free morphemes are prefixed and then adjoined to create a compound. In the second

---

[14] It is difficult to know whether these stems are actually related to an attested root elsewhere in the language. To find this out, one would need to look into the historical semantics of these roots and these compounds.

example, [pəre-pərɔŋ], a filler-stem is created that mimics the bimorphemic free stem. Both these patterns converge by having the same prosodic template.

If we view all alliterative compounds in Mon as belonging to a prosodic template, then we can account for the many patterns we have observed. Prosodic templates, or CV skeletons, are representations of the syllabic or segmental structure of a (morphological) word, represented by a string of consonants (C) and vowels (V)(McCarthy, 1981). The grammar determines which prosodic string is associated to a particular morphological construction. In the case of Mon, the morphological construction is the alliterative compound. The prosodic structure of such compounds is outlined in (14).

(14)   *Prosodic Template: Alliterative Compounds in Mon*
         di-sesquisyllabic: presyllable$_1$-(onset$_1$rime$_1$) + presyllable$_1$-(onset$_1$rime$_2$)
         di-syllabic: (onset$_1$rime$_1$) + (onset$_1$rime$_2$)

The alliterative pattern in Mon compounds consists of either repeated onsets or repeated pre-syllables in both stems. In the template above, the only contrasting component in these compounds is the rime of the final syllable, indicated by different subscripts.

There are a number of generalizations that fall out from this prosodic template in Mon. First, the prosodic template is insensitive to the morphological structure of the stems. The template may be filled with free or bound morphemes. This accounts nicely for the distinct patterns that we find in the data, where compounds have all types of morphological composition; F-F, F-B, B-F, and B-B. Since the F-F pattern occurs frequently, it appears that Mon prefers to insert alliterating morphemes already attested in the lexicon to create a compound. However, if such morphemes are not available, then the language creates them to fit within the phonological requirements of the template. Second, the reapplication of prefixes in di-sesquisyllabic compounds occurs in order to agree with the phonological requirements of the template. Thus, we expect prefixes to re-apply in such a template due to their phonological structure, not their morphology. Finally, this template fits with the stress system of Mon (and other Mon-Khmer) words. Mon has binary foot-structure (Bauer, 1982) which is iambic. The contrasting rime occurs in the stressed syllables of a di-sesquisyllabic compound, which is a prosodically strong position. As a result, the prosodic template not only accounts for the variation and complexity of the data but also correlates with general prosodic principles in Mon. Alliterative compounding in Mon is a case of phonologically-conditioned morphology.

3.21   Borrowing in Alliterative Compounds

Another interesting feature of these patterns is the degree to which they have been borrowed from other languages. There has been a long history of contact between Burmese and Mon, which has resulted in both acquiring a great many words and expressions from each other. This is particularly apparent in the set of alliterative and ablauting compounds in Mon. A great many of these compounds have been borrowed as a whole from Burmese, but others consist of combinations of native Mon stems with borrowed Burmese stems. Some examples are shown in (15):

(15)a. *Allit. Compound*   coa-cɔp
        Gloss   'to have affection for; cleave to'

	Composition	(O // 'to touch, adjoin')
	Derivation	Burmese /cwai/ 'to cleave to' + Mon /cɑp/ 'to touch, adjoin'

b. *Allit. Compound* phui-phɛk
   Gloss 'to mix together'
   Composition ('to mix in' // to mix, put together')
   Derivation Mon /phui/ 'be mixed' + Burmese /phak/ 'to associate'

c. *Allit. Compound* chen-chɛk
   Gloss 'to descend, to be handed down'
   Composition (O // 'to join or connect')
   Derivation Burmese /chaň/ 'to descend' + Burmese /chak/ 'to join or connect'

In this data, the composition of the compound represents how its stems relate to its meaning *synchronically*, while the derivation of the compound represents the *historical* origin of the stems. We see in (15a) that a native Mon stem was suffixed to an alliterating Burmese stem with similar semantics. The Burmese stem in fact does not occur in Mon as a free morpheme. In (15b), we see that a native Burmese stem is suffixed to a native and alliterating Mon stem with similar semantics. Both stems in fact occur as free morphemes in Mon. In (15c) we see that both stems come from Burmese, but only one occurs as a free morpheme in Mon.

There are a number of interesting generalizations to be made here. First, these data show us that the pattern of compounding is able to use both native and borrowed lexical stems, some of which even become bound as a part of a lexicalized compound, as in (15a) and (15c). Perhaps historically Mon contained the Burmese stem /cwai/ 'to cleave to' as a free morpheme, but it has since been lost or frozen in this alliterative compound. This explanation would closely gibe with what I suggested above; that certain stems may be frozen in alliterative compounds. Second, examples like (15c) may have been borrowed as complete units into Mon from Burmese. Since this example consists of two Burmese stems, this would not be surprising. The same thing has occurred with a handful of Mon alliterative compounds which come directly from Pali. I show some examples in (16):

(16) a. *Allit. Compound* thı̆ʔthan
   Gloss 'vow'
   Composition (O // O)
   Derivation from Pali /adhiṭṭhāna/ 'determination'

b. *Allit. Compound* pəɲɔe-pəɲɔp
   Gloss 'to pacify'
   Composition ('to make level or even' // 'to command, ordain')
   Derivation Mon /pɲı̄/ 'to be level' + Skt/Pali /pɲap, paɲap/ 'to command'

c. *Allit. Compound* cùn-cɛ̀a
   Gloss 'baby's swing'
   Composition (O // 'net, lattice, latticework')

Derivation   Mon */juɲ/ + Skt/Pali /jāla/ 'net'

In (16a), both stems of the compound are bound. This is not surprising given the compound's origin as a single word in Pali. In both (16b) and (16c), we notice that the Pali words, being borrowed at a very early period in Mon (at least before 700 C.E. (Diffloth, 1984)), are used to make up alliterative words in Mon. Alliterative compounding in Mon is not limited to Mon stems, but may contain stems borrowed from other languages. Compounds which consist of two bound stems often originate from compounds borrowed from Burmese or Pali.

The order of stems in such compounds appears to be always Mon stem + Sanskrit/Pali stem. However, there are non-alliterative compounds in Mon which have the opposite order. For instance, the compound [kaoʔsɒ-khɒh] 'to have good luck' is composed of two stems (freely attested in Mon), the first which derives from the Sanskrit stem [kuśala] and the latter of which derives from the Mon stem [khuih]. Given the presence of the Sanskrit + Mon pattern in regular, non-alliterative compounding, I am unable to make any claim on stem order in compounding here.

3.3 *Semantics:*

The most elusive aspect of alliterative compounds in Mon is semantic. The compounds look like phonological reduplication which would make us think that they have reduplicative-type semantics. However, these compounds do not fit neatly in such categories. A large part of this research is investigating just what categories seem to fit best with the data. Earlier in section 2.3, I mentioned the work of Vuori (2000) on the semantics of reduplication in Mon-Khmer. I first divided the data into semantic categories based on his work. Then, I found that I needed to include more categories and exclude others which did not fit. Finally, I realized that it is best to include two semantic categories for any one compound, one which categorizes the compound as a whole and the other which categorizes the derivational process or semantic change that the stems undergo when compounded. I did this because a compound may consist of bound stems which are unattested in the language. Such compounds do not therefore have any derivational semantics (classified as 'idiosyncratic'), but still fit within a combined compound semantic category. The table in (17) refers to the compound's meaning but the table in (18) refers to the change in semantics from a stem to a compound, i.e, it is derivational. I have two categories because I wanted to include all alliterative compounds in a study of their semantics. I outline the categories in (17) and (18):

	Combined Semantics Category	*Meaning*
(17)	stative	An attributive/stative verb
	motion mimetic	A compound representing a particular movement
	oscillatory	A compound representing a back and forth movement
	motion verb	A verb involving any other type of motion
	sound mimetic	Either an onomatopoeic compound or a compound describing manner of speech.
	shape/texture	A compound involving physical shape or texture
	plant/animal	A compound representing a plant or animal

object	A compound representing a physical or abstract object
entity	A compound representing a person or occupation
miscellaneous	An additional category comprising some transitive verbs, abstract nominals and abstract adverbs

(18)   *Derived Semantics*   *Meaning*

Reduplication-like

intensive	A compound with more intensive meaning than its stems
continuative	A compound with semantics of continuity
delimitative / singularizing	A compound with semantics which pick out a more specific referent or context than its stems
connected	A compound which involves the connection of two objects or beings
repetitive	A compound which involves a repetition of any sort
plural	A compound with pluralizing semantics
generic	A compound with a more generic reference or scope
reciprocal	A compound which involves reciprocal motion or reference

Other categories:

attributivizing	Compounding derives an attributive verb
causativizing	Compounding derives a transitive verb
adverbializing	Compounding derives an adverb
nominalizing	Compounding derives a noun
identical	Compounding does not change the semantics of the constituents in any way
idiosyncratic	Compounding derives idiosyncratic semantics from the stems

The sheer number of categories that I include in the tables above reflects the difficulty in analyzing the semantics of alliterative compounds. Rather than placing more compounds in the category "other," I chose to have a more detailed classification. Aside from (17) and (18) each compound was labelled for its lexical category: predicate, noun. Out of 202 alliterative compounds in Mon, 80 are nouns and 122 are either stative or non-stative predicates. The predicates have more complex semantics than nouns. Since nouns are much easier to analyze, I will begin with them.

3.31   Nominal Alliterative Compounds in Mon

Most of the nominal alliterative compounds in Mon fall within three whole compound semantic fields: plant or animal names (26), objects (37), and entities (8). The remaining nominal compounds are in the following categories: other (6), stative predicate (1), motion (1), and

oscillation (1).[15] It is relatively common to find plant or animal names in Southeast Asia which are formed by alliterative or reduplicative patterns. This often occurs for phonaesthetic or visually imitative reasons; the presence of phonological repetition mirrors the sounds or motion made by animals. As for plants, their location or visible characteristics have something to do with the resulting alliteration. These compounds tend to be composed of two bound stems which are unattested outside of the particular alliterative compound. The plant or animal compounds which are not comprised of compound-specific bound stems include some type of plant or animal prefix which functions like a semantic formative[16]. I show examples in (19) of both patterns.

(19)	*Alliterative Compound*	*Gloss*	*Composition (Stem 1 // Stem 2)*
	a. dɒp-de	'butterfly'	(O // O)
	b. kɛŋ-ki	'centipede'	(O // O)
	c. caiŋ-coiŋ	'turkey'	('fowl' // O)
	d. klɜ-kloiŋ	'wolf'	('dog' // O)

Both (19a) and (19b) are animal names consisting of two bound stems which are unattested outside of these compounds; they are idiosyncratic. By contrast, both (19c) and (19d) consist of one stem relating to the animal type and one bound stem which alliterates. The process of compounding here picks out a specific type of plant or animal; it is delimitative.

The nominal alliterative compounds which refer to objects are the largest group. Many of these (17/37) are coordinate compounds where both stems are freely attested in the language. This contrasts with the pattern in the plant or animal names where the compounds are idiosyncratic, usually consisting of two bound stems. I list two examples in (20):

(20)	*Alliterative Compound*	*Gloss*	*Composition (Stem 1 // Stem 2)*
	a. həmɜ̀ŋ-həmàt	'bellows'	('bellows' // 'metal or smith worker')
	b. kɛ̀a-kɔʔ	'cangue'	('cangue' // 'neck, throat')

Alliterative compounding in objects does not usually involve a change in semantics from the stem meanings; fitting into the category 'identical.' For instance, (20a) has a gloss identical to one of its stems. The remainder of these compounds are either idiosyncratic or repetitive in their semantics. I show some examples in (21):

(21)	*Alliterative Compound*	*Gloss*	*Composition (Stem 1 // Stem 2)*
	a. ròk-rɛ̀ak	'thin, crisp kind of cake'	(O // O)
	b. cɔk-caŋ	'shuttlecock'	(O // O)

---

[15] The numbers in parentheses represent the frequency of the compounds in the category.
[16] Such 'formatives' are words which determine a semantic field. For example, the Thai word /plaa-/ 'fish' is a free root but always prefixed onto different fish species' names. (Matisoff, p.c.) Khmer also has a large number of such formatives. For example, the word /neək/ 'person' is prefixed onto many other roots; /neək-miən/ 'rich/wealthy' (lit. 'person' + 'to have').

    c. plɔŋ-plop   'bicycle pump'   ('tube, pipe' // 'to insert, introduce')

While (21a) is idiosyncratic, both (21b) and (21c) involve repetition. A 'shuttlecock' travels over a net in badminton repeatedly (if one is good enough) while a bicycle pump requires repetitive movement to fill a tire with air. These compounds demonstrate that alliteration in Mon may result in reduplicative-like semantics involving repetition.

        Nominal alliterative compounds may also refer to humans. In some cases, the first stem refers to the category of the individual, while the second specifies their particular area of work. For example, in the compound [həma-həmòiŋ] 'tailor,' the first stem is a morpheme that usually occurs in occupational nouns while the second means 'seam or hem.' In the compound [kəmɛŋ-kəmoa], the first stem means 'guard' while the second is related to the stem 'to watch.' This pattern is similar in nature to the general pattern of noun classifiers used in Mon-Khmer languages, all of which precede the word that specifies the compound's meaning. For instance, in Khmer, the compound [neək leeŋ-lbaeŋ] 'gambler' contains the person classifier [neak] along with the compound 'to play a game.' It is not uncommon to find occupations in reduplicative-like constructions. Any person in an occupation repeats an action over and over as a part of their job (Garrett, 2001). Nominal alliterative compounds in Mon have repetitive semantics in both these entities and the objects shown in (21).

3.32   *Verbal Alliterative Compounds in Mon*

        I previously mentioned that there are more predicative alliterative compounds in Mon than nominal alliterative compounds. This may have something to do with the nature of these compounds; alliteration in many languages tends to be used to describe movement, states, or onomatopoeia. The description of such things tends to be verbal or predicative in languages of the world. Predicative alliterative compounds have a broad semantic range but tend to be attributive verbs (80/122). Of the combined semantics category are the following: stative (57), motion verb (39), sound mimetic (12), motion mimetic (3), shape/texture (5), oscillatory (2), and other (4).

        The category stative predicate is the most numerous and includes compounds of an evaluative nature. These compounds refer to exaggerated physical or internal characteristics. For example, the word [kak-kun] means 'to be humpbacked.' The quality of being 'humpbacked' is an exaggerated physical characteristic. Many (28/57) of these compounds fall in the category 'identical' where the stem meanings are identical to the compound meaning. Others have to do with states involving connectedness (8/57), intensity (6/57), or delimitation (4/57). The remainder of the stative compounds have derivational or idiosyncratic semantics. I show examples of each type in (22)

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(22)	a. kəsɛk-kəsa	'to be in continued	(O // O)[17]	stative;

---

[17] Shorto mentions that the second stem here is related to the form /pəsa/ 'illness,' but has undergone phonological mutation. I found a handful of compounds whose stems underwent such

	pain'		idiosyncratic
b. meaŋ-mɔh	'to be remarkable or amazing'	('to wonder, be amazed // O)	stative; attributivizing
c. coa-cɔp	'to have affection for or cleave to'	(O // 'to touch or adjoin')	stative; connected
d. sàŋ-sùi	'to be deliberate in movement; to be ponderous'	('to be heavy or weighty' // O)	stative; delimitative
e. tòiŋ-tὲʔ	'to be dazed'	('to be dazed' // 'to be numb')	stative; identical
f. kətao-kətan	'to be overheated'	('to be hot' // O)	stative; intensive

The compound in (22a) is idiosyncratic (because compounding cannot derive anything from the unattested stems), but involves a sense of 'continuity,' which is often found in reduplicative words. The compound in (22b) is an attributive verb. In this example, a transitive verb stem is de-transitivized to create an attributive compound. Since alliterative compounds tend to be attributive, it is not surprising that compounding results in a change in valency here. The compound in (22c) has semantics that have to do with 'connection.' This semantic type resembles reduplicative semantics because 'connecting' involves the concatenation of two objects or many objects together; it is repetitive. The compound in (22d) fits the category 'delimitative' because the compound meaning is more specific than the stem meaning 'to be heavy or weighty.' Whereas the stem may be used to refer to anything that is 'heavy,' the compound refers only to 'heavy movement.' This semantic category is common in reduplicative constructions (Vuori, 2000).[18] The compound in (22f) has an intensive meaning where the stem 'to be hot' becomes 'to be overheated.' This semantic category is common in reduplication and is found in the verbal alliterative compounds in Mon. While half of the stative compounds do not involve reduplicative-like semantics, the other half do.

The second largest category of verbal alliterative compounds involves motion. Many verbs of motion here involve a sense of continuity, repetition, connection, or delimitation. These semantic categories fit with reduplicative-type semantics. In addition to these categories, there are also a handful of verbal motion compounds which fit the category 'identical.' I outline these types in (23):

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(23)	a. toiŋ-tɛk	'to pluck a stringed instrument'	('to pluck an instrument' // 'to strike or hit')	motion; repetitive

---

mutation in the creation of alliterative vowel or rime changing compounds. The stem here does not appear to be productively used in the language as a free morpheme.

[18] In fact, the English whole word reduplicative construction has these semantics. Observe the following sentences "You're not reading a book. You're reading a magazine! You should read a *book-book*.' In this context the speaker repeats the word 'book' in order to specify a more prototypical instance of the category 'book.' Hence, the word 'book' is delimited.

b. pao-pàk	'to wrap around'	('to bind a cloth or cord around' // 'to put round')	motion; connected
c. lɜ̀-l̃in	'to roll about'	('to roll' // O)	motion; continuative
d. kàŋ-ket	'to take or pick out'	('to grasp or take' // 'to take or obtain')	motion; delimitative
e. krat-krao	'to wash or swill out'	('to wash by pouring water over' // 'to rinse, wash')	motion; identical

The compound in (23a) refers to the motion of 'plucking' in reference to making music. This action involves repetition. The compound in (23b) involves the connection of some object o another object. The compound in (23c) means 'to roll about.' The act of 'rolling about' conveys the sense that the action is continuous and exaggerated. The compound in (23d) is more specific than the meaning of its stems, making it delimitative. The compound in (23e) has identical semantics as its stems. However, the action of 'washing' involves repetitive movement. While compounding does not alter the meaning of the stems in this example, the repetitive action is exemplified by their concatenation. This is an interesting observation; compounding occurs with these stems for a purely iconic reason. Since the verb involves repetition, it occurs in a reduplicative-like process.

In general, alliteration signals a semantic category that will depend on the semantics of the verb. When two verbal stems alliterate, the semantics are "to do more of the action denoted by the verb." If the action is short and abrupt (telic), like 'striking' in (23a), that action can be repeated numerous times (Garrett, 2001). If the action is more continuous (atelic), then it can be extended or intensified. If we look at the compound [klài-klàk] 'to inquire or research into,' we see that the first stem in this compound means 'to look for, search for, or seek,' while the second stem is bound. The action of 'searching' is atelic. Therefore, repetition signals increased intensity. The action of "research" is certainly more intense than the act of "searching"! Verbal alliterative compounds of motion in Mon involve semantics which match reduplicative semantics in other languages.

The remainder of the alliterative verbal compounds in Mon involve mimesis of some sort. Mimesis is the imitation of nature. It may include any aspect of nature: sound, shape, or motion. The categories of sound mimesis, shape/texture, and motion mimesis reflect the three ways in which alliteration imitates natural phenomena. I give examples of each type in (24):

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(24) a.	kəsɔn-kəsɔ	'to keen'	('to mourn or lament' // 'to murmur')	sound mimetic; identical
b.	krɛ̀ʔ-krùt-krɛ̀ʔ-krùt	'the movement of a turtle'	(O // O)	motion mimetic; idiosyncratic
c.	hətot-həton	'to be wrinkled'	('to be wrinkled' // O)	shape/texture; identical
d.	krùʔkrùʔ-krɛ̀akkrɛ̀ak	'the noise of a	('the sound of rain' //	sound mimetic;

            crowd'   'to be constricted')   plural

The category 'sound mimetic' includes both (24a) and (24d), the first describing the manner of lamentation, and the second of which is onomatopoeic. These semantics are often found in reduplicative constructions in languages of the world. For instance, in English we may either use the word 'jibber-jabber' as an onomatope or as a verb of speaking, i.e. 'He just jibber-jabbers all day long.' The category 'motion mimetic' is exemplified in (24b) where the manner of motion of a turtle is described using alliteration. While there are not many of these types of compounds in the Mon dictionary I used (Shorto, 1962), Mon-Khmer languages have an abundance of such words that refer to manner of motion using alliteration. In English, words like 'zig-zag' and 'flip-flop' exemplify this pattern. The verbal compound in (24c) involves texture. This category is probably better seen as a subset of the stative predicates which I described above. These categories represent mimetic characteristics that are often found in reduplicative-type semantics.

      Verbal alliterative compounding in Mon results in either no change of the stem meaning (identical) or reduplicative-type semantics involving repetition, continuity, delimitation, intensity, or mimesis. Out of 122 predicate alliterative compounds, 49 fit the category 'identical' and 60 fit reduplicative-type semantic categories. The category 'identical' is composed of mostly coordinate compounds. The fact that Mon alliteration comprises both semantically coordinate and reduplicative-like compounds is evidence in favor of treating alliteration as compounding here and exemplifies how the data behaves semantically like reduplication. It is difficult to put a single semantic label on such a phenomenon, but there is an interesting parallel between the semantics described in this section and the morphological types in section 3.2. Coordinate compounds are usually of the pattern F-F. Mon prefers to make a coordinate compound out of two free morphemes; while the combination of a bound and free morpheme or two bound morphemes will more likely result in reduplicative-like semantics.

IV.   <u>Khmer:</u>

<u>4.1   Phonology</u>
<u>4.11   General Patterns</u>

      In this section, I will examine the phonological patterns in Khmer alliterative compounds. These compounds occur with a variety of stem types and show an interesting vowel and consonant pattern with respect to ablaut. Khmer has a greater number of alliterative compounds than Mon. A quick glance at Headley's (1977) extensive (60,000 word) dictionary is not possible without seeing at least a handful of these compounds on every page. As I mentioned in section 3.1, whole rime-alternating alliteration is more common in Mon and Khmer than vowel-alternating alliteration, or 'chiming,' as Jacob calls it (1979). While vowel-alternating alliteration occurs somewhat marginally in Mon, it is more frequent in Khmer. I collected 1184 Khmer compounds, 423 of which were vowel-alternating and 779 of which were whole rime-alternating. Additionally, I collected 75 words which show a historical pattern of reduplication in Khmer. I will describe all the patterns of reduplication and alliteration here and go into more detail on the alliterative types.

      As described by Jacob (1979), there are five types of "reduplication" in Khmer, four of which are lexicalized. The first and only productive process is full word reduplication, which conveys either a plural or generalized meaning. I list some examples in (25):

	Word	Gloss		Reduplication	Gloss of Reduplication
(25)	a. tooc	'small'	>>	tooc-tooc	'small and numerous; very small'
	b. srəy	'girl, woman'	>>	srəy-srəy	'women (in general; plural)'

In (25a), the reduplicated adjective has a more intensified meaning than the base. In (25b), the reduplicated noun has a more generalized or plural meaning. The second process that Jacob describes is a rhyming pattern where both the reduplicant and the base are identical except for their initial consonant; the "rhyming" pattern. This pattern is relatively rare (ibid) when compared with the alliterative and ablaut patterns shown in (1) and (2). I show examples of this process in (26):

	Reduplicated Word	Gloss	Composition
(26)	a. mɛɛn-tɛɛn	'really'	('true' // O)
	b. sruuəl-buuəl	'with ease'	('comfortable, easy' // O)

In both examples in (26), the stems have alternating onsets; they rime. The stems of these compounds do not always occur as free morphemes however. In both cases above, the first stem has an independent meaning while the second appears to have resulted from a historically productive process of rhyming reduplication or compounding. There are not nearly as many of these types of reduplication in Khmer as those which alternate the stem vowel.

The third type of reduplication in Khmer is a historically productive process where the initial consonant was prefixed with regular vowel epenthesis, resulting in a sesquisyllable. This pattern is shown in (27) (data taken from Headley, 1977):

Sesquisyllabizing Reduplication

(27)	a. tɔ-tuuc	' to do something stubbornly when told not to; to insist'	>>	from /tooc ~ tuuc/ 'small'[19]
	b. ŋɔ-ŋɔək	'noddingly, in a nodding manner'	>>	from /ŋɔək/ 'to nod, to bob'
	c. cɑ-cak	'to provoke, irritate, or excite'	>>	from /cak/ 'to stab, to pierce'
	d. kɑ-kaeh	'to scratch often, to insist often on something (as a ruse)'	>>	from /kaeh/ ' to scratch once or lift with a finger or fingernail'

In (27) we notice that the base of the reduplicative form has its onset reduplicated. The vowel that is epenthesized here is either /ɑ/ for words whose base is in the first register or /ɔ/ for words whose base in the second register.[20] The vowel type is predictable from the base. The reduplicant of these words is often pronounced in a very reduced manner, making them into sesquisyllables

---

[19] This stem is mentioned to be related to the reduplicated form by Headley (1977). I find this questionable however.

[20] With a few exceptions, the register complex in Khmer is a word-level phonological property. It has been for the most part lost in standard Khmer, but still exists in some dialects, most notably Chanthaburi Khmer (Wayland and Jongman, 2003).

(section 2.1). This prosodic pattern of reduplicated words is distinct from the alliterative compounds that I have been talking about. In Section 2.2, I mentioned that alliterative compounds are never pronounced as sesquisyllables, but as a sequence of two major syllables or two sesquisyllables (for 4 syllable compounds). These cases of sesquisyllabizing reduplication are prosodically distinct from alliteration.

       All of these words are lexicalized in Khmer. While some of the bases seem to correspond in meaning to the reduplicated form, i.e. with a related or more "specific meaning" (Jacob, 1979), many bases are completely unrelated to their reduplicated form. In (27a), the stem /tooc/ 'small' is unrelated to the reduplicated meaning "to do something stubbornly when told not to; to insist." As a result of the lexicalization that has occurred for these words, there is a disjunct between the stem meaning and the reduplicated meaning. I collected 75 of these words in Khmer to compare with the alliterative rime and vowel alternating types. Since these reduplicative forms in (27) are phonologically regular, I will not say more about their structure than what I have just mentioned.

       The vowel and rime alternating types of "reduplication" are the final two patterns that Jacob mentions and the main focus of our attention in Khmer. Similarly to Mon, most alliterative compounds in Khmer have the same register across stems. The voice register distinction in Khmer synchronically involves a vowel quality difference, with higher vowels occurring in the second, "breathy" register and lower vowels occurring in the first, "clear" register.[21] However, a voice quality distinction is still heard in some speakers' voices (Huffman, 1967; Jacob, 1968). Since the stems in each alliterative compound have the same register of which voice quality is no longer the main component, we can conclude that these types of compounds have existed in Khmer since at least the Middle Period (1400 – 1600 C.E.) when a difference in voice quality and vowel quality were both consistently produced between registers (Schiller, 1999). The following data set shows cases of alliteration which change register within the compound. I have included a grave (`) accent to mark the stems in the second register here.

	Alliterative compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(28)	a. kməy-kmiət	'to try very hard to do something'	('hastily, quickly, promptly' // O)
	(or) kmiət-kməy		
	b. ciəm-coàm	'to cause to be tired of; frequently'	('to moderate, temper' // 'saturated, sopping, or wet')
	c. taam-toàn	'to catch up with'	('to follow or agree' // 'to catch up with, be on time for')
	d. toàl-tae	'until, up to the time that'	('to run out of; exhaust' // 'only; however')

In (28a), we notice that both stem orders are possible. While the ordering of stems is usually fixed in Khmer, there are a few like this one which vary. I found only 32/1184 (2.7%) cases where the order of stems in a compound varied, but the remainder 1170/1184 (97.3%) had fixed

---

[21] The historical development of register is a complicated one where voiced initials conditioned a change in voice quality on the vowel. These initial voiced stops were then devoiced. The voice quality distinction eventually led to a difference in vowel quality. For an excellent overview of this process in Khmer and other Mon-Khmer languages, see Ferlus (1979).

stem ordering. The order of the registers in this limited set of compounds is also not fixed. In (28b) and (28c), we notice that the second register stem occurs as the final stem in the compound. In (28d), the stem in the second register is the initial stem of the compound.

With a few exceptions like (28a), the alliterative compounds in Khmer which cross register are composed of two morphologically free stems or have been borrowed as whole units from Sanskrit or Pali. For example, the compound /toàn-tek/ means 'dental,' but has been borrowed in its entirety from the Pali word /danta/, meaning 'tooth.' This pattern of stems which change register in compounds is similar to what we found in Mon where compounds which cross register are always composed of two free morphemes.

With respect to syllable type, we find monosyllables, sesquisyllables, and disyllables in alliterative compounds in Khmer. Out of 1184 alliterative compounds, I found 412 or 34% to be composed of stems larger than monosyllables, while the majority (790/1184), 66% are composed of monosyllabic stems. Most of these larger stems (349/412) are sesquisyllables, while the rest of stems are disyllabic (63/412). I show examples of each type in (29):

	Allit. Compound	Gloss	Composition	Stem Syllable Type
(29)	a. caa-caeŋ	'to explain in detail'	('to engrave' // 'to tell, inform, or clarify')	monosyllabic
	b. ŋɔɲii-ŋɔŋəə	'stupid, not mentally alert'	(O // O)	sesquisyllabic
	c. bantəc-bantuəc	'very little; in a small quantity'	('small, little, few' // O)	sesquisyllabic
	d. craŋap-craŋəl	'in a rolling or tumbling manner'	(O // O)	disyllabic[22]

In (29a), two monosyllabic stems are adjoined to create an alliterative compound. In (29b), we notice the pattern of sesquisyllabic repetition, where two reduplicative words alliterate with only vowel alternation. Similar to Mon, alliterative compounds with stems made up of sesquisyllables and disyllables comprise a smaller portion of the data than those composed of monosyllables.

Many (80/348) of the sesquisyllabic alliterative forms consist of two reduplicated stems. In many cases these compounds come from monosyllabic alliterative compounds which were reduplicated, while in other cases, no un-reduplicated form is found. For example, in (29b), the stem morphemes come from the compound /ɲii-ŋəə/ 'nodding from side to side (of the head).' Each syllable was reduplicated in this compound to create a novel compound with slightly different, but related semantics. However, there is no morpheme /ɲii/ or /ŋəə/ in Khmer. It is notable that in many cases reduplication has applied to a compound whose stems do not occur freely. Reduplication would conceivably yield the form */ŋɔɲii-ŋəə/, a legitimate compound type in Khmer[23], but this compound does not alliterate or "chime" in the same way as the other

---

[22] See section 2.2.

[23] For example, /tɔtiŋ-tɨh/ 'not traditional, not according to one's goals,' from /tɔtiŋ/ 'to disagree or refuse to go along with,' and /tɨh/ 'direction;' or /saʔtɨh-sap/ 'homonym' from /saʔtɨh/ 'homologous, similar' and /sap/ 'sound, noise, voice, word, language.' Such compounds may have

compounds do. It is precisely this salient "chiming" pattern in alliteration in Khmer that makes such compounds phonologically unique with respect to other compounds. In (29c) we see a sesquisyllabic compound which is not created by the adjoining of two reduplicative stems. The final example, (29d), shows a disyllabic stem compound.[24]

   The alternating vowels in Khmer compounds are always on the second syllable in disyllabic stems or on the major syllable in sesquisyllabic stems. This fits the stress pattern in Khmer which is the iambic word and foot. As in Mon, Khmer alliterative compounds vary in the vowel and rime that alternate between stems. I show some compounds exemplifying various vowel and rime alternations in (30) and (31):

"Chiming" allit. Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(30) a. kɑntreɛk-kɑntraak	'very ragged; tattered'	(O // 'torn apart, ragged')
b. krɑɲeɛŋ-krɑɲaaŋ	'angry, bad-tempered'	(O // 'disorderly, separate')
c. crɑkiiŋ-crɑkiəŋ	'unbecoming, bad--mannered'	(O // 'disorderly, old-looking')
d. thiiŋ-thooŋ	'dizzy'	(O // O)

Rime-altered allit compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(31) a. kɑɲceah-kɑɲciəy	'to neglect, waste, be disorganized'	(O // O)
b. kvat-kvaeŋ	'intertwined, meandering, crossing'	(O // 'to criss-cross, to be intertwined')
c. cvat-cvaeŋ	'in circles, circling around'	(O // 'left (direction)')
d. cvat-cviəl	'in circles, circling'	(O // 'around in circles')

In (30c) and (30d) we notice that the stem vowel /ii/ alternates unpredictably with either /iə/ or /oo/. Similarly, in (31c) and (31d) the rime /at/ corresponds unpredictably to either /aeŋ/ or /iəl/. I will demonstrate, however, that the data is not so irregular if we look at the statistical likelihood of certain patterns.

4.12 *Vowel Correspondences*

   The data in (30) and (31) make us think that alliteration is a completely irregular and unpredictable process in Khmer. But despite this irregularity, there do appear to be some recurrent patterns of vowel or rime alternation. In (30a) and (30b), the same initial stem vowel /eɛ/ alternates with the same final stem vowel /aa/. Similarly, in (31b) and (31c), the same rime

---

accidental alliteration, but they are rare and do not alternate rimes the same way as the the rime-alternating alliteratives do.

[24] Henderson (1952) defines "minor disyllables" (i.e. sesquisyllables) as those which never occur with complex onsets and show vowel reduction, while "major disyllables" have complex onsets and no vowel reduction.

alternations occur. The Khmer data have more clear correspondences in the way in which stems alliterate in compounds than the Mon data. There are other compounds that show this same vowel alternation in Khmer. Some are shown in (32):

	Alliterative compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(32)	a. pɑpak-pɑpaek	'tremblingly, shakingly'	(O // O)
	b. caa-caeŋ	'to explain in detail'	('to engrave'// 'to tell, inform')
	c. knah-knaeŋ	'to try hard'	(O // O)

The pattern between /a/ and /ae/ above does not hold for all alliterative compounds which contain /a/ in the first stem. For instance, the compound /tlah-tlooy/ 'to fail or go wrong because of negligence,' alternates the vowel /a/ with /oo/. However, when patterns like this occurs repeatedly, one is tempted to put them together to see how things correspond. Khmer has 20 monophthongal vowels and 12 diphthongal vowels, all of which are phonemic (Henderson, 1952). Since there are so many vowels, there are many possible vowel combinations that can occur. Nevertheless, there are many cases where one vowel will predominantly alternate with another specific vowel in Khmer. The table in (33) lists some common correspondences in the data.

		Initial Stem Vowel		*Final Stem Vowel*	*Frequency (N=1,184)*
(33)	a.)	ae	-->	ɑ, ɑɑ, ao	57
	b.)	ə	-->	o, oo	49
	c.)	eɛ	-->	aa	40
	d.)	ii	-->	iə	35
	e.)	a	-->	aə	11
	f.)	ee	-->	oo	9
	g.)	ɨŋ	-->	eaŋ	9

One thing that we notice from the chart in (33) is that low vowels in initial stems tend to stay low on the final stem or diphthongize, as in (33a) and (33e). Another interesting characteristic of these correspondences is that initial stem mid vowels alternate with final stem back vowels. This is the case in (33c) and (33f). Finally, in (33d) we see that the initial stem high front vowel /ii/ tends to diphthongize in the final stem. Taken as a whole, there are tendencies for the final stem vowel of an alliterative compound to be lower, longer, and diphthongized. Some examples are shown in (34).

	Alliterative compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(34)	a. kɑkeɛp-kɑkaap	'feverishly'	(O // O)
	b. krɑɲeɛŋ-krɑɲaaŋ	'angry, bad-tempered'	(O // 'disorderly, separate')
	c. cɑcreɛh-cɑcraah	'helter-skelter'	(O // 'to reverse one's direction; to move in reverse')

d. tɔtiim-tɔtiəm	'slowly, hesitantly, and indecisively'	(O // O)
e. baŋkii-baŋkiə	'torpid, sluggish, slow'	(O // O)
f. kcii-kciə	'lazy, careless, negligent'	(O // O)

The examples in (34a) – (34c) show the alternation between /eɛ/ and /aa/. It is interesting that these words have similar semantics. The compound in (34c) refers to a "mad" state, similar to the meaning of (34a) 'feverishly' and (34b) 'angry.' The examples in (34d) – (34f) show the alternation between /ii/ and /iə/. These compounds all have something to do with 'slowness' or 'laziness' but have no phonological similarity outside of the fact that they share the same "chiming" vowel alternation. It is often the case that one finds a particular rime or vowel alternation in Khmer to correspond to a particular meaning.

These examples are evidence of what is known as a phonaesthetic process. Onsets or rimes which happen to share some semantics tend to be associated with a particular meaning over time (or are etymologically related). In English, for example, words like 'glitter,' 'glisten,' 'gleam,' and 'glow' all pertain to 'light' or 'vision.' Similarly, words which end in the rime /æʃ/ tend to involve hitting, e.g. 'bash,' 'crash,' 'smash,' 'lash,' and 'mash.' We find similar types of phonaesthemes in non-repetitive words in Khmer as well. For instance, the words /krəluəc/ 'high-pitched,' /sruəc/ 'pointed,' and /huəc/ 'to whistle,' all have a similar meaning and share the same rime (Jacob, 1992). It is not surprising therefore that phonaesthesia is found in alliterative compounds. In fact, the sheer number of these types of expressions in Khmer may increase the likelihood that they be phonaesthetic (Bergen, 2004). These compounds have a prosodic similarity which is salient to the speaker of Khmer and share the same range of meanings associated with manner of motion, qualitative state, repetition, etc. Vowel patterns show correspondences in alliterative forms from which speakers make a phonaesthetic connection.

In the present study, I investigated this tendency for vowels and rimes to alternate regularly in Khmer by looking at the number of vowels and their varying qualities across the stems in alliterative compounds. This was inspired by the work of Eric Schiller (1999), who compared alternating vowels in chiming "reduplicants" for Middle Khmer. He found that there is a tendency for high, front vowels to appear on the initial stems in such compounds while lower, and back vowels tend to appear in final stems. While I replicate Schiller's study with more data, I also have included whole-rime alliterating compounds where he only considered "chiming" or vowel-alternating ones. The chart in (35) shows vowel quality in both the initial stem and the final stem of alliterative compounds. It includes only the monophthongal vowel types. I discuss diphthongs later.

(35)   *Stem and Vowel Quality in Khmer Alliterative Compounds* (N = 861 stem1; 764 stem2)[25]

---

[25] I excluded 33 compounds without fixed stem order, e.g. (28a) /kməy-kmiət ~ kmiət-kməy/ 'to try very hard to do something,' as well as 308 examples where initial stems consisted of diphthongs and 405 cases where final stems consisted of diphthongs. I discuss diphthongs in (37).

Front Vowels	Frequency Stem1/Stem2	Central Vowels	Frequency Stem1/Stem2	Back Vowels	Frequency Stem1/Stem2
i, ii	91 / 9	ɨ, ɨɨ	63 / 28	u, uu	58 / 35
e, ee	61 / 21			o, oo	97 / 124
eɛ	69 / 6	ə, əə	182 / 86		
ɛɛ	6 / 21			ɔ, ɔɔ	8 / 33
a, aa	127 / 255			ɑ, ɑɑ	99 / 145

There are two patterns in the data in (35). First, all front, non-low vowels occur more commonly in the initial stem when compared to the final stem. These vowels are above the darkened line in the table. I counted 524 cases of /i, ii, e, ee, eɛ, ɨ, ɨɨ, ə, əə, u, uu/ on initial stems but only 185 on final stems. As the null hypothesis, we expect these vowels to appear with equal likelihood on either stem. However, higher and more front vowels show a significant tendency to occur on the initial stems of alliterative compounds. The result of a $\chi^2(1)$ is 160.1, ($p < 0.00001$). The highest and most front, /i/, almost never occurs in final stems in alliterative compounds. The second pattern is for non-low back vowels to occur with greater likelihood on the final stem of a compound. These vowels appear below the bold line in the table. I counted 337 cases of lower and further back vowels on the first stem (/ɛɛ, a, aa, ɔ, ɔɔ, ɑ, ɑɑ, o, oo/), but 578 on the final stem. There is a significant tendency for vowels which are lower and further back to occur on the second stem of alliterative compounds. The result of a $\chi^2(1)$ test is 63.5, ($p < 0.00001$). In sum, initial stem vowels tend to be non-back and non-low in vowel quality while final stem vowels tend to be back and low in vowel quality. These findings are significant and agree with Schiller's claim. What exactly does a vowel pattern like this suggest? The presence of lower and back vowels occurs in the compound-final position, a prosodically-strong domain in Khmer. It is the prosodic structure of the alliterative compound which determines its pattern of stem combination.

While the patterns in (35) relate to vowel quality, I also looked at vowel duration. I noticed that short vowels appear on initial stems with a higher frequency than long vowels or diphthongs. Conversely, long vowels and diphthongs occur on final stems with greater frequency than short vowels do. This data is organized in (36):

(36)   *Stem and Vowel Length in Khmer Alliterative Compounds* (N = 1,170)

	Initial Stem	*Proportion*	*Final Stem*	*Proportion*
Short Vowels:	495	42.3%	341	29.1%
Long Vowels:	366	31.3%	423	36.2%
Diphthongs:	309	26.4%	406	34.7%

Initial stems show a significant tendency to contain short vowels in their rimes. The result of a $\chi^2(2)$ test is 46.6, ($p < 0.00001$). There is also a tendency for final stems to have longer and more diphthongized vowels, although this tendency is weaker than the initial stem pattern; $\chi^2(2) =$

9.61, (p < 0.01). Between the two stems, there is a greater tendency for shorter vowels to occur in the initial stem than in the final stem; $\chi^2(1)$ test is 28.4, (p < 0.00001). The initial stems of alliterative compounds typically have shorter, fronter, and higher vowels while the final stems typically have longer, diphthongized, lower, and backer vowels. Stems which have longer or diphthongized vowels tend to pattern initially in alliterative compounds. This is another example of how the prosody of word-formation determines the structure of the compounds. Longer vowels are prosodically strong, having a longer duration than short vowels.

       Similarly to Mon, Khmer has a large inventory of diphthongs: /iə/, /ɨə/, /uə/, /ae/, /ao/, /aə/, /ea/, /oa/, and /ɔə/. In (36) we find that a larger proportion of diphthongs occur on the final stems of alliterative compounds than on the initial stems. I provide the specific frequencies of different diphthong types in (37).

(37)   *Stem and Vowel Diphthong Type in Khmer Alliterative Compounds* (N = 1,169)

Initial Stem Vowels		Final Stem Vowels	
Diphthong	**Frequency**	**Diphthong**	**Frequency**
backing ao	9	backing ao	49
raising ae	91	raising ae	41
lowering oa, ea	46	lowering oa, ea	79
centralizing iə, aə, ɨə, **ɔə**, uə	162	centralizing iə, aə, ɨə, **ɔə**, uə	237

The table in (37) shows the diphthongs organized by their direction of movement (or offset vowel). Diphthongs which back, lower, or centralize a vowel occur with greater likelihood on the final stems of alliterative compounds in Khmer than on the initial stems; $\chi^2(1) = 37.6$, (p < 0.00001). This observation correlates with the pattern shown in (35) where the final stem prefers back and low vowels.

       Two diphthongs, /ae/ and /ɔə/, occur with greater frequency on the initial stem of an alliterative compound (91:41, and 47:34, respectively). This appears to be the opposite of the pattern in (35) where diphthongs occur with higher frequency on final stems. However, these diphthongs do not lower or back a vowel like the others. The vowel /ɔə/ centralizes from a back vowel while /ae/ raises. These two diphthongs are not actually exceptions to the tendency in (35), but reflect the preference for back and low vowels in the final stems of alliterative compounds in Khmer.

       While these tendencies are synchronically gradient, they were much more regular in Middle Khmer. Schiller makes this point in his article and reconstructs the vowels in Middle Khmer based on a corpus of about 450 compounds. I have reproduced Schiller's findings in the chart in (38):

       Middle Khmer Vowel Distribution in Alliterative Compounds (Schiller, 1999)
Stem Vowels based on backness   **Stem Vowels based on height**

(38)

	Initial Stem Vowel	Final Stem Vowel		Initial Stem Vowel	Final Stem Vowel
Front	268	17	High	178	70
Central	30	155	Mid	94	71
Back	34	149	Low	60	191

We notice that front vowels occur in a large majority of initial stems in "chiming" compounds in Middle Khmer, but are rare in final stems, where central or back vowels are more common. Similarly, high vowels are more common in initial stems than in final stems in Middle Khmer. While the set of sound changes involved between Middle Khmer and Modern Khmer is quite complex, certain sound changes would have affected this pattern greatly. Breathy phonation, or the second voice register, in Khmer conditioned a change in vowel height between breathy and clear vowels. Breathy vowels were perceived as having higher vowel quality. As a result, the modern reflex of certain historically low vowels is a higher vowel (Ferlus, 1979). For example, the Middle Khmer vowel */a̤/ is /iə/ in Modern Khmer. Vowels such as */ɑ̤/ and */ṳ/ are /ɔ/ and /o/, respectively, in Modern Khmer. This first example is significant; we find a great number of cases (123/1169) where /iə/ occurs as the vowel in the final stem of an alliterative compound. Since this vowel was historically a low vowel, it would support the tendency for low vowels to occur in final stems of alliterative compounds, as Schiller and I suggest.

4.13   Summary and Discussion

Aside from these vowel quality and length distinctions, there are many interesting generalizations regarding the phonology of alliterative compounds in Khmer that I have covered here. First, compounds which change register are composed of two free stems in most cases. Second, alliterative compounds in Khmer are composed of monosyllables, sesquisyllables, and disyllables, in this order of decreasing frequency. Finally, there is an abundance of vowel patterns which commonly occur and converge semantically, forming phonaesthetic alliteratives.

The data from (35) – (37) show there to be tendencies in the vowel patterns in alliterative compounds. Final stems tend to have vowels which are diphthongized, longer, lower, and back. Initial stems tend to have vowels which are monophthongal, shorter, higher, and non-back. These tendencies suggest that there is a relationship between the prosodic pattern of alliteration and the vowels which are found on alternating stems. The rhythmic structure of the Khmer word is iambic. Stressed syllables occur in word-final position. Domain-final positions such as these are the locus of many processes of vowel change, such as final-lengthening and vowel laxing (Barnes, 2002). Barnes gives examples of processes which involve "final-strengthening" in a number of languages. Final position and stress overlap in Khmer words. These stressed positions attract vowels which are more open, further back, and longer (ibid). Final stems in Khmer alliterative compounds correlate with this pattern of final strengthening. However, this tendency is not limited to vowel type, but also includes the whole rimes of such words.

4.14   Rime Correspondences

As I mentioned in (31), the whole rimes which alternate in alliterative compounds do so in an unpredictable fashion. Hence, /kɑɲceah-kɑɲciəy/ 'to neglect, waste, be disorganized' alternates two rimes /eah/ and /iəy/. However, these rimes do not always co-occur. For example, the compound /ʔanteah-ʔantɛɛŋ/ 'to experience great emotional agitation' includes a different rime /ɛɛŋ/. Data like this make us believe that rimes are unpredictable in their patterning in Khmer alliteration. However, the comparison of coda types across the stems in these compounds shows a tendency for final stems to include more sonorants than initial stems. The data on final consonant type is shown in (39).

(39) *Coda Consonant*
      Across Stems:

Final Consonant	Initial Stem	Final Stem
/n/	46	80
/m/	60	63
/ɲ/	23	19
/ŋ/	72	127
/l/	18	45
/ʋ/ [ɰ]	14	30
/y/	22	95
/p/	59	45
/t/	79	50
/c/	36	33
/k/	119	52
/ʔ/	5	3
/h/	90	35
/V-final/	108	74
Total	751	751

Coda Type:

	Sonorant	Obstruent	None (Vowel)
Initial Stem Frequency	255/751	388/751	108/751
Percentage of coda type	34.0%	51.7%	14.4%
Final Stem Frequency	459/751	218/751	74/751
Percentage of coda type	61.1%	29.0%	9.9%

For stems which end in a coda, the null hypothesis is that they will end in either a sonorant or an obstruent, with equal likelihood. However, from the data in (39) we see that obstruent codas occur with greater frequency on the initial stem in an alliterative compounds than sonorant codas do; $\chi^2(1) = 27.5$, (p < 0.00001). Also, sonorant codas occur with a greater frequency on the final stems in alliterative compounds than obstruent codas do; $\chi^2(1) = 85.8$, (p < 0.00001). The prototypical alliterative compound in Khmer alternates an obstruent coda with a

sonorant coda, in that order. Thus, a compound like /cvat-cviəl/ 'in circles, circling' is typical of the data while a compound like /riəŋ-riəp/ 'continuously, always' is relatively uncommon.[26]

In the previous section, I mentioned that the structure of Khmer prosody is related to the pattern of vowel alternation that is found. The data in (39) also supports this perspective. Linguistic stress involves higher pitch, greater vowel length, and increased intensity. The rime is the prosodic unit over which stress or tone are realized in language. Rimes are lengthened as a result of stress. Since Khmer has no word-final voiced obstruents, the only voiced finals are sonorants. These sonorants extend the duration over which pitch may be realized in the rime. Obstruent codas, because they are voiceless, do not. In addition to vowel lengthening and vowel quality differences on the final stem of alliterative compounds, we also find increased rime duration in these stems with sonorant codas. This increased duration of the entire rime in the final stem correlates with the iambic structure of Khmer words; it is a rhythmic property. Additionally, the vocalic and coda patterns act together in a unified way which supports the notion of the rime as a structurally relevant prosodic unit in Khmer.

4.2   *Morphology*
4.21   *General Patterns*

The morphology of alliterative compounds in Khmer is similar to what we find in Mon. First of all, we find the same degree of variation in the order of free and bound stems in Khmer. Compounds may be composed of two free stems, one free and one bound stem, or two bound stems. However, more compounds consist of two bound stems in Khmer than in Mon. Furthermore, Mon compounds have a tendency for final stems to be bound while Khmer compounds do not. The four patterns are exemplified in (40):

	Alliterative Compound	*Order*	*Gloss*	*Composition (Initial-Final)*
(40)	a. papək-papak	B-B	'in a constantly wet and/or dirty manner'	(O // O)
	b. ŋək-ŋak	B-F	'to swing or move the body in order to attract attention'	(O // 'abruptly, brusquely; to stop suddenly')
	c. cae-cəʊ	F-B	to be a matchmaker between bride and groom; to be a slanderer'	('older sister' // O)
	d. bandah-banday	F-F	'to use a pretext to get out of some commitment; to be excused from a deadline'	('to deliver or save someone from something' // 'to let go; to allow')
	e. kan-kap	F-B	'to maintain, control, supervise or operate'	('to hold something, to adhere to a philosophy or policy' // 'to chop, to cut')

---

[26] It is possible that there is also a preference in Khmer for these codas to be homorganic. This is an unresolved issue.

Contrary to Mon (and Vietnamese), there is no preference in Khmer for the F/B pattern shown in (40c). The B/B pattern in (40a) and the F/F pattern in (40d) are common in Khmer. Khmer prefers to combine two bound or two free stems when creating an alliterative compound; $\chi^2(3) = 95.5$, $(p < 0.00001)$. The table in (41) shows these tendencies in the corpus.

(41)

Morphological Pattern (N=1,153)	*Frequency*	*Percentage*
Free+Free	354	**30.7%**
Free+Bound	185	16.0%
Bound+Bound	384	**33.3%**
Bound+Free	230	19.9%

The four possible patterns in (41) can be further divided into three types of semantic composition: exocentric, endocentric, or partially-endocentric. The example in (40a) is exocentric while (40b-d) are endocentric. The example in (40e) is partially-endocentric. The initial stem has related semantics to the compound's meaning, but the final stem 'to chop or cut' is clearly unrelated. While the initial stem is the semantic head, the final stem is a non-head.

There are two possible types of bound stems in exocentric and endocentric compounds. A bound stem may be consist of a freely-attested stem which is a semantic non-head, as in (40e), or an unattested stem which alliterates with the semantic head, as in (40b). This latter type is more common in the Khmer data. The former type may be the result of accidental homophony (as mentioned earlier for Mon). The table in (42) shows these patterns, where A represents an attested stem and U represents an unattested stem.

(42)

Composition (N=799)	*Frequency*	*Percentage*
Bound + Bound		
A+A	33/384	8.6%
A+U	36/384	9.4%
U+A	42/384	10.9%
U+U	273/384	71.1%
Bound + Free		
A+A	54/230	23.5%
U+A	176/230	76.5%
Free + Bound		
A+A	57/185	30.8%
A+U	128/185	69.2%
Total: with unattested bound stem:	**577/799**	**72.2%**
with attested bound stem:	**222/799**	**27.8%**

The table in (42) shows that most alliterative compounds in Khmer with bound stems do not in fact contain those that are attested outside of the compound. Rather, alliterative compounds are composed of bound roots which are unattested outside of the compound; i.e. they were created via a productive process like reduplication. A greater percentage of compounds of the shape B/B, F/B, and B/F contain bound stems which are unattested. This pattern is significant; $\chi^2(1) = 157.7$, $(p < 0.00001)$. The compounds which have attested stems which are semantic non-heads comprise only 28% of this data.

One wonders if the many stems which are labelled as attested semantic non-heads are actually the result of a productive reduplication-like process at one time but *accidentally* became homophonous with attested roots. For example, a word like "pish-posh" in English contains the syllable 'posh,' which has a lexical meaning 'hip, or stylish' as an independent morpheme. Most people would say that these two homophonous syllables are not homo-morphous (composed from the same morpheme). However, there is no a priori reason to reject the hypothesis that homophonous syllables correspond to the same morpheme in Khmer. Many compounds become exocentric due to semantic change; morphemes which were historically semantic heads of a compound become bound to it. Much historical work would be required to resolve this issue. For the purposes of this paper, I assume all stem homophony is homomorphous where an independent morpheme exists.

4.22   *Comments on Khmer Morphology*

Since Khmer has a large set of derivational (but unproductive) prefixes and infixes, many alliterative compounds consist of two derived stems. The functions of these derivational morphemes are sometimes predictable, but often an affix will have a variety of functions which make a derived stem semantically opaque from its base (Huffman, 1970). Haiman and Ourn (2003) suggest that some affixes may have applied to stems for *stylistic* reasons, where a full, sesquisyllabic or disyllabic form signalled a derived stem, and the underived root was the reduced form. Since so many roots in Khmer appear to be related by derivation, many linguists have attempted to define the function of these morphemes (Huffman, 1967, 1970; Jacob 1963, 1968, 1976, 1979; Nacaskul, 1978; Haiman and Ourn, 2003). It is not the goal of this paper to analyze these morphemes generally, but to look at which ones occur with alliterative compounds.

The prefixes in Khmer are composed of the shape /C/, /CV/, /CrV/, or /CVN/ (where the nasal is homorganic to the initial of the major syllable). The /C/ prefixes are /p, t, c, k, s/, where /p/ is the most common, having the function of 'causation.' For example, the word /psaŋ/ 'to tame' is derived from /saŋ/ 'to be tame' (from Huffman, 1970). There are two prefixes with the shape /CV/, a reduplicative prefix and an attributive prefix /rɔ/. The data in (27) exemplify the first pattern. For instance (27d) /kɑ-kaeh/ 'to scratch often, to insist often' comes from the stem /kaeh/ 'to scratch once.' These reduplicative prefixes often carry the meaning of intensification or repetition. The /CrV/ prefixes are prV-, trV-, crV-, krV-, and srV-. The most common of these prefixes is prV-, which has the typical function of reciprocity. Out of the many /CVN/ prefixes, the most common is /bɑN/ which has the function of 'causation.'

The infixes in Khmer are composed of the shape /C/, /Vm/, /VN/, or /rVN/, all of which apply after the initial consonant of the stem. The infixes -C-, -VN-, and -rVN- are all nominalizing while the infix -Vm- is causativizing. All of these affixes may occur in the stems of alliterative compounds, as in (43):

(43)   *Derived Stems in Alliterative Compounds*

Alliterative	Meaning	Affix	Composition
a. pkoap-pkun	'to try to please in order to gain a favor or benefit'	/p/	/pkoap/ 'to satisfy, please' </koap/ 'right, satisfied' /pkun/ (unattested)

b. rɔɲəm-rɔɲam[27]	'messy, messed up, very fast and repetitious (of speech)'	/rɔ/	/ɲəm/ 'to nurse, suckle'
			/rɔɲam/ 'repeatedly, fast'
			< /ɲam/ 'to eat'
c. pranəŋ-pranoŋ	'rough, uneven, disorderly sloppy, improperly'	/prV/	from /nəŋ-noŋ/ 'music on small gongs'
d. bantəc-bantuəc	'very little, in a small quantity'	/baN/	/bantəc/ 'small, little, few'
			< /təc/ 'little, few, small'
			/bantuəc/ (unattested)
			< /tuəc/ (unattested)
e. lbəc-lbaaŋ	'misleading, sly, crafty, tricky, a trick'	/-b-/	/lbəc/ 'trick, trickery'
			< /ləc/ 'to appear from a hidden spot, to leak'
			/lbaaŋ/ 'to test (someone)'
			< /lɔɔŋ/ 'to test'
f. samraap-samruəl	'to arrange, coordinate'	/-Vm-/	/samraap/ 'to cause to agree'
			< /sraap/ 'agreeing'
			/samruəl/ 'to make easy'
			< /sruəl/ 'easy'
g. sraneh-sranaok	'to repent, to regret, to feel sorry for something'	/-rVN-/	/sraneh/ (unattested)
			< /seh/ 'horse'
			/sranaok/ 'sorrow, remorse'
			< /saok/ 'to regret; sorrow'
h. papŋak-papŋaək	'bouncing, moving up and down'	Redup.	/papŋak/ (unattested)
			< /pŋak-pŋaək/ 'balancingly'
			/papŋaək/ 'totteringly'

The table in (43) shows that stems which are morphologically complex can alliterate in Khmer compounds. This is particularly interesting with regard to the two types of bound stems, mentioned in section 4.21. As we recall, a bound stem is a semantic non-head which is either attested or unattested in Khmer. We expect that derivational morphemes apply only to attested roots because the meaning of a derived stem form requires semantic content in the base. However, there are many alliterative compounds which have stems that appear to be morphologically complex, but lack semantic content. For instance, in (43d), the word /bantuəc/ is unattested, but appears to contain the prefix /baN/. In (43a), the word /pkun/ is unattested, but appears to contain the prefix /p-/. In (43c) and (43h), an un-derived alliterative compound is attested, but the derived stems are not. Derivational morphology has applied in alliterative

---

[27] Many compounds relating to speech are also related to the mouth. The compound *mɔət-piət* 'argument, quarrel' contains the word for mouth *mɔət* and the word for 'word' *piət*.

compounds where one or both of the roots are unattested. It even may apply to an entire alliterative compound, as in (43c) and (43h).

These examples show an interesting pattern of derivation. In (43c), affixation applies to both stems of an alliterative compound. In (43d), affixation applies to the bound and unattested root /tuəc/. This root is a "filler" syllable, in the same sense that we mentioned in Mon. In (43g), affixation applies to the bound and attested root, /seh/. In each of these cases, affixation occurs in both the stems of an alliterative compound. The reason for this is clearly not morphological. Rather, the phonological symmetry of these compounds is one of their most salient features which must over-ride the morphological requirements in Khmer, where affixation would normally be restricted to free stems. This symmetry is entirely prosodic, where the stem shape of an alliterative compound consists of either a major syllable, a sesquisyllable, or a full di-syllable. Two stems are required in alliterative compounding to have the same onset or pre-syllable. This structuring can be represented with the same prosodic template that I gave for Mon in section 3.2. I repeat it in (44):

(44)   *Prosodic Template: Alliterative Compounds in Khmer*
         2 sesquisyllables/syllables: $presyllable_1$-($onset_1rime_1$) + $presyllable_1$-($onset_1rime_2$)
         2 syllable: ($onset_1rime_1$) + ($onset_1rime_2$)

A number of generalizations fall out of this prosodic template in Khmer. First of all, bound stems which are unattested are created via a process resembling reduplication to fit this template. Khmer prefers to compound with mostly attested bound stems, as we saw in section 4.21. Stems which are semantic non-heads are selected because they happen to alliterate, according to the template, with the semantic head of the compound. In this way, alliteration happens to function like compounding where an alliterating stem is available. Where one is unavailable, it is created to fit the template. Second of all, this prosodic template for alliteration predicts that we will find derivational morphology applying to unattested roots. If a root does not occur with certain derivational morphology, the template will match the pre-syllable and onset portion of this bound root to alliterate with the semantic head of the compound. It does not matter for the template whether or not a compound has one or more bound roots, as long as they alliterate. Hence, we predict that alliterative compounds will have all types of morphological constituency, consisting of free stems, bound and free stems, and bound stems. Finally, this template corresponds with the iambic structure of Khmer words. Alliterative compounds are composed of a single iambic foot, as in the 2 syllable pattern in (44), or two iambic feet, as in the 2 sesquisyllable/syllable pattern.

This template also sheds light on how other derivational processes work in Khmer. Outside of alliteration, there are many "derived" words in Khmer where affixation applies to an unattested root. Let us take reduplication as an example. I previously mentioned in (27) that there was a historically productive process of reduplication where a root initial was prefixed with an epenthetic /ɑ/ or /ɔ/ vowel. I showed examples where the semantics of the reduplicative word was related to the semantics of the root. For example, the word /cɑ-cak/ 'to provoke, irritate, or excite' comes from /cak/ 'to stab, to pierce.' The meaning of the root is related to the meaning of the reduplicative word; it is semantically transparent. However, this transparency is rare in Khmer reduplication. For instance, out of 75 reduplicated words I collected, only 25 have free roots. There are 50 reduplicated words which are composed of a reduplicant and a bound root.

Similarly to Mon, some of the roots in reduplicative words are attested, but are semantic non-heads. In other cases, no meaning exists for the root of a reduplicative. I show some examples in (45):

		Reduplicative	*Gloss*	*Composition (Stem 1 // Stem 2)*
(45)	a.	cɑ-ceɛh	'to be stubborn or unruly; to refuse to listen to advice'	(O // O; stem = 'thread')
	b.	tɔ-tuu	'to cover one's head (with something)'	(O // O; stem = 'to compare')
	c.	ŋɔ-ŋul	'to be bent down (of the head)'	(O // O)
	d.	pɑ-paoc	'noisily, loudly'	(O // O)

Both (45a) and (45b) have attested roots which have different meanings from the derived form. The word for 'thread' in (45a) is unrelated to 'a refusal to listen to advice.' These roots are quite possibly accidentally homophonous. If this root were related to the reduplicative, then the relationship in meaning has been lost and the reduplicative is completely semantically opaque. These cases contrast with those in (45c) and (45d), where the stem of the reduplicative word is completely unattested in the language. Both reduplication and alliteration involve derivational morphemes that may apply to unattested roots. The reason for this is in alliteration is that the prosodic template conditions the morphological process of compounding. In reduplication, the same may apply. Roots which are unattested in the language are reduplicated to fit a sesquisyllabic pattern.

4.23   Borrowing in Alliterative Compounds

There are a handful of alliterative compounds in Khmer which have been borrowed from Sanskrit, Pali, or other Southeast Asian languages. There are some interesting patterns regarding how the stems are combined in compounds like these. I show some examples in (46):

(46)a. *Allit. Compound*   bɑɑ-boo
      Gloss   'plentiful, abundant'
      Composition   (O // O)
      Derivation   from Skt. /paripūrṇa/ 'full'

   b. *Allit. Compound*   kseɛm-ksaan
      Gloss   'peaceful, pleasant, enjoyable'
      Composition   ('happy and healthy; free from enemies; secure' // 'peaceful, pleasant, enjoyable')
      Derivation   both stems from Skt. compound /kshema-kshānta/; /kshema/
                           'good, peaceful, prosperous, happy,' /kshānta/ Ø (derived?)

   c. *Allit. Compound*   kriəm-krɑh
      Gloss   'dried out; meager (of a meal)'

Composition	('dry, dried' // 'dry')
Derivation	first stem /kriəm/ from Thai /kriam/, second stem is Khmer

d. *Allit. Compound*   ciəm-coam
    Gloss   'to cause to be tired of (food or doing something); frequently'
    Composition ('to moderate, to temper' // 'saturated, sopping wet')
    Derivation   first stem /ciəm/ from Thai /ciam/ 'humble,' second stem is Khmer

e. *Allit. Compound*   chaep-chɑɑp
    Gloss   'to praise, to flatter'
    Composition (O // 'to be compatible, to get along with or love each other')
    Derivation   /chaep/ is unattested, second stem /chɑɑp/ from Thai /chɔɔb/; may come from Thai compound /chɔɔb-cay/ 'to be pleased; happy' or predicate /cay chɔɔb/ 'the heart likes something' (Headley, 1977)

The examples in (46a) and (46b) come from Sanskrit. In the first example, the Sanskrit word /paripūrṇa/ is actually derived from two morphemes; the prefix (upasarga) /pari/ meaning 'around or about' and the stem adjective /pūrṇa/ meaning 'full.' The final syllable of both these morphemes is lost in Khmer. As a result, the syllables are alliterative but have no meaning as free morphemes in the language. In (46b), both Sanskrit stems are attested in Khmer as free morphemes, but are combined together to make an alliterative compound. This is not surprising however; an abundance of words which begin with the onset /ks/ are from Sanskrit. The combination of these two phonologically-related words to create an alliterative compound may have been directly borrowed as a whole from Sanskrit, but I have no data to support this claim. It appears that the compound was created in Khmer. These two examples show us the types of alliterative compounds that we find from Sanskrit in Khmer; they may either consist of two bound stems, such as the case in (46a), or two free stems as in (46b).

The situation for Thai borrowings into Khmer is somewhat different than the Sanskrit examples. While (46e) is potentially an entire compound borrowed from Thai, both (46c) and (46d) contain two stems, one of which is Thai and the other Khmer. While I only have a handful of compounds of this sort, the pattern is always Thai stem first, Khmer stem second. In both (46c) and (46d) the Thai stems are free morphemes in Khmer. While these compounds consist of two stems of different origin, they still consist of either two free or two bound morphemes. When borrowed, alliterative compounds in Khmer are either taken as a whole, or combined via free word compounding.

4.3   *Semantics in Khmer*

Alliterative compounds are "par excellence" attributive (stative) verbs in Khmer. Nominal alliterative compounds do occur, but with less frequency in Khmer than in Mon. Recall that nominal forms make up 40% of the alliterative vocabulary in Mon. Khmer, on the other hand, has 186/1184 nominal forms, which make up only 15.7% of its alliterative vocabulary.

Most alliterative compounds are verbal or polyfunctional[28] predicates. The word 'predicate' is a large grammatical category, but I use it because it is difficult to categorize Khmer alliterative compounds (and Khmer words in general) into smaller categories like verb, adjective, and adverb. For instance, the compound /rɔʔak-rɔʔaək/ means 'nauseated,' 'to feel like vomiting,' or 'in a nauseated manner.' The first gloss is adjectival, the second verbal, and the third, adverbial. There is much overlap between these three categories in the Khmer alliterative compounds. Nominal compounds do not overlap much with these categories, with the exception of a few (16) deverbal nouns.

The semantics of alliterative compounds in Khmer are very similar to what we found in Mon. I utilized the same semantic categories that I show in (17) and (18). The semantic categories which are more "reduplication-like," as in (18), involve repetition, continuity, intensity, delimitation, connectivity, generality, and reciprocity. These categories comprise 444/1184 compounds, or 37.5% of the data. Khmer alliterative compounds tend to have semantics similar to reduplication. Certain compounds, namely mimeticizing and those which add manner to a verb, required additional categories. Mimeticizing compounds occur where two stems, which may or may not be onomatopoeic, combine to create an onomatope. The other categories include derivational semantics like verbalization, zero derivation, or idiosyncratic semantics. These last two semantic categories make up a large percentage of the Khmer data (508/1184; 42.9%). In many cases the semantics of a compound do not vary from the semantics of one of its stems. Compounding results in zero derivation. In other cases the semantics of the compound are very different from the stems. I have classified these compounds as having idiosyncratic semantics.

4.31   *Nominal Alliterative Compounds in Khmer*

Nominal alliteratives in Khmer consist of objects (100), plant and animal names (32), entities (22), and deverbal nouns (32). Objects are the most common type (100/186; 53.8%), most of which are coordinate compounds with transparent semantics. I show some examples in the table in (47).

Object Alliterative Compounds

	Alliterative Compound	*Gloss*	*Category*	*Composition*
(47)	a. soo-saah	'phonetics, phonology'	identical (pseudo)	('sound, voice' // 'science, knowledge')
	b. bən-baat	'daily morning food quest of Buddhist monks'	idiosyncratic (pseudo)	(O // O)
	c. naeŋ-nɑɑŋ	'kind of shackle of wood or bamboo put on cattle'	idiosyncratic	(O // O)
	d. kɑmpek-kɑmpok	'small and uneven; odds and ends'	generic	(O // 'amphora, small urn, bottle')
	e. hun-hɑɑy	'small share in an investment'	delimitative	('invested money, a share' // O)

---

[28] For discussion of the polyfunctionality of Khmer predicates, see Haiman and Ourn (2003).

The example in (47b) is a pseudo-compound originating from the Sanskrit word /piṇḍapāta/. It appears like other alliterative compounds but its repetition was never productive in Khmer. Rather, the second and fourth syllables of the word were elided to create what looks like a two syllable alliterative compound.[29] Pseudo-compounds are alliterative compounds which come from Sanskrit or Pali (Huffman, 1970). The stems of the pseudo-compounds happen to alliterate in the same way as other alliterative compounds in Khmer. The example in (47a) is another pseudo-compound, which comes from the Pali words /sara/ 'sound, voice' and /sastra/ 'science, knowledge.' Both of these stems are attested as independent morphemes in Khmer, but their order is clearly Indic, with the modifier 'sound' preceding a head 'science' (ibid). Khmer word order in noun phrases is always head-initial (Huffman, 1967, 1970; Jacob, 1968; Nacaskul, 1978).

The example (47c) is idiosyncratic in meaning because its meaning does not derive from any change of the semantics of its stems. In fact, the stems are unattested as independent morphemes in Khmer. This compound is not Indic though, making it distinct from the pseudo-compounds. The examples in (47d) and (47e) represent some of the objects with more 'reduplicative-like' semantics. In (47d), for instance, the alliterative derives a more generic meaning from the stem 'bottle.'

As I noted for Mon, it is common to find alliterative compounds for plant or animal names because they tend to be related to the sound the animal makes or its visual characteristics. These qualities often use onomatopoeic mechanisms in languages of the world and especially in Southeast Asia. Animal calls are often repeated, so phonological repetition is used to name the animal. I provide some examples of these nominal forms in (48):

	Allit. Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(48)	a. kɑŋ-kaep	'frog'	('echoing, reverberating, resounding, loudly' // from /ʔaŋkaep/ 'frog')
	b. krɔleeŋ-krɔlooŋ	'kind of grey and black bird'	(O // O)
	c. sɑk-seeh	'kind of plant that has flowers resembling a horse's mane and edible roots'	('hair' // 'horse')
	d. keɛŋ-kɑɑŋ	'a hornbill (bird); a kind of black and white snake'	('well-stuffed, bulging' // 'echoing, reverberating')
	e. ciiŋ-cɑk	'kind of small house lizard'	(O //'sound made by certain geckos')

In (48b) and (48e), the compounds are sound symbolic; mimicking the sounds made by a 'kind of grey and black bird' or a 'house lizard.' In the compounds in (48a) and (48c), the stems are unattested with related semantics. In the former compound, the word for 'frog' is modified. In the latter, the compound is related to a plant type by its similar shape. The alliterative in (48d) nominalizes the stem meanings to create an animal name. The first stem of this compound relates

---

[29] Since I did not exclude words which happen to include phonological repetition by chance, such words are included in my study. They make up only 12/1184 of the words in the database (1.0%).

to the appearance of an animal, a hornbill, which has a very large beak. The second stem relates to the sound of the animal. Indeed, the English word 'hornbill' is similar to the Khmer word in this regard.

Nominal alliterative compounds in Khmer may also be deverbal nouns. Many of these nouns are actually polyfunctional, behaving as both verbs and nouns. The table in (49) gives some examples.

	Allit. Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(49)	a. sɑnthɨk-sɑnthoap	'very loud noise; on a large scale'	('loud noise, abundant' // 'to threaten or intimidate by shouting')
	b. ksən-ksay	'quick disappearance, death'	('destruction, loss' // O)
	c. roam-rɛɛk	'dance which imitates carrying gesture'	('to dance' // 'to carry with a yoke across the shoulders')

In (49a) and (49b), the de-verbal noun has intensified semantics from its stems. In (49c), the compound is simply a nominalized coordinate compound with semantics that are transparently derivable from the stems.

The final category of nominal alliterative compounds in Khmer is entities. As I mentioned in Section 3.31, entities or occupations are commonly found in reduplicative-like expressions because they refer to a person who does an action repeatedly. Khmer has a handful (22) of such compounds. The compound /mee-mɔət/ 'witch, medium,' from "maid/girl' + 'to have a secret understanding," includes a person classifier /mee/ and a modifier /mɔət/. With the exception of the pseudo-compounds, this classifier+modifier pattern is representative of the Khmer alliterative data for entities.

4.32   Verbal Alliterative Compounds in Khmer: Quantitative Observations

Alliterative compounds in Khmer are mainly predicative, not nominal. As I mentioned at the beginning of section 4.3, it is difficult to organize predicative alliterative compounds into discrete lexical categories, for they tend to vary in their syntactic function. Using the semantic categories from (17), we find that most alliterative compounds are stative with evaluative semantics. As in Mon, these compounds tend to refer to the state of a person or object, a description of their characteristics (including shape), or a description of some type of motion. The frequency of the combined semantic categories for predicative alliterative compounds are shown in (50).

	*Combined Semantics (*N=1,020*)*	*Frequency*	*Percentage*
(50)	Stative (evaluative)	586	57.5%
	Shape/Texture (evaluative)	44	4.3%
	Non-motion verb	136	13.3%
	Motion verb	79	7.7%
	Motion Mimetic	68	6.7%
	Sound Mimetic	60	5.9%
	Oscillation	44	4.3%

Aside from stative predicates, alliterative compounds in Khmer may be associated with motion. Each of the categories 'motion verb,' 'motion mimetic' and 'oscillation' involve motion, comprising a large part of the data that is not strictly of the stative category. The remainder of the data consists of either non-motion verbs, which include things like 'to sleep,' 'to insinuate,' or 'to abandon,' and sound mimetics, which are onomatopoeic.

There are many predicative alliterative compounds which have "reduplicative-like" semantics, as in (18), involving intensification, continuity, delimitation, connectivity, repetition, plurality, generality, or reciprocity. Alliterative compounds also may have undergone derivational morphological processes via compounding, such as attributivization, adverbialization, causativization, mimeticization, or the addition of manner to a verb. There is some overlap between the former, more semantic derivations and the latter grammatical derivations. I attempted to categorize the compounds into semantic categories first, as the more grammatical derivations are not as common in the data as the semantic processes. While many compounds have idiosyncratic or identical (zero-derivation) semantics, where there is either an unclear semantic relationship between the stems and the compound or there is no derivation, most alliterative compounds in Khmer do involve some semantic change in compounding that is non-idiosyncratic. The frequency of these semantic categories in the predicative data is shown in (51).

(51)   *Derived Semantics (*N=1,020*)*   *Frequency*   *Percentage*

Zero or Idiosyncratic Derivation		
Idiosyncratic	201	19.7%
Identical / Zero	198	19.4%
Derivational		
Mannerizing	74	7.3%
Attributivizing	70	6.9%
Mimeticizing	25	2.5%
Adverbializing	23	2.3%
Causativizing	20	2.0%
Nominalizing	4	0.4%
Reduplication-like:		
Intensive	151	14.8%
Repetitive	70	6.9%
Continuative	61	6.0%
Generic	47	4.6%
Plural	38	3.7%
Delimitative	16	1.6%
Connective	14	1.4%
Reciprocal	8	0.8%
Total Derivational-type:	**216**	**21.2%**
Total Reduplication-type:	**405**	**39.7%**

From this data, we can see that many (39.7%) of the compounds have reduplicative-like semantics. The rest have either idiosyncratic/identical semantics from the stem meanings (39.1%), or more derivational-type semantics (21.2%). Similarly to Mon, the alliterative compounds in Khmer include both semantics which we expect to find in processes of phonological repetition and other semantics which involve grammatical processes. The semantic categories that Vuori (2000) uses in his dissertation are useful in determining what is a typical semantic pattern for reduplication to have in Southeast Asia. In Khmer, alliteration matches up semantically with these categories.

4.32   *Verbal Alliterative Compounds in Khmer: Qualitative Observations*

All of the semantic fields (stative, mimetic, etc) include compounds which are endocentric and exocentric. Exocentric compounds are semantically idiosyncratic while endocentric and partially-endocentric compounds derive meaning from one or more stems; they are compositional. Many compounds which are endocentric do not appear to have a derived meaning outside of the stem semantics; they result in zero semantic derivation. This category I have labelled 'identical.' Exocentric compounds make up 201/1020 (19.7%) of the predicative alliteratives while the identical category comprises 198/1020 (19.4%). I show examples in (52).

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(52)	a. pɑpleh-pɑplɑh	'frivolous, joking'	(O // O)	idiosyncratic (exo)
	b. thaa-thaaŋ	'having fancy manners'	('to say' // O)[30]	idiosyncratic (exo)
	c. rɔyiiŋ-rɔyooŋ	'drooping or hanging down unevenly; in a drooping manner'	(O // 'dangling or hanging down in disorder; in a dangling or hanging manner')	identical
	d. cbah-cbɑɑŋ	'definitely oldest'	('clear, distinct, definitely' // 'elder, eldest')	identical

The examples in (52a) and (52b) are exocentric compounds. In (52a) the stems are unattested but in (52b), one of the stems is attested, with distinct semantics from the compound meaning. These idiosyncratic compounds either consist of stems which were historically lost, or arose via phonaesthesia. Since there are many compounds which have related semantics and very similar phonology, new compounds may be created with bound roots that have identical meaning. For instance, there are 15 compounds which alternate the rimes /eh/ and /ɑh/ like (52a). The meanings associated with all of these compounds involve 'childishness,' 'playfulness,' or 'triviality.' These meanings are identical to the (52a) 'frivolous, joking.' For instance, the compound /mɔmeh-mɔmɑh/ 'childish, infantile, or frivolous' and is phonologically distinct from (52a) with the exception of the final stem rimes. These rimes have semantic content which make new compounds like (52a) semantically possible. The only criterion for the creation of new

---

[30] The stem /thaa/ meaning 'to say' is bound but attested. Its meaning is unrelated to the compound meaning.

compounds is that the phonological shape be consistent with the template in (44). This same generalization holds for (52b), for which there are a number of compounds with the final rime /aaŋ/ having related semantics.[31] The development of new words or compounds via phonaesthesia is probably more complicated than this, but the presence of so many phonaesthetic compounds in the data makes this hypothesis plausible.

The examples in (52c) and (52d) are endocentric, but have zero derivation between the meaning of the stems and the meaning of the compound. For instance, the stem in (52c) 'dangling' is identical to the compound meaning. If no change in meaning occurs, why do we have alliteration in (52c)? The answer is that the semantics of the stems in such compounds usually include reduplication-like semantics. For instance, the act of 'dangling' involves a back-and-forth movement. Many compounds, categorized as 'oscillating,' derive these exact semantics from stem meanings. The compound /yeek-yook/ 'rocking back and forth' has the bound stem /yeek/ and the free stem /yook/ meaning 'to rock, to balance.' The back-and-forth motion of the compound is derived via the process of ablauting alliteration. Since the *stem* in (52c), [rɔyooŋ], already includes these back-and-forth semantics, a zero derivation occurs. This is the case in (52d) as well, where the stem 'definitely' has an intensifying meaning. The process of compounding itself does not result in intensification here. Rather, it is implicit in the meaning of the stems. Compounds categorized as 'identical' have stems which include alliterative or reduplicative-like meaning. This is an important observation because it unites a large portion of the alliterative data with other data having related semantics.

The largest semantic field of alliterative compounds in Khmer is stative. These compounds have an evaluative nature, referring to the physical condition of an object or person. Some examples are shown in (53).

Stative Alliteratives with Reduplicative Semantics

	Alliterative	Gloss	Composition	Semantic Category
(53)	a. thət-theɛ	'long-lasting, permanent, unceasingly'	('to reside, stay, or remain' // 'firm, solid')	continuative
	b. rɔleak-rɔləə	'confused, lost, disoriented'	('to be shaking, to be bouncing along' // 'uprooted, falling over with the roots pulled out')	generic
	c. rɔloat-rɔlooc	'split or cut in many places, having many scratches'	('peeling off of the skin' // O)	plural
	d. cruət-criəp	'well-soaked or absorbed; thoroughly or completely'	(O // 'to absorb, impregnate, or infiltrate')	intensive

---

[31] Some English ablauting compounds appear to have some phonological regularity. For instance, 'dilly-dally,' 'mish-mash,' and 'flim-flam' all involve an alternation between [ɪ] and [æ]. While there are not many such words in English, one could say that the limited set has some vocalic template.

e. leɛm-ləəm	'appearing or disappearing frequently'	(O // 'in an appearing or disappearing manner')	repetitive

The evaluative nature of stative predicates usually has an intensified quality. For instance, the example in (53d) involves the intensification of the stem 'to absorb.' (At the same time, the compound is a stative verb derived from a transitive verb.) The example in (53c) does not just mean 'cut,' but rather 'cut in many places.' Each of these examples involve semantics which are commonly found in processes of reduplication. Hence, the categories continuative, generic, plural, intensive, and repetitive occur throughout the set of stative predicates.

One could conceivably put all the "reduplicative-like" semantics into one category 'intensive.' However, the lexical semantics of each verb make the specific type of intensification visible. For instance, the intensification of the stem 'to remain,' as in (53a), could only result in the continuity of the action. Since the verb is atelic, an increase in its 'intensity' increases its duration. For the stem in (53e), 'in an appearing or disappering manner,' the action involves a movement between two states (appeared and disappeared). Hence, this change in state is intensified by its repetition. Thus, a verb like 'go' may become 'go far' when reduplicated, but a verb like 'consider' will become 'consider for a long time.' In the first case, the continuity is spatial. In the latter, it is temporal. The semantics of these stative alliterative compounds converge on being 'intensified' in some way, but depend on the lexical semantics of the stems for how this intensification is realized.

In many compounds intensified or repetitive semantics are present when the compound, on the surface, does not appear to involve repetition or intensification. For instance, the verb 'try' entails an action that is often repeated. Verbs like 'lift' and 'reach' do not necessarily involve repetition. A person usually has to "try" something several times. There are 14 alliterative compounds which include the word 'try,' as in /kmiət-kməy/ 'to try hard, to make serious efforts.' However, there are none which include the words 'lift' or 'reach.' The verbs 'try,' 'lift,' and 'reach' are all telic, but only the first entails repetition. Mimetics and verbs of manner also often involve some type of repetition. For instance, the compound /chaep-chɑɑp/ 'to praise, to flatter' entails repetition in the type of action. One must often flatter someone several times before achieving success. Finally, verbs which include the notion of 'hesitation' or 'slow movement' have lexicalized the notion of 'continuity of action.' There are 51 compounds which are glossed with the word 'hesistantly' or 'slowly.' Both concepts involve the continuation of time, via manner of movement. Phonological repetition denotes actions or states involving repetition (as the most iconic), intensity, or some sense of continuity in space or time. Alliterative compounds in Khmer include these common semantic categories, be they lexically entailed, as in 'try,' or transparent, as in (53d) '*well*-absorbed.'

Now we turn to the set of alliterative compounds which are non-stative, those which involve motion, non-motion, or are mimetic. Motion verbs and non-motion verbs are mainly transitive. Stem compounding for these verbs has intensifying semantics or results in syntactic derivation. The former process adds reduplicative-like semantics to the compound, while the latter results in valence-change or the addition of manner to a verb (discussed in (59)). I show some examples in (54).

Allit. Compound	*Gloss*	*Composition*	*Semantic Category*

(54) a. yiəŋ-yiəʋ   'to take long steps in   ('to walk, go' // O)   motion: continuative
walking, to do some-
-thing at long intervals'

b. ʔaət-ʔaəm   'to glance about   ('to raise one's head   motion: repetitive
fearfully and repeat-   to look' // O)
-edly'

c. vɔəŋ-viən   'to make into a curvy   ('circle; to make a   motion: manner
shape, form a spiral'   circle' // 'to bend or
twist; to roll up')

d. tnət-tnɑɑm   'to take extremely   (O // 'to take good   non-motion:
good care of'   care of; carefully')   intensive

e. pkoap-pkun   'to try to please in   ('to satisfy or please   non-motion:
order to gain favor   someone' // O)   repetitive
or benefit'

f. kɔkrɨk-kɔkreeŋ   'to celebrate by   ('sound of many   non-motion:
having many   people walking' //   causativizing
different activities   'sound of waves')
going on at the same time'

The examples in (54a) and (54b) are both motion verbs which involve some reduplicative-like semantics, the former involving continuity, the latter, repetition. The examples in (54d) and (54e) are non-motion verbs which also involve some reduplicative-like semantics; the former is intensive while the latter is repetitive. While these verbs are distinct from the stative verbs, they still encode the same range of "reduplicative-like" semantics involving repetition, intensity, and continuity. The example in (54c) involves a manner of movement, 'making a curve,' while (54f) causativizes onomatopoeic stems. Both of these compounds have semantics related to the others, but include some syntactic derivation. In (54c), the act of "making a spiral" involves a repetitive circular movement. In (54f), a pluralized meaning is present. Repetition and plurality are semantic categories present in the rest of the alliterative data. While I have categorized these two compounds as syntactically-derivational, they still encode reduplicative-like semantics.

Another category which involves explicit repetitive motion is 'oscillation.' These compounds all involve a back-and-forth movement. These alliterative compounds are distinct from motion verbs because they are usually stative.

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(55)	a. rɔyeek-rɔyook	'balanced, rocking or swinging back and forth'	(O // 'to be balanced')	oscillating: repetition
	b. theeŋ-thooŋ	'dizzy, giddy'	(O // 'dizzy')	oscillating: identical (Ø-deriv.)
	c. rɔleak-rɔlɛɛk	'tremulous (of the voice while singing)'	('to be shaking or bouncing' // O)	oscillating: repetition

In (55a), the action of oscillation is explicit in the verb gloss. In (55b) and (55c), the action of oscillation is implicit in the glosses 'dizzy' and 'tremulous.'

There are many alliterative compounds in Khmer which do not directly involve the notion of repetition, but are mimetic. Mimesis is an iconicity between the motion or sound of an object or action and the phonological shape of a word. In this case, such iconicity occurs in alliterative compounding. Mimesis occurs not only with plant and animal names, but also with other sound types and types of motion. I have labelled these categories sound mimetic and motion mimetic. Motion mimetics are different from motion verbs because they are stative and consist only of manners of motion which are iconic with the phonological repetition in alliteration. Motion verbs are not iconic like motion mimetics. Sound mimetics are onomatopoeic. The two tables provide some examples.

<u>*Sound Mimetics*</u>

	Allit. Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(56)	a. pɑpraek-pɑpraok	'continuous crunching sound (leaves)'	(from /praek/ 'wooden frame around wheels of Khmer oxcart' // from /praok/ 'sound of something falling into dry leaves or brushwood.')
	b. kriŋ-kreaŋ	'noisy'	(O // 'sound of something heavy falling')
	c. crɘc-crooc	'sound of a whistle'	('chirping sound, a sharp sound' // 'the sound of water being poured through a spout.')

<u>*Motion Mimetics*</u>

	Allit. Compound	*Gloss*	*Composition (Stem 1 // Stem 2)*
(57)	a. rɔpət-rɔpot	'unruly; walking with the thighs pressed together'	(from attributive prefix /rɔ/ and compound /pət-pot/ 'with a tense gluteus maximus (ass)'
	b. ɲək-ɲɑk	'to swing or move the body in order to attract attention'	(O // O)
	c. ktɔɔ-ktoa	'in a widespread, rumbling manner'	('rumbling' // 'to vibrate, reverberate')

These onomatopoeic compounds in Khmer are interesting for a number of reasons. Morphologically, they are composed of onomatopes themselves. In all cases in (56), the stem morphemes are onomatopoeic and the compound meaning is onomatopoeic as well. For instance, in (56c), the 'sound of a whistle' is high-pitched or 'chirping' and contains noise which sounds like 'water pouring through a spout.' This is remarkable because we do not normally consider sound symbolic words to take part in morphological processes. Indeed, their status in the lexicon is marginal in most languages of the world. Words like 'ribbit' (for a frog) can not take the repetitive prefix in English; i.e. one can not say 're-ribbit.' However, the polyfunctionality of Khmer words makes such morphology possible (Haiman and Ourn, 2003). Many of these compounds are glossed as onomatopes or stative verbs. For instance, the first stem in the compound /haah-haay/ 'a laugh, laughing' means both 'the sound of laughter' and 'laughter.' Of

course, onomatopoeic stems do not always result in onomatopoeic compounds. In (56b) for instance, the stems combine to make a compound meaning 'noisy.' Often, onomatopoeic words occur in phonological repetition because sound may be repeated. Things like chirps or barks do not usually occur in isolation. Birds continue to chirp and dogs continue to bark, in a repeated fashion. In (56a) for instance, a sense of 'continuity' of action is present in the gloss of the alliterative compound. The 'crunching' sound continues to happen. Onomatopoeic compounds share these reduplicative-like semantics with the rest of the alliterative compounds.

The data in (57) shows motion mimetics. These compounds represent particular manners of motion and are syntactically polyfunctional. They can be used as stative predicates, adjectives, or adverbs. For example, in the sentence in (58) (Ñok Thaem, 1960, p.43) the motion mimetic can be understood as any of these lexical categories.

(58) beh-dooŋ niəŋ kɑɑ naen *taan-tiŋ*   knoŋ   əwraa ciə khlaŋ
      heart   her then firm tightening inside body be strong
      "Her heart then was strongly firm and tight inside her body."   (adjectival)
      "Her heart then was firm and tightening strongly inside her body."   (stative verb)
      "Her heart then was firm in a strongly tightening manner inside her body." (adverb)

The compound /taan-tiŋ/ can be translated as 'to be tight,' 'tight,' or 'in a tightening manner.' The glosses of the examples in (57) have the same degree of polyfunctionality. For instance, (57b) can occur alone as an intransitive verb, or with another verb like 'walk,' where the meaning becomes "to walk swinging the body in order to attract attention.' Here, the compound adds manner to the verb 'walk;' it is adverbial. It is important to note that these motion mimetics also seem to involve some notion of repetition. For instance, 'swinging the body' involves an oscillation. The gloss in (57c) 'in a widespread, rumbling manner' involves repetitive movement and sound. While both sound and motion mimetics are a distinct class of alliterative compounds in Khmer, they share semantics with the rest of the compounds.

The final category of Khmer predicative alliteratives are those which involve some syntactic derivation. The process of compounding sometimes results in a distinct change in the part of speech of a compound. These compounds overlap with the other semantic categories, so for instance, compounding could result in both adverbialization and repetition. As I showed in (51), there are five derivational processes that may occur in predicative compounding: the addition of manner, attributivization (stative-verb formation), mimeticization, adverbialization, and causativization. Two of these categories are only present in the Khmer data; they are absent in Mon. There are many alliterative compounds which specify manner on a verb stem which previously did not specify it. This is the most common type of "derivation" like process which occurs in compounding. The other process is mimeticization. This occurs where two stems which are not onomatopes combine to create an onomatopoeic compound. I give examples of all five processes in (59).

Derivational Processes in Compounding

	Allit. Compound	*Gloss*	*Composition*	*Semantic Category*
(59)	a. baŋʔuət-baŋʔɑɑ	'to brag in order to please or impress someone'	('to vaunt or praise' // 'to cause to be joyful;rejoice')	adds manner

b. cɑcak-cɑcoc	'to talk in such a way as to make people agree with one'	('to provoke, irritate, or excite' // O)	adds manner
c. ʔaŋ-ʔaac	'audacious, brave'	(O // 'to dare to do, to be capable')	attributivizing
d. toŋ-taŋ	'sound made by hard object falling'	('large can, barrel, or cask' // 'small bench or chair')	mimeticizing
e. rɔsəp-rɔsiəp	'in a whispering manner; by gossip'	('whispering sound; lightly' // O)	adverbializing
f. sɑmʔət-sɑmʔaaŋ	'to adorn; to deck out carefully or beautifully'	(O // 'embellishing or decorative materials; decorative')	causativizing

In (59a) and (59b) the verb stems do not include manner, but the compounds do. This process is not limited to coordinate compounds, as (59b) consists of a free stem and a bound stem. Since so many alliterative compounds are of an evaluative manner, describing manner of motion or appearance, it is not surprising to find compounds which derive these semantics. In (59c), the compound derives a stative verb from a transitive. This process of derivation is relatively common (70/1020; 6.9%). Earlier, we observed that most of the alliterative predicates in Khmer are stative verbs. The fact that compounding derives this part of speech correlates with its presence in the rest of the data. In (59d), we see that two nominal stems have combined to create a sound mimetic. The process of mimeticization and verb-formation out of nouns is much more frequent than nominalization in alliterative compounds. Alliterative compounds are "par excellence" predicative. In (59e), an onomatopoeic stem becomes an adverbial compound. In (59f), a stative stem becomes a causative verb.

The derivational-type processes that occur in alliterative compounds in Khmer are primarily predicative. Nominalization is rare in the data in comparison to processes which attributivize or adverbialize. There are only 16 nominalized alliteratives, but 93 which attributivize or adverbialize the stem(s). Furthermore, the semantics of these derivational compounds correlate with the rest of the data. There are many compounds which include manner of motion or appearance, similar to how many of these compounds derive it.

In sum, predicative alliteratives are canonically stative and include reduplication-like semantics, involving repetition, intensity, continuity, etc. There are many fields of predicative alliteratives, most of which include these semantic properties. Identical, or zero-derivation compounds entail these semantic properties in one of their stems. Stative, oscillating, motion, and mimetic compounds include these properties as well. Derivational compounding tends to result in evaluative stative verbs, similar to the rest of the data. This makes the Khmer data semantically similar to what we found in Mon.

Khmer alliterative compounds show the same array of meaning types that the Mon compounds do, but also have some differences. First, nominal alliteratives are not as common in Khmer as they are in Mon. Furthermore, Khmer contains compounds which mimeticize or add manner to stems, while Mon does not. While I have focused on the specifics of these semantic categories here, I discuss the entirety of the data in the following section.

V.   Summary and Discussion
5.1   Phonology

We have looked at the data on alliterative compounds in terms of their phonology, morphology, and semantics. While I have made some narrow generalizations on these characteristics in each section, I generalize and offer some explanations as to their structure here.

To begin, alliterative compounds in Mon and Khmer share some phonological characteristics that converge in their prosodic nature. First, both languages have alliterative compounds which are composed of monosyllables or disyllables. The Khmer data includes sesquisyllables as well. There are very few Mon or Khmer compounds with an odd syllable pattern that appear in alliteration; there are few 3 or 5 syllable compounds. Both languages have polysyllabic words consisting of three or more syllables (mostly Indic). Yet, neither language, composes alliterative compounds out of them. There are many monosyllabic words in both language, but these are rarely combined with disyllables (or sesquisyllables) in processes of alliteration. Second, alliterative compounds contain bound and free stems. It is tempting to think of the bound stems as the result of a productive reduplication process. However, in both Mon and Khmer, the phonological shape of these stems is unpredictable. Yet, there is a regularity to such compounds. Alliterating monosyllabic stems always alternate their rime or vowel, while in alliterating disyllabic stems, only the final syllables alternate. In Khmer, specific rime types tend to alternate with other specific rime types, as we noted in (33). Finally, we observed that stems with different registers usually result coordinate compounding. Such compounding is rare in both Mon and Khmer; compounding tends to select stems of the same register. Since register is a word-level effect in Mon and Khmer, there is evidence to consider these compounds as a structural unit. The phonology of these compounds only seems completely irregular; a closer inspection of the data yields some regularities.

While the exact phonemes of the alternating rimes in alliterative compounds are unpredictable across the stems, the Khmer data shows a significant tendency to favor higher, fronter, and shorter vowels in initial stems while lower, backer, and longer vowels are more likely to appear in final stems. Furthermore, there is a preference for codas to consist of sonorants in compound-final position, while the initial stem usually ends in an obstruent. These patterns were present in Middle Khmer as well (Schiller, 1999). The Mon data does not lead us to posit these tendencies.

The Khmer pattern leads us to conclude that final rimes in alliterative compounds are more sonorous than initial rimes. Here, I assume that sonority corresponds to a duration over which pitch, signalling stress in Khmer, can occur. Higher and fronter vowels tend to be shorter in duration than low and back vowels. Low vowels are more sonorous (Barnes, 2002). Sonorants also extend the rime duration over which pitch is realized in Khmer. Sonorants, being voiced continuants, allow this duration to be extended. There is evidence of rime sonority being significant in final positions in Khmer versification (Jacob, 1979). Khmer songs which have lines ending in an obstruent always change the obstruent to a homorganic nasal sonorant.

We would expect this sonority to play a role in the final stem of Khmer alliterative compounds because this position is also the final rime of the compound. Word-final position is prosodically marked in many languages (Barnes, 2002). We can look at these changes in sonority across Khmer stems as a process of final stem fortition. Since Khmer polysyllabic words are iambic disyllables or sesquisyllables, word-final strengthening correlates with the position of

stress in the language. Word-final fortition is a characteristic of Khmer alliterative compounds, which set them apart from alliterative compounds in Mon.

The structure of monosyllabic and disyllabic/sesquisyllabic alliterative compounds in Mon and Khmer is given by the templates in (60) and (61)

Monosyllabic Stem Pattern:

(60)   Initial Stem   +   Final Stem
       $Onset_1$ + Vowel/$Rime_1$ + $Onset_1$ + Vowel/$Rime_2$

Disyllabic/Sesquisyllabic Stem Pattern:

(61)   Initial Stem   +   Final Stem
       $syllable_1$ + $Onset_1$ + Vowel/$Rime_1$ + $syllable_1$ + $Onset_1$ + Vowel/$Rime_2$

In both (60) and (61), the vowel or rime is different in the second stem from the first stem. In monosyllables, only the onsets of both stems are the same. In disyllables, the initial syllable is the same across stems as well. In Khmer, the final rime of these compounds is strengthened by increasing its sonority. In the patterns above, the contrasting rime occurs in the word-final position of the compound. Alliteration is the repetition of the initial parts of the stems. There is a "same" + "different" pattern in the phonological shape of these compounds.

Haiman and Ourn (2000) state that the tendency for Khmer to be a prefixing language relates to a general pattern of Same + Different compounding in Khmer. Prefixes normally have a fixed phonological shape (given no morphophonological process) while the stem may vary. Alliterative compounding matches this pattern prosodically. The initial syllables/onsets of the stems are the same, while the final rimes of the stems are different. This phonological pattern in both Mon and Khmer is related to this morphological pattern.

The patterns in (60) and (61) are also related to a much more general pattern in Southeast Asia, elaborate expressions. The four syllable elaborate expression is a common morphological sequence in most Southeast Asian languages (Vuori, 2000). The pattern in most languages is a sequence where the first and third words alternate while the second and fourth are the same, or a pattern where the second and fourth words alternate while the first and third are the same. Khmer has its fair share of these types of expressions (Jacob, 1968), all of which are of the shape A-B-A-C (with alternating second and fourth words). Mon has the same A-B-A-C type of elaborate expressions (Bauer, p.c.). How is this related to alliteration? We noticed that neither Mon or Khmer use larger stem types (of 3 or more syllables) in creating alliterative compounds. Recall from section 3.2 and 4.2 that derivational morphology can apply to each stem of an alliterative compound, regardless of whether that stem is a free morpheme. For instance, (43c) /prɑnəŋ-prɑnoŋ/ 'rough, uneven, disorderly' is derived from the reciprocal prefix /prV/ and the compound /nəŋ-noŋ/ 'music played on small gongs.' The prefix applies to both /nəŋ/ and /noŋ/, even though they are unattested as separate morphemes. The A-B-A-C pattern in Mon and Khmer match the template in (61). I show some elaborate expressions from Khmer in (62).

(62)   *Elaborate Expression*   *Gloss*   *Transliteration*
       a. təʊ veɲ təʊ mɔɔk   'go back and forth'   go-back-go-come
       b. lah coh lah laəŋ   'scurry up and down'   scurry-go.down-scurry-go.up
       c. stəə ruəh stəə slap   'at death's door'   be.on.the.point-alive-

			-be.on.the.point-die
d. piən traok piən traŋ		'to act thoughtlessly'	climb.on-mounds(earth)
			climb.on- (bound root)

In (62), we notice that the expression /təʊ-veɲ-təʊ-mɔɔk/ literally means 'go-back-go-come,' but as an expression means 'to go back and forth.' This expression uses phonological repetition to convey a repetitive manner of motion; it is not sufficient to use only the word /təʊ/ 'go' (Jacob, 1968).[32] In all of these expressions, the first and third word need to be repeated while the second and fourth alternate. There is no syntactic reason for repeating the first and third word, it is a purely prosodic governed, identical to what we find in alliterative compounding.

There is also a morphological and semantic similarity between alliterative compounds in Mon and Khmer and elaborate expressions. Elaborate expressions are used to convey repetitive back-and-forth movement. For example, (62a)-(62c) each involve a movement between two locations (or states, as in (62c). Alliteration achieves the same semantic effect. Furthermore, alliteration often includes bound roots which are unattested outside outside of the compound. The example in (62d) includes a bound root. Indeed, this example looks phonologically identical to the two stem alliteration pattern.

If there is a relationship between alliteratives and elaborate expressions, it would nicely explain why we only find monosyllabic or disyllabic stems. Simply put, the first and third syllables in alliterative compounds cannot be identical if the stems are larger. Furthermore, we find an explanation for why derivational morphology has applied to many bound stems in Khmer alliteratives. In order to maintain this 1-3 same, 2-4 different pattern, a repeated "dummy" syllable is inserted. There is a prosodic, 4 syllable structure, in both elaborate expressions and alliteratives which motivates this morphology and this restriction on stem syllables.[33]

5.2   *Morphology*

Both Mon and Khmer show much variation in how alliterative compounds are formed. Both languages have a pattern where two bound, two free, or one bound and one free stem combine to create a compound. In Mon, there is a significant tendency to find a F/B pattern. In both languages, the F/F pattern is more prevalent than other patterns. While the Vietnamese data in (5) suggests that the F/B pattern may be more general across Mon-Khmer languages, it is not

---

[32] One can say /təʊ veɲ/ meaning 'go back,' so the word /veɲ/ does not require /mɔɔk./.

[33] There is also a relationship between the presence of phonaesthetic rimes in Khmer and the alliterative pattern. If we are to look at the repeated phonological material in such compounds, then we find that the rime, being the only part which alternates, could easily be associated with the meaning-bearing element of such compounds. For instance, consider the compound / baŋkii-baŋkiə / 'torpid, sluggish, slow' in (34e). Given that speakers often pay closer attention to parts of an acoustic signal which contrast, one's attention is often guided towards the portion of the compound which contrasts, the vowel or rime. This salient part of the compound may have a tendency to be semanticized due to its prominence. Such a hypothesis needs to be worked out more fully though.

shown in Khmer. In both languages, bound roots tend to be unattested outside of the alliterative compounds where they occur. The languages have three patterns of compounding: exocentric, endocentric, and partially-endocentric. This last category applies to compounds for which one of the stems is a semantic head while the other is not.

One interesting characteristic of the morphology of these compounds is the application of derivational morphemes which are bleached of their semantics. As we saw in section 3.2 for Mon and in 4.2 for Khmer, derivational morphemes apply to both stems of the alliterative compound. This double application of prefixes or infixes has no morphological purpose, but occurs to fill a prosodic template, shown in (60) and (61). This gibes closely with the finding of Haiman and Ourn (2003) who discuss the polyfunctionality of derivational morphemes in Khmer, which often apply with zero-derivation. Morphologically, Khmer differs from Mon in having a set of compounds which derive a different part of speech from the stem types. Such processes derive more verbs than nouns in Khmer, which matches the tendency for Khmer alliteratives to be predicative.

5.3   Semantics

Mon and Khmer show many similarities as to what type of thing will be encoded by alliteration. First of all, nouns appear to be less common in alliterative compounds than predicatives. In Mon, 80/202 (39.6%) compounds were nominal while in Khmer, only 186/1184 (15.7%) were nominal. It appears that alliteration favors predicatives over nominals for both Mon and Khmer. The most common semantics for alliterative nouns are objects, followed by plants and animals, followed by entities. I show data for both Mon and Khmer in (64).

(64)	*Nominal Semantics*:	*Mon*	*Khmer*
	Objects	37	100
	Plant/Animal	26	32
	Entities	8	22

Mon and Khmer have the same tendency to encode objects with alliterative expressions over plants and animals and entities. Many of these objects have reduplicative-like semantics which usually consist of some type of repetition. Such semantics are usually entailed by the type of object. For instance, a word like 'sewing machine' or 'shuttlecock' have something to do with repetition.

In both Mon and Khmer, the typical predicative alliterative is a stative verb. The other types of predicates encoded by alliteration are motion verbs, non-motion verbs, and mimetics. All of these types has some reduplicative-like semantic property, including repetition, intensity, continuation, generality, etc. These semantics may be hidden in a compound, such as in the word 'try,' which implies a set of repetitions. Many of the compounds labelled as 'identical' have stems which imply these semantics as well. In (65) I show the semantic types in both languages.

(65)	*Predicative Semantics*:		*Mon*	*Khmer*
	Stative:	Evaluative	58	586
		Shape/Texture	5	44
	Non-stative:	Motion verb	24	79

	Non-motion verb	16	136
	Oscillating	3	44
Mimetic:	Sound Mimetic	12	60
	Motion Mimetic	3	68

The data in (65) show that stative predicates are the most frequent type of alliterative compounds, followed by non-stative predicates, followed by mimetics. Mimetic compounding is more prevalent in the Khmer data than in the Mon data. All the semantic types in (65) include reduplicative-like semantics as proposed by Vuori (2000).

Both Mon and Khmer have an interesting pattern of semantics in compounding. If a stem already includes some repetitive or intensive semantics, then the alliterative compound which contains the stem may not alter the meaning between the stem and the compound. For instance, if we have a stem meaning 'to jump up and down,' the compound may mean the same thing. In this case, the semantics involving repetition are implied by one of the stems. This is also present where a lexeme implies repetition or intensity. For instance, reduplicative-like processes often represent occupations because they a person must repeat the same action over and over in a particular occupation. Furthermore, verbs like 'try' implies this sense of repetition. In Khmer, we find compounds which derive the word 'try' from stems which do not contain it. For instance, the compound /pkoap-pkuɲ/ 'to try hard to please someone' is derived from the free stem /pkoap/ meaning 'to satisfy.' The semantics of words like 'try' include repetition. These implied semantics are an interesting characteristic of the alliterative data.

Alliterative compounds have mostly 'expressive' meaning in Mon and Khmer. Expressives in Mon-Khmer languages are a particular type of word that may in fact have independent status as a lexical category (Diffloth, 1976). Elaborate expressions are phonologically and semantically related to alliterative compounds and are another type of expressive in Mon-Khmer languages. The fact that the entire category 'expressive' shows a great deal of repetition in these languages is perhaps good evidence for their being treated as a separate lexical category. Indeed, their semantics make them stand out from the rest of the words in the lexica and make them intriguing for both a learner of a Mon-Khmer language and the linguist who tries to understand them.

VI. Final Remarks

The study of alliteration in Mon-Khmer languages is interesting because it is so prevalent throughout the family. The phonological tendencies regarding sonority in Khmer inspire serious investigation into how an alliterative pattern would have looked in the historical phonology of Mon, Vietnamese, and other languages. My main choice for looking at only Mon and Khmer in this study is to limit the data. One could write volumes on these types of compounds/words. Indeed, the work of Vuori, focusing mainly on semantics is a dissertation on its own. In this paper I have provided a much needed, in-depth look at how alliteration and "chiming" compounds work in Khmer and Mon. While there is a history of work in this area in Khmer, I hope to have provided some empirical data on how Mon works as well. Alliterative compounds in Khmer and Mon are a unified category which is organized by its phonological structure first, then by a related set of semantics. Alliterative morphology is diverse because compounding must adhere to the templates as I have shown in (60) and (61). Alliterative compounding is an

example of a phonologically-conditioned morphological process that looks like reduplication without any fixed segmentism, but is actually a type of compounding which follows a phonological template.

There are many areas that I did not cover in this paper. For instance, while I mentioned the similarity between elaborate expressions and alliterative compounds based on their "same+different" pattern, I wonder whether languages with a A-B-C-B elaborate expression pattern show a similar correlation with their alliterative compounds, i.e. do we expect a "different+same" pattern in alliteration if the language has elaborate expressions like these? The relationship between "same+different" pattern and prefixation in Khmer has been established here and in Haiman and Ourn (2000). One corollary of this view is that a 2-4 same, 1-3 different pattern in elaborate expressions must only occur in languages which have a tendency for suffixation. This corollary remains an open question of further research. The study of alliterative compounds and expressives in any Mon-Khmer language is a rewarding endeavor that is bound to increase our understanding of both linguistic structure in Southeast Asia and the role of compounding in linguistic systems.

References:

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk (1997) *Reduplication with Fixed Segmentalism.* Ms., University of Massachusetts, Amherst. [ROA-226, http://ruccs.rutgers.edu/roa.html]
Apte, Vasudeo Govind (1996) *Sanskrit-English Dictionary*, Hippocrene, New York
Banker, Elizabeth (1964) *Bahnar Reduplication*, in Mon-Khmer Studies, Volume 1.
Barnes, Jonathan (2002) *Positional Neutralization: A Phonologization Approach to Typological Patterns*, doctoral dissertation, University of California, Berkeley.
Bauer, Christian (1989) *The verb in spoken Mon.* from Mon-Khmer Studies 15.
Bergen, Benjamin K. (2004) *The Psychological Reality of Phonaesthemes*, in Language, 80, 2.
Diffloth, Gérard (1976) *Expressives in Semai,* in Austroasiatic Studies, I, Phillip Jenner, Laurence Thompson, and Stanley Starosta (eds). University of Hawaii Press.
_____ (1984) *The Dvaravati Old Mon Language and Nyah Kur*, Monic Language Studies, Vol. 1, Chulalongkorn University.
_____ (1994) *i: big, a: small*, in Sound Symbolism, Leanne Hinton, Johanna Nichols, and John J. Ohala (eds), Cambridge University Press.
Ferlus, Michel (1979) *Formation des Registres et Mutations Consonantiques dans les Langues Mon-Khmer*, in Mon-Khmer Studies, VIII: 1-76.
Garrett, Ancdrew (2001) *Reduplication and Infixation in Yurok: Morphology, Semantics, and Diachrony* in International Journal of American Linguistics, 67, 3, 264 – 312.
Goldman, Robert P. and Sally J. Sutherland Goldman (1999) *Devavāṇīpraveśikā: An Introduction to the Sanskrit Language,* University of California Press, Berkeley.
Haiman, John and Noeurng Ourn, (2000) *Symmetrical Compounds in Khmer*, in Studies in Language 24:3, 483-514.
Headley, Robert K., Kylin Chhor, Lam Kheng Lim, Lim Hak Kheang, Chen Chun (1977) *Cambodian-English Dictionary*, Catholic University of America, Washington D.C.
Henderson, Eugénie J. A. (1952) *The Main Features of Cambodian Pronunciation*, in Bulletin of the School of Oriental and African Studies, University of London, Volume XIV, part 1.

_____ (1965) *The Topography of Certain Phonetic and Morphological Characteristics of South East Asian Languages*, in Lingua, 15, 400-434, special edition "Indo-Pacific Linguistic Studies Part II: Descriptive Linguistics, G.B. Milner and Eugénie J.A. Henderson (eds). North Holland Publishing.

Huffman, Franklin E. (1967) *An Outline of Cambodian Grammar*. Phd Dissertation, Cornell University.

_____ (1970) *Modern Spoken Cambodian*, Yale University Press, New Haven.

Jacob, Judith (1963) *Prefixation and Infixation in Old Mon, Old Khmer, and Modern Khmer*, in Cambodian Linguistics, Literature, and History: Collected Articles of Judith Jacob, Smyth, David A. (ed). SOAS, 1993.

_____ (1968) *Introduction to Cambodian*. Oxford University Press, London.

_____ (1976) *Affixation in Middle Khmer with Old and Modern Comparisons*, in Cambodian Linguistics, Literature, and History: Collected Articles of Judith Jacob, Smyth, David A. (ed). SOAS, 1993.

_____ (1979) *Observations on the Uses of Reduplication as a Poetic Device in Khmer*, in Cambodian Linguistics, Literature, and History: Collected Articles of Judith Jacob, Smyth, David A. (ed). SOAS, 1993.

_____ (1992) *Some Comments on the Relationship between Khmer words having Identical Vowel Nuclei and Final Consonants*, in Cambodian Linguistics, Literature, and History: Collected Articles of Judith Jacob, Smyth, David A. (ed). SOAS, 1993.

Matisoff, James A. (1973) Tonogenesis in Southeast Asia'. In Larry M. Hyman (ed.), *Consonant types and tone* (*Southern California Occasional Papers in Linguistics* 1), 71-95. Los Angeles: University of Southern California.

_____ (1982) *The Grammar of Lahu*, University of California Press, Berkeley

_____ (1994) *Tone, Intonation, and sound symbolism in Lahu: loading the syllable canon,* in Sound Symbolism, Leanne Hinton, Johanna Nichols, and John J. Ohala (eds), Cambridge University Press.

_____ (2003) *Handbook of Proto-Tibeto-Burman*, University of California Press, Berkeley.

Nacaskul, Karnchana (1978) *The Syllabic and Morphological Structure of Cambodian Words*, in Mon-Khmer Studies, VII, 183-200.

Rischel, Jørgen (1995) *Minor Mlabri: A Hunter-Gatherer Language of Northern Indo-China*, Museum Tusculanum Press, University of Copenhagen.

Schiller, Eric (1999) *Reduplication in Khmer Morphology*, Linguistics Unlimited

Shorto, H.L. (1962) *A dictionary of modern spoken Mon*. Oxford University Press. London

Varasarin, Uraisi (1984) *Les Éléments Khmers dans la formation de la langue siamoise*, from "Langues et civilisations de l'asie du sud-est et du monde insulindien," 15, SELAF.

Vuori, Vesa-Jussi (2000) *Repetitive Structures in the Languages of East and South-east Asia*, Studia Orientalia, 88, Finnish Oriental Society, Helsinki

Wayland, Ratree and Allard Jongman (2003) *Acoustic Correlates of Breathy and Clear Vowels: The Case of Khmer*, in Journal of Phonetics, 31, 181-201.