A Conservative Vowel Phoneme Inventory of Sumatra: The Case of Besemah

Bradley McDonnell

ARIZONA STATE UNIVERSITY

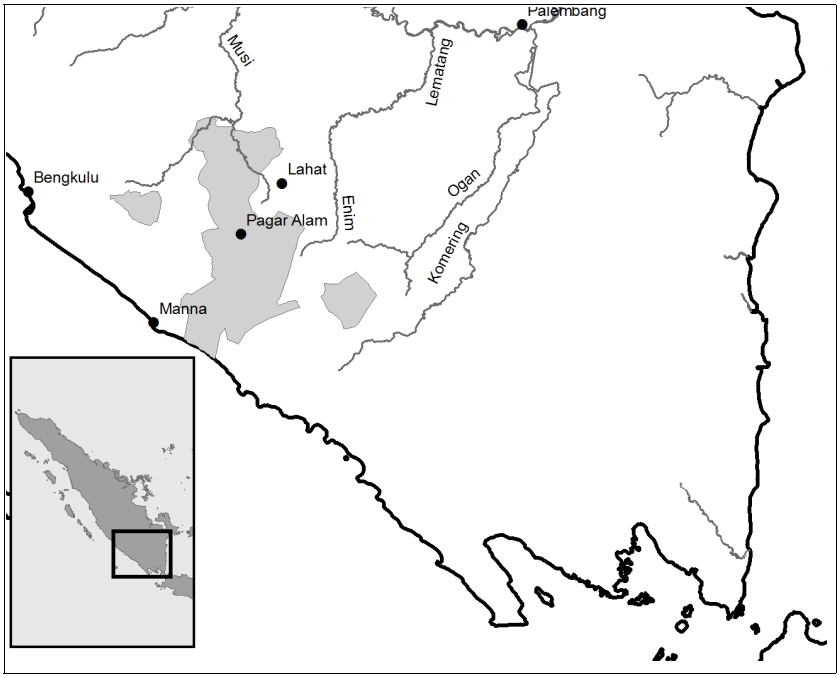
In general, the Malayic languages of Sumatra show a vowel phoneme inventory that is either equal to or greater than that of the five phonemes in Standard Malay/Indonesian (not including the so-called *pepet* vowel); dialects of Jambi Malay, Palembang Malay, and Minangkabau all show five phonemes. However, from recent fieldwork in Besemah, a Malayic language in the highlands of southwest Sumatra, I shall describe an inventory that evinces the smallest number of vowel phonemes in any Malayic language of Sumatra described to date. That is, Besemah has three vowel phonemes. Nevertheless, the analysis of Besemah vowel phonemes is not straightforward; vowel lowering in closed syllables, vowel harmony, and raising of the word-final low central vowel all cloud the analysis of a three-vowel system. Furthermore, Besemah currently is experiencing intense pressure from other standard and nonstandard varieties of Malay/Indonesian, which demonstrates how intense contact involving diglossia and increasing bilingualism in these other varieties of Malay/Indonesian appear to be leading to an emergence of mid vowels in Besemah.

1. INTRODUCTION.[1] The Besemah[2] vowel phoneme inventory is striking as it evinces the smallest vowel inventory of any Malayic language of Sumatra described to date: the vowel phonemes of Besemah are /i, u, a/ with the addition of the phonetically inserted *pepet* vowel occurring with a limited distribution.[3] In comparison to all other Malayic languages, only the vowel inventories of the Malayic languages of Borneo, including Brunei Malay, Berau Malay, and Banjar Hulu Malay could be considered smaller, as these languages lack the pepet vowel that appears in Besemah (cf. Adelaar 2005:206). It is far more common that a Malayic isolect of Sumatra shows mid vowel phonemes, although the degree to which these permeate the language varies. Prentice

1. I would like to acknowledge Dawn Bates, two anonymous reviewers, and in particular Uri Tadmor for extensive discussion and comments on this paper. I would like also to acknowledge the following individuals and institutions for their support of this research: Karen Adams, Hengky Firmasyah, Elly van Gelderen, David Gil, Bambang Kaswanti Purwo, Yassir Nasanius Tjung, and Peter Suwarno, as well as the Max Planck Institute for Evolutionary Anthropology Field Stations in Jakarta and Padang, the Center for Languages and Cultures at Atma Jaya Catholic University in Jakarta, and the American Indonesian Exchange Foundation. Lastly, I would like to thank Matthew Connor for providing the map of Sumatra. This research was made possible by a Fulbright research grant. I am solely responsible for all content in this paper.

Oceanic Linguistics, Volume 47, no. 2 (December 2008)

MAP 1. THE BESEMAH LANGUAGE AREA†



† Source: McDowell (2007)

and Usman (1978:134), in a discussion of Kerinci, a Malayic language of West Sumatra that lacks mid vowels in penultimate syllables, observe that "in common with other Western Indonesian languages such as Javanese and Sundanese, almost all Malay dialects have undergone a phonemic split whereby *u and *i are represented in the modern language by two phonemes: *i by *i* and *e* and *u by *u* and *o*." That is, Malayic languages of the eastern lowlands of Sumatra, including dialects of Palembang Malay and Jambi Malay,[4] and languages of the western coast of Sumatra, including Minangkabau and Ser-

2. Besemah (alternatively, Pasemah) is a little-known Malayic language spoken in the remote highlands of South Sumatra in western Indonesia. It is considered a "vernacular" Malay by Adelaar and Prentice (1996) and Adelaar (2005), and appears to form a subgroup with the languages of the highlands and lowlands of southwest Sumatra. This group of dialects/languages has been traditionally referred to as Middle Malay by Brandes (1884), and subsequently labeled Central Malay by Adelaar (1992). See Mintani (1980) for an initial survey and McDowell (2007) for a recent survey of Central Malay and other South Sumatran languages. Voorhoeve (1955) also discusses earlier research on these languages. Linguistic publications on Central Malay have been limited and in recent time almost nonexistent, but the lexicographical work, short grammatical sketch, and texts by the late O. L. Helfrich (1904, 1915, 1921, 1927, 1933) form by far the most significant work on these languages. Even though I reference Helfrich (1904), this study is based primarily on my own fieldwork in a number of villages in the Besemah highlands about ten kilometers outside of the city of Pagar Alam.
3. I refer to vowel phoneme systems here as all vowel phonemes with the exclusion of the mid central vowel, which I claim is phonetically inserted (see fn. 5 below).
4. This is the so-called city dialect of Palembang Malay from Aliana et al. (1978) and the *Tanjung Raden* and *Mudung Darat* dialects of Jambi Malay from Regina Yanti (pers. comm.).

away, all show mid vowels, although the latter appears to show "less-developed" mid vowels, while the former shows "well-developed" mid vowels (see section 4 for possible implications; cf. Adelaar 1992:45–46). Prentice and Usman (1978:158, fn. 17) and later Prentice (1990:191) also point out that the rise of mid vowels in Standard Malay/Indonesian results from a universal lowering of high vowels in final closed syllables and an unpredictable lowering in nonfinal syllables, together with a vowel harmony rule that gives rise to doublets and minimal pairs in addition to loan words now entering into the language with mid vowels intact.

Besemah demonstrates many of the characteristics listed above by Prentice and Usman, but with the important difference that, in the case of Besemah, all vowel changes can still be explained through conditioned allophonic variation. That is, all high vowels in Besemah lower in closed syllables, and vowel harmony lowers high vowels in open syllables in the penult when followed by a lowered high vowel in a root-final closed syllable (see 2.1.1). Unrelated to high vowels, but important to the analysis of the vowel inventory, Besemah evinces two phonetically overlapping mid central segments showing complex characteristics, which warrant a distinction between mid central segments that occur root/word-finally and mid central segments that occur in penultimate and prepenultimate syllables. This analysis demonstrates that the latter of these is the pepet, the phonetically short epenthetic vowel, while the prior is an allophone of the word-final /a/, which undergoes raising in the final position of the root and the word (see 2.2).

Many descriptions of Malay/Indonesian do not discuss the status of mid central segments and usually consider the mid central vowel to be a phoneme with little discussion. However, Cohn (1989:175) mentions the possibility that the mid central vowel in Standard Indonesian (analogous to the *pepet* here) is epenthetic. I do not intend to defend this point fully here for other Malay(ic) dialects, but see 2.2.1 for a discussion of mid central vowels in Besemah.

Besides factors of allophonic variation, the three-vowel phoneme inventory of Besemah appears to be under intense pressure from other Malayic languages with larger vowel phoneme inventories, such as Standard Indonesian, the language of education, Jakarta Indonesian, the language of the media, and Palembang Indonesian, the language of wider communication.[5] The Besemah speaking community demonstrates a clear case of diglossia. Younger speakers of Besemah are becoming increasingly proficient in all three varieties of Indonesian through increased access to lowland cities—namely Palembang—universal education, and media, such as Indonesian television, all of which were greatly restricted until recent decades.[6] The current situation creates a stark contrast between older speakers, who either cannot speak these varieties of Indonesian or have limited ability in them, and younger speakers who have the ability to speak both languages at varying levels. This is relevant to the discussion of vowel phonemes, given that a clear sign of this change is evinced in the (in)ability of younger speakers and older

5. This naming convention follows Gil (1994) and Tadmor and Gil (1997).
6. It is reported that paved roads and electricity came to the villages surrounding Pagar Alam in the 1980s, and beginning in the 1990s television became available to more and more of the residents in these villages. At the present time, buses and vans pass through these villages many times a day destined for Jakarta, Bengkulu, Palembang, and other major cities of Sumatra.

speakers to produce mid vowels in particular positions in the roots of loan words from these Indonesian varieties (see 3.2).

To date, there exists no careful description of Besemah that accounts for the allophonic variation above. Therefore, the bulk of this paper, in section 2, presents a synchronic description of the Besemah vowel phonemes, maintaining that Besemah indeed has a three-vowel phoneme inventory accompanied by the phonetically inserted pepet vowel. Building on this description, 3.1 presents a discussion that highlights the retention of high vowels in Besemah from Proto-Malayic. Section 3.2 addresses issues of contact between Besemah and other varieties of Indonesian, which raises the question, could Besemah be developing mid vowels? Finally, 3.3 poses questions regarding the phonemic nature of final vowels and the status of mid vowels in Malayic languages of Sumatra that result from this analysis of Besemah vowels.

2. VOWEL PHONEMES IN BESEMAH.[7] Besemah has three vowel phonemes /a, i, u/ and a phonetically inserted mid central vowel, henceforth the pepet, represented here as [ə]. The three phonemes and the pepet are in table 1. The nine allophones for the three vowel phonemes are in table 2. In this discussion of Besemah vowels, it is convenient to split vowels into two groups: *high* vowel phonemes (/i, u/) and their respective allophones ([i, ɪ, e, u, ʊ, o]), and the *central* vowel phoneme (/a/) and its respective allophones ([a, ə, ɨ]); the pepet [ə] will be included in the discussion of the latter due to some overlapping features with the low vowel allophones. All discussion here of mid vowels refers to the mid-front vowel [e] and mid-back vowel [o]; in Besemah these are always allophones, but in other Malayic varieties these may represent phonemes /e/ and /o/.

2.1 HIGH VOWELS. There are two high vowel phonemes in Besemah: the high front unrounded /i/ and the high back rounded /u/. High vowels occur freely in the ultimate and penultimate positions of disyllabic roots, but only occur in the antepenultimate position of trisyllabic roots when followed in the next syllable by a low vowel /a/, shown in (1) below. Antepenultimate high vowels, however, are deleted in some cases; /siapa/ 'who' com-

TABLE 1. VOWEL PHONEMES

	FRONT	CENTRAL	BACK
HIGH	/i/		/u/
MID		(ə)	
LOW		/a/	

TABLE 2. VOWEL ALLOPHONES

	FRONT	CENTRAL	BACK
HIGH	[i] [ɪ]	[ɨ]	[u] [ʊ]
MID	[e]	[ə], [ə]	[o]
LOW		[a]	

7. Appendix 1 presents the consonant phonemes of Besemah. The consonant phonemes are annotated with the main alternations and distributions of some of the phonemes.

monly occurs as [sapə], and /tuapa/ 'what' commonly occurs as [tapə].[8] The high front vowels show contrast with each other and all other vowel phonemes, including the pepet.

(1)

SYLLABLE	ROOT (/i/)		ROOT (/u/)	
ANTEPENULTIMATE	[**pi**yapʊŋ]	'fishing floater'	[**su**waɣə]	'voice'
PENULTIMATE	[ʧə**ni**lah]	'sandal'	[b**u**bu]	'fish trap'
ULTIMATE	[ba**ɣi**]	'old'	[sik**u**]	'elbow'

2.1.1 Allophonic Variation in High Vowels. Allophonic variation in high vowels is predictably determined by two principles: the presence (or absence) of a coda, and vowel harmony. In the first principle, /i/ and /u/ surface, respectively, as [i] and [u] in open syllables, but undergo lowering to [ɪ, e] and [ʊ, o] in closed syllables. The height of lowered high vowel allophones, whether /i/ lowers to [ɪ] or [e] and /u/ lower to [ʊ] or [o], is conditioned by the place of articulation of the following coda consonant. Impressionistically, /i/ and /u/ are lowered to mid-high vowel allophones [ɪ] and [ʊ] preceding oral/nasal conditioning consonants, which include [p, t, k, m, n, ŋ, l, r, s, ɣ], while /i/ and /u/ are lowered to mid vowel allophones [e] and [o] preceding a glottal conditioning consonant [h, ʔ].[9] Allophonic variation in penultimate and ultimate high vowels is illustrated in (2) below (where CVC-O = oral/nasal codas and CVC-G = glottal codas).[10]

(2)

	/i/	/u/
CV	/bini/ > [**bi**ni] 'wife'	/kutu/ > [**ku**tu] 'louse'
	/rami/ > [ra**mi**] 'busy'	/kamu/ > [ka**mu**] 'you (PL)'
CVC-O	/riŋkih/ > [**rɪŋ**keh] 'great'	/puntuŋ/ > [**pʊnt**ʊŋ] 'firewood'
	/kisit/ > [ki**sɪt**] 'go'	/abut/ > [a**bʊt**] 'heavy'
CVC-G	/ɣusik/ >[ɣu**seʔ**] 'visit'	/basuh/ > [ba**soh**] 'wash'

In the second principle, high vowels in open penultimate syllables are lowered by the presence of a final high vowel that has been lowered via a closed syllable; that is, /gigi/ 'teeth' and /kutu/ 'louse' always surface as [gigi] and [kutu] respectively, while /gigit/ 'bite' and /pitung/ 'hold' surface as [gɪgɪt] and [pɪtʊŋ] in normal speech. Vowel harmony is discussed in more detail in the next section.[11]

8. As in many other Malayic languages of Sumatra (except Minangkabau), the pepet is almost exclusively the only vowel present in the antepenultimate syllable; the phoneme /a/ is barred from the antepenultimate position and the phonemes /i/ and /u/ are restricted to vowel sequences in the antepenultimate position. The pepet even arises in loan words, such as *polisi* 'police' via Standard Indonesian from English, which is realized [pəlisi].
9. Vowel lowering can be stated by the following informal rules:

   (i) /i/ → [ɪ], /u/ → [ʊ] / __{p, t, k, m, n, ŋ, l, r, s, ɣ}, and

   (ii) /i/→ [e], /u/ → [o] / __ {h, ʔ}.

   In (i), the high vowel phonemes /i/ and /u/ are lowered to mid-high vowel allophones [ɪ] and [ʊ] in a closed syllable when followed by an oral/nasal coda consonant. In (ii), the high vowel phonemes /i/ and /u/ are lowered to [e] and [o], respectively. It is also important to note that [ʔ] is an allophone of the phoneme /k/. As is noted in appendix 1, /k/ generally changes to [ʔ] in root/word-final positions, such as /ambik/ 'take' that is realized as [ambeʔ]. There is, however, a small number of roots that maintain a final [k], such as /ibik/ 'duck', which is realized as [ibɪk]. It is noteworthy that in the case of /ibik/ 'duck' the final allophone of /k/ is [k] and that the allophone of /i/ is [ɪ], while in the case of /ambik/ 'take' the final allophone of /k/ is [ʔ] and the allophone of /i/ is [e], yielding [ambeʔ].
10. Glottal sounds [h, ʔ] primarily occur in the root/word-final position and are usually prohibited from the word-medial position.

Vowel harmony is one of two phonological alternations that complicate the analysis of Besemah high vowels; the so-called *phantom* phoneme /ɣ/ is the other (see 2.1.1.2). Essentially, these two processes create the illusion that mid vowels occur in phonetically unconditioned environments, such as open syllables. The next two sections illustrate that these two processes demonstrate principled accounts of the allophonic variation of high vowels.

2.1.1.1 Vowel Harmony in High Vowels. Vowel harmony is a common property of Western Austronesian languages, including Malayic languages such as Standard Malay (cf. Teoh 1988:25–36).[12] Vowel harmony only occurs in roots that have any combination of high vowels with a final closed syllable. Vowel harmony occurs when a high vowel in the penultimate position is lowered due to an already lowered high vowel in the ultimate syllable, which underwent lowering by the final coda consonant of the root. For instance, the final vowel /i/ in /bukit/ 'hill, jungle' is lowered to [ɪ] due to conditioning oral coda /t/, thus triggering a back mutation of the first syllable /u/, yielding [bʊkɪt]. That is, vowel harmony in reality only affects the height of high vowel phonemes /i/ and /u/ in open penultimate syllables. The lowered final vowel resulting from the closed syllable is then the so-called *trigger* of vowel harmony.

Besemah vowel harmony generally demonstrates the following three phonetic characteristics. First, vowel harmony is not prominent or does not surface in words pronounced in isolation. For instance, /busuk/ 'rotten smell' is pronounced [busoʔ] in isolation, but in sentential contexts it is often pronounced [bʊsoʔ] or even [bosoʔ]. Second, the height of the penultimate vowel is relative to the height of the ultimate vowel; the lower the ultimate vowel is, the lower the penultimate vowel will be. For example, the final /i/ in /gigit/ 'tooth' surfaces as [ɪ] because it is followed by an oral conditioning consonant, yielding [gɪgɪt] and not *[gegɪt], while the final /i/ in /dirik/ 'hoist' surfaces as [e] because it is followed by a glottal conditioning coda, yielding [dɪreʔ] or [dereʔ]. Similarly, the penultimate vowel is generally slightly higher than the ultimate vowel; the final [o] in /busuk/ [bʊsoʔ] 'rotten smell' is lower than the first vowel [ʊ]. The first vowel may surface as [o] as well, yielding [bosoʔ], but the first vowel is still impressionistically higher. Finally, vowel harmony does not apply across the morpheme boundary. That is, the /i/ in the undergoer voice prefix *di-* does not lower; /dirik/ 'hoist' surfaces as [didɪreʔ] 'to be hoisted' when affixed with *di-*.

2.1.1.2 The Phantom Phoneme /ɣ/. The voiced velar fricative /ɣ/ is not realized in word-final position after the high back vowel /u/, but surfaces as an off-glide, as in /ʤmuɣ/ 'dry', which surfaces as [ʤəmʊw]. Even though the final coda does not surface, the phonemes are still lowered to [ʊ] and the /ɣ/ is subsequently realized as the off-glide [w]. However, upon the addition of a vowel-initial suffix *-i, -an,* or *-nya*, the phantom /ɣ/ is

11. Words in citation form usually do not evince vowel harmony, but still show vowel lowering in closed syllables. For example, in the cases of /gigit/ 'bite' and /pituŋ/ 'hold' above, the citation forms yield [gigɪt] and [pitʊŋ], respectively. In both examples, the closed final syllable is lowered from /i/ to [ɪ] and /u/ to [ʊ], while the vowel /i/ in penultimate syllable does not undergo vowel harmony and remains [i]. Word forms in all other sections of this article follow the citation form pronunciation, unless specified otherwise.

12. An anonymous reviewer points out that this is a common process in many Western Austronesian languages including Indonesian, Javanese, Madurese, and Balinese.

realized. For instance, /niuɣ/ is realized as [niyʊw] 'coconut', but as [niyʊɣan] 'coconut farm' when suffixed with *-an*.[13] Some speakers will occasionally pronounce the affixed root with an epenthetic [w], instead of the [ɣ], yielding [niyʊwan], but this appears to be a less common occurrence.[14] There are some roots that have already merged the /ɣ/ after /u/ with [ʔ], as in [tidoʔ] 'sleep' or [təloʔ] 'egg', which surface as [pənidoʔan] 'sleeping' upon the circumfixation of nominalizer *peN-...-an* or [təloʔan] 'many eggs' upon affixation of *-an*.

2.1.1.3 High Vowels in Loan Words. Recent loan words reveal a lot about the status of high vowels in Besemah. Words are borrowed from Standard Indonesian, Palembang Indonesian, English, and Javanese, but most of these appear to be taken via Standard Indonesian or Palembang Indonesian, even if the word is of a different origin. Loan words containing mid vowels from Standard Indonesian, including *sekolah* 'school', *merdeka* 'independent', and *dosen* 'professor', are most commonly pronounced with all high vowels, yielding [səkulah], [mərdikah], and [dusɪn]. In some cases, a word-final mid vowel is appended with [ʔ] or [h]. For example, *bule* 'foreigner' and *America* 'the United States' are realized as [buleʔ] and [amerikah], respectively. If there are mid vowels in the penultimate and ultimate syllable, the first vowel is realized as a high vowel, but the final vowel is often appended with /h/, as in the case of *soto* 'k.o. soup' from Javanese which is realized as [sutoh]. This, however, is not without variation and often depends on the speaker, the interlocutor, and the language situation. However, when the speech situation involves two speakers of Besemah, the pronunciation of mid vowels (nearly) always demonstrates the pronunciation above. This issue is taken up again in the discussion of contact and mid vowels in 3.2.

2.2 THE CENTRAL VOWEL(S). At first blush, Besemah appears to be a Malayic language that has a less restricted mid central vowel, as it can surface in the ultimate syllable. There are even a number of roots that contain all mid central vowels on the surface, such as [sələmə] 'a cold', [kələ] 'later', and [ʤəmə] 'person'.[15] From this, one might posit a mid central vowel phoneme /ə/. However, upon further investigation, it is clear that what appears to be the same mid central vowel is in reality two different entities: the phonetically inserted pepet (represented as [*ə*]) and the mid central allophone of the low central vowel /a/ (represented as [ə]). Essentially, Besemah has only one central vowel phoneme, the low central vowel /a/ with the allophones [a], [ə], and less commonly [ɨ], which has a different phonological status from the phonetically inserted pepet [*ə*]. From the phonetic form [ʤəmə] 'person', I posit a phonemic representation /ʤma/, by which the penultimate mid central vowel seg-

13. Upon the suffixation of *-an*, the final vowel /u/ in the root /niuɣ/ is still lowered to [ʊ], even in the derived form /niuɣan/ 'coconut farm' when the /u/ is in an open syllable [ni.yʊ.ɣan]. In other roots, the root final consonant (most commonly a stop consonant) forms a geminate when affixed with a vowel-initial suffix. For instance, /ambik/ 'take' is realized as [ambeʔ] in bare form, but upon suffixation of *-an*, the final glottal stop forms a geminate, yielding [am.beʔ.ʔan] 's.t. taken'.
14. The variation between the epenthetic [w] and the [ɣ] appears to be variable from speaker to speaker, but it appears that /ɣ/ surfaces far more often among older speakers.
15. The phonemic representations of these lexical items are as follows: /slma/ 'a cold', /kla/ 'later', and /ʤma/ 'person'. Reasons for these representations are given in 2.2.1.

ment is absent from the phonemic form and the final mid central vowel segment [ə] is raised in the root/word-final position from /a/ to [ə]. In the sections that follow, I first provide the motivation for treating the mid central vowel sounds as separate entities, the mid central allophone [ə] and the pepet [*ə*]. Subsequently, I provide evidence that the mid central allophone [ə] is in turn the allophone of the phoneme /a/.

2.2.1 Distinguishing the Pepet and Mid Central Allophone. This section puts forth the claim that the mid central vowel segments in roots such as [kələ] 'later' are in fact two distinct entities, despite their apparent overlap in the vowel space. The first vowel is the phonetically inserted pepet [*ə*] and the second vowel is an allophone [ə] of the low central vowel phoneme /a/, yielding a phonemic form /kla/ 'later'. This claim is based on four characteristics of the pepet and mid central allophone that warrant a distinction between the two.

First, the pepet only surfaces as a mid central vowel, while the mid central allophone can surface as [ə] or [ɨ], the latter usually occurring in a stressed utterance-final syllable.[16] For example, /ʤma/ 'people' normally surfaces as [ʤəmə] and occasionally as [ʤəmɨ] in the utterance-final syllable, but neither *[ʤɨmə] nor *[ʤɨmɨ] are possible.

Second, the pepet never receives any relative prominence in the word. In fact, the pepet is consistently shorter than all other vowels and is commonly not realized in stop-liquid sequences or word-initially before a nasal or a nasal-stop sequence. For instance, /bliŋ/ 'glass shards' can surface as [bəliŋ] with the pepet or [bliŋ] without the pepet, /mas/ 'gold' as [*ə*mas] with the pepet or [mas] without the pepet, and /mbau/ 'smell' as [*ə*mbaw] with the pepet or [m̩baw] with a syllabic nasal.[17] In the same manner, the pepet is also frequently unrealized in certain morphophonemic processes. The pepet that intervenes between the verbal prefixes *b-* and *t-* and a consonant-initial root is commonly not realized before a vowel-initial root. For example, /capak/ 'discard' is affixed with *t-*, yielding a [tə*ə*capaʔ] 'discard (accidently)', but when *t-* is affixed to a vowel-initial root, an intervening pepet does not always surface, and in the case of /umban/ 'fall' and /iŋat/ 'remember' the pepet commonly is not realized, yielding [tʊmban] 'fall (accidentally)' and [tiŋat] 'remember' with the verbal affix *t-*.[18] In other cases, the pepet is not realized in a root with an initial nasal-stop sequence when it is affixed with the vowel-final prefix *di-*, the undergoer voice marker. Therefore, the root /nʤuk/ 'give' is realized as [*ə*nʤoʔ] in its bare form, but is commonly realized as [diʔ*ə*nʤoʔ] 'to be given' with the pepet or [dinʤoʔ] without the pepet. The word-final mid central allophone [ə], on the other hand, is never deleted upon affixation. Therefore, /niaga/ 'to conduct business' is realized as [niyagə] in its bare form and as [niyagəɣan] 'business conducted' when affixed with the nominalizer *-an* (see 2.2.2 below for more discussion of affixation of mid central allophone final roots).

16. Besemah does not have word-level stress. Rather Besemah evinces utterance-final stress as has been found in other Malayic languages, Indonesian, and Betawi (cf. van Zanten and van Heuven 1998; van Heuven, Roosman, and van Zanten 2008). However, the details concerning stress in Besemah are not yet clear.
17. Note that [*ə*mbaw] here is an unaffixed root. The verbal form of [*ə*mbaw] is [ŋ*ə*mbawi] 'smell s.t'.
18. The prefixes *t-* and *b-* most commonly surface with an epenthetic voiced velar fricative [ɣ] or a glottal stop [ʔ] intervening before the vowel roots /umban/ 'fall' and /iŋat/ 'remember', yielding [tə*ʔ*ʊmban] or [tə*ɣ*ʊmban] and [tə*ʔ*iŋat] or [tə*ɣ*iŋat]. It is important to note that the epenthetic voiced velar fricative [ɣ] also follows the mid central allophone [ə] when followed by vowel-initial suffixes, such as the nominal suffix *-an* in roots appended to the root /kaua/ 'coffee', which is [kawə] in its bare form and [kawə*ɣ*an] 'coffee field' in its derived form.

Third, the distribution of the pepet is predictable; it is restricted to open syllables between consonants or preceding a nasal (in bisyllabic roots) or a nasal-stop cluster in root/word initial position. Conversely, the mid central allophone is not predictable in the same manner. The mid central allophone is only predictable insofar as it always occurs in open syllables word-finally. This is actually a property of its status as an allophone of /a/, which is discussed in the following section.

Fourth, the pepet shows minimal contrast with /a/, while there are no cases in which the mid central allophone [ə] shows minimal contrast with /a/; /raga/ [ragə] 'basket' contrasts with /rga/ [rəgə] 'price' and /baliŋ/ [balɪŋ] 'behind' contrasts with /bliŋ/ [bəlɪŋ] 'glass shard'. Only near minimal contrasts exist for [ə] and [a]: [katə] 'word' and [katah] 'many'.

There are, however, certain characteristics of the mid central vowel segments that would challenge treating the pepet and mid central allophone as separate segments. First, these mid central segments are in complementary distribution; the mid central allophone [ə] occurs in final position and can follow a vowel in a vowel sequence, while the pepet is barred word-finally and only occurs between two consonant phonemes, shown in (3) below. It is important to note, however, that there is a gap in the distribution of the mid central vowel segments. Mid vowels do not occur in the first vowel of a vowel sequence; there are no examples of [əu], [əi], or [əa] in Besemah. Conversely, /a, i, u/ all occur as the first vowel in a vowel sequence; /dai/ [dayi] 'face', /tiaŋ/ [tiyaŋ] 'pillar', and /ɣuas/ [ɣuwas] 'young bamboo'.

(3)

SYLLABLE	PEPET [*ə*]		ALLOPHONE [ə]
ultimate (open)	—		[tuwapə]'what'
penultimate - *C__C*	[**gə**daŋ]	'strong'	—
word-initial - *__NC*	[**ənt**uwə]	'parents-in-law'	—
antepenultimate	[**ʧə**nilah]	'sandal'	—

Furthermore, the mid central allophone and the pepet show considerable overlap in the vowel space; both sounds most commonly surface phonetically as a mid central vowel. Even though the distinction between the mid central allophone and the pepet cannot be based on phonetic or contrastive evidence, they are clearly phonologically separate entities based on the evidence of the (non)occurrence of each segment, the phonetic length, and the distribution of each segment. Ultimately, the distinction between these two segments is due to the fact that the mid central allophone [ə] is in reality an allophone of /a/, overlapping in the phonetic space of the epenthetic segment. The next section turns to this issue.

2.2.2 Establishing Final [ə] as an Allophone of /a/. The root/word-final [ə] is an earmark of Besemah phonetic representations. Speakers of Besemah commonly say that the main difference between Besemah and Standard Indonesian is that there is an [ə] at the end of a word instead of an [a], most commonly citing the Standard Indonesian phrase *ke mana* 'to where' as [kəmanə] in Besemah.[19] This section, however, sets out to demonstrate that this distinctive root/word-final mid central vowel is actually a case of allophonic variation rather than a phonemic contrast. Essentially, the /a/ raises to [ə] root/word-finally; /tuapa/ surfaces as [tuwapə] 'what'.

19. Interestingly, speakers of Besemah almost never comment on the fact the mid vowels /e/ and /o/ do not really occur in the language.

This analysis seems abstract because the allophone [a] of the phoneme /a/ rarely surfaces root/word-finally. Rather, the mid central allophone [ə] surfaces in roots and affixes in either the root and/or word-final position in nearly every case. In the case of the third person singular suffix *-nye* /ɲa/, the cognate of the Standard Indonesian *-nya*, the final vowel is realized as [ɲə̃] after vowel-final roots, as in /badʒu/ 'clothes', surfacing as [badʒuɲə̃] 'his/her/the clothes', or simply as a nasalized mid central allophone [ə̃] after consonant final roots, including the final /h/; /kakaŋ/ 'older sibling' and /ɣumah/ 'house' surface as [kakaŋə̃] 'his/her/the older sibling' and [ɣumahə̃] 'his/her/the house', respectively. In root-final position, the [ə] surfaces in bare forms and in affixed forms followed by an epenthetic [ɣ]. For instance, /kaua/ 'coffee' surfaces as [kawə] 'coffee' in its bare form and [kawəɣan] 'coffee plantation' when affixed with *-an*. In all cases of affixation, the final vowel phoneme /a/ surfaces as [ə] in the root-final position. Affixation can also be extended to reduplication, as in /gala/ 'all', which still surfaces as [galə-galəɲə̃] 'everything' with the addition of *-nye*. It is important to note that it is much more common to use partial reduplication, yielding [gəgaləɲə̃] 'everything'.

In spite of this distribution, there is evidence that the word-final [ə] is indeed an allophone of /a/. First, the distribution of the low back vowel /a/ evinces a gap resulting in a lack of minimal contrast between /a/ and [ə]. The low back vowel /a/ occurs in ultimate and penultimate positions within the root; no examples show the back vowel occurring in the antepenultimate syllable. Phonetically, lexical items with a final /ah/ sequence occasionally end in the back vowel [a]; /laŋkah/ 'step' can surface as [laŋkah] with the final [h] or [laŋka] without the final [h]. However, phonologically, it appears that all apparent [a]-final roots occur with the glottal fricative /h/.[20] The glottal fricative /h/, for the most part, surfaces unpredictably unless it is in careful speech, stressed utterance-final syllables, or in the suffixation of the vowel-initial suffixes *-an*, *-i,* or *-nye*; /laŋkah/ with affixation of the applicative suffix *-i* always surfaces with the root-final /h/ intact, as in [laŋkahi] 'to step over s.t.'. This, then, establishes that there are no truly [a]-final roots, affixes, or words. The lack of [a]-final roots supports an analysis in which the phoneme /a/ in the root/word-final position raises to the allophone [ə].

Further support for this position is provided by compounds and loan words. Compounds, including numeral compounds, demonstrate the only phonetic realization of [a] in roots that normally surface as [ə] in root-final position. That is, roots, such as /katʃa/ 'glass' and /mata/ 'eyes', surface in their bare forms as [katʃə] and [matə], but when compounded surface as [katʃamatə] 'glasses', not *[katʃəmatə]. Furthermore, numerals including /dua/ 'two' and /tiga/ 'three' surface as [duwə] and [tigə] in their bare forms, but when the numeral is compounded with /blas/ [bəlas] 'teen' or /puluh/ [puloh] 'ten', the numeral compounds surface as [duabəlas] 'twelve' and [tigapuloh] 'thirty', respectively. This is significant in light of other Malayic varieties, such as Palembang Malay, which yield a final [o] (analogous to [ə]) in numerals [duo] 'two' and [tigo] 'three' in bare forms and in numeral compounds [duobəlas] 'twelve' and [tigobəlas] 'thirteen'. This means that the final [o] in Palembang Malay may in fact be a phoneme /o/, yielding a phonemic form /duo/ 'two' and /tigo/ 'three'.

20. Knowledge of the presence of /h/ is very intuitive for speakers of Besemah and appears to be quite consistent from speaker to speaker. However, this distinction is not easily differentiated by nonnative speakers of Besemah.

The final piece of evidence comes from loan words ending in [a], which do not commonly undergo raising, but are appended with /h/. Such cases occur in loan words from Standard Indonesian; *ceria*, the name of an Indonesian cell phone provider, and *juta* 'million' are realized as [ʧəriyah] and [ʤutah] or [ʤutahan] 'millions' when affixed with *-an*. The treatment of [a]-final loan words demonstrates that there is a strong restriction on the occurrence of [a] in the final position of the root or the word, thus supporting an analysis that considers the word-final [ə] to be an allophone of /a/.

2.3 DIPHTHONGS. Besemah evinces three diphthongs [ay], [aw], and [ʊy], the latter being restricted to a small number of roots, such as /smɣui/ [səməɣʊy] 'messy hair' and the exclamative /ui/ [ʊy]. The diphthongs in Besemah are analyzed as phonemic sequences of vowels /ai/, /au/, and /ui/. In the majority of cases, this involves the low central vowel /a/ and either the high front vowel /i/ or the high back vowel /u/, which in the final syllable diphthongizes, phonetically emerging as a consonant. Diphthongs in Besemah are considered vowel sequences for the following reasons. First, diphthongs only occur in word-final position and almost exclusively in disyllabic words, (i.e., [kərbay] 'female' and [kidaw] 'left'). Roots that contain an /a/ + a high vowel sequence in a nonfinal position are most always two syllables. The only two examples collected thus far are [tawu] 'to be able' and [dayi] 'face'. When the exclamatives /ui/ and /au/ are stressed in an utterance-final position, the final high vowel is lengthened yielding the pronunciation [uyi:] and [awu:], respectively.[21]

3. DISCUSSION. A description of Besemah vowels would not be complete without some discussion of (1) Besemah high vowel correlates in Proto-Malayic and Standard Indonesian, especially roots that show mid vowels; (2) the current contact situation that appears to be pressuring the current three-vowel phoneme inventory of Besemah; and (3) the implications of the analysis in section 2 for other Malayic languages of Sumatra. In what follows, there is a brief note on the correspondence between Proto-Malayic and Besemah high vowels, maintaining that in almost all cognates Besemah retains the Proto-Malayic high vowels. Then, after a brief outline of the sociolinguistic situation in Besemah, there is a discussion that highlights points of increasing contact with Standard Indonesian, Palembang Indonesian, and Jakarta Indonesian, all of which appear to be pressuring Besemah to develop mid vowel phonemes. Even though there is no conclusive assessment of the future status of mid vowels, a future high vowel split into high and mid vowels in Besemah seems apparent. Finally, I conclude the discussion with the implications of the current analysis of Besemah regarding the areal distribution and permeation of mid vowels in the Malayic languages of Sumatra.

21. Further evidence for the lack of glide consonants in the language comes from the morphophonology of certain supposedly [w]-initial roots. When a consonant-initial root /bataŋ/ [bataŋ] 'stick' is affixed with *s-*, a pepet is inserted between the consonant cluster, yielding [səbataŋ] 'one/a stick'. However, when a [w]-initial root, [waraŋ] 'father-in-law' is affixed with *s-*, the pepet does not surface, but a high back vowel surfaces instead, yielding [suwaraŋan] 'the relationship between fathers of a married couple'. Therefore, the phonemic form of [waraŋ] is /uaraŋ/. The morphophonemic process by which a pepet vowel is not realized is strictly a characteristic of vowel-initial roots and never occurs with consonant-initial roots.

3.1 A NOTE ON THE RETENTION OF HIGH VOWELS. Many scholars have asked whether Malayic languages made an initial split of high vowels into mid and high vowels and at a later time re-merged these into high vowels, or if high vowels split in some dialects, but not others (see Adelaar 1992:45–46, Collins 1986b, Prentice 1994, and Prentice and Usman 1978). The overall consensus is the latter, and Besemah shows no evidence to the contrary. In fact, Besemah can add a Malayic variety of Sumatra (that is currently spoken) to the number of Malayic isolects that have attested a retention of high vowels, supporting the claim that the development of mid vowels in many Malayic dialects was not autochthonous. Comparing Standard Indonesian and Besemah, mid vowels show correlates to both high and, in one case, low vowels in Besemah, as is shown in (4).

(4)

BESEMAH	STANDARD INDONESIAN	PROTO-MALAYIC	
[ikoʔ]	[ekor]	*ikur	'tail'
[pandaʔ]	[pendeʔ]	*pandak	'short'
[libaɣ]	[lebar]	*libar	'wide'
[uɣaŋ]	[oraŋ]	*uraŋ	'person'
[biloʔ]	[beloʔ]	*biluk	'turn'

As might be expected from the vowel inventories of Besemah and Proto-Malayic, the same correlate high vowels are shared in Besemah and in many of the Proto-Malayic reconstructions in Adelaar (1992). In fact, out of the 191 reconstructed words given in Adelaar (1992), 138 of these contain one root with at least one high vowel; of these 138, 111 Besemah roots show matching high vowel correlates. Appendix 2 presents the 200-item wordlist from Adelaar (1992) for Besemah, highlighting high vowel correlates of Besemah and Proto-Malayic. The lexical items containing high vowel correlates show no evidence that Besemah split high vowels and subsequently remerged these high and mid vowels. Rather, the Besemah wordlist demonstrates that Besemah vowel phonemes are exceptionally conservative, compared to the other Malayic isolects of Sumatra (see 3.3 for more discussion on this point).

3.2 CONTACT AND THE RISE OF MID VOWELS. The Besemah appear to have long had contact with the downriver city of Palembang, the historic home of the maritime empire of Srivijaya, and wider Southeast Asia. The most well-known example of this contact is seen in the prehistoric megalithic stone monuments scattered throughout the Besemah plateau and a long tradition of a sung oral epic, or *guritan* (cf. van der Hoop 1932, Collins 1998). These remnants of the past demonstrate long contact with languages such as Javanese, in the case of the *guritan*, and wider Southeast Asia in the case of the megalithic stone monuments (cf. Collins 1979:13). Despite what appears to be sustained contact with languages that evince mid vowels, Besemah, at least for the time being, has resisted the emergence of mid vowel phonemes. However, currently this appears to be changing for some speakers of Besemah due to the increasing pressure from both standard and nonstandard varieties of Indonesian that already have well-developed mid vowels. After a cursory description of Besemah villages in the highlands of South Sumatra, issues relating to the rise of mid vowels in Besemah are treated in the following section.

3.2.1 Diglossia, Bilingualism, and the Possibility of Emerging Mid Vowels. Currently, many Besemah villages of the highlands of South Sumatra line the main roads that lead to major cities, such as Palembang and Bengkulu.[22] Most Besemah farm rice and coffee, which are subsequently sold at local markets or sent to major cities of southern Sumatra. Many Besemah work for a period of time in major cities of Indonesia or Malaysia, or send their children to secondary school or university in Bengkulu, Palembang, or several cities on Java. There appear to be only a handful of university educated Besemah who have returned to live in the villages, although many have university educated relatives living in cities of Sumatra and Java. The Besemah largely look towards the city of Palembang for commerce, work, and education. Many households now own at least one television, by which speakers are exposed to Standard Indonesian and Jakarta Indonesian on a daily basis. The city that is considered to be the center of Besemah culture is Pagar Alam,[23] where a koine is spoken that most closely represents Palembang Indonesian with influence from Besemah and other highland Central Malay dialects. Pagar Alam sees far more outsiders, including ethnically Chinese Indonesians, Minangkabau, and Javanese. Merchants from Java and other areas of Sumatra circulate through the villages to sell goods at weekly markets held in prominent villages. There are some households in which a Besemah has married someone from a neighboring ethnolinguistic group, most commonly from the neighboring Lintang, but occasionally from other ethnic groups of Sumatra. There are, of course, Javanese transmigrant villages as well, mostly working at the tea plantations started by the Dutch at the base of South Sumatra's highest peak, Mount Dempo. Adjacent villages have more intense contact with Javanese, but there is little contact outside of this and most Javanese transmigrants learn and use Besemah outside of their village.

The Besemah speaking region represents a clear case of diglossia as first proposed by Ferguson (1959), but as noted by Sneddon (2003:520) for other Indonesian varieties, this diglossia appears to be on a continuum with Standard Indonesian on one end, Besemah on the other, with Palembang Indonesian and the koine of Pagar Alam somewhere in the middle. In other words, Standard Indonesian is the acrolect, Palembang Indonesian is the mesolect, and Besemah is the basilect. In formal ceremonies, such as local elections, weddings, and the Friday sermons at the Mosque, Standard Indonesian (with formulaic phrases from Arabic, in the case of the latter two) is primarily used by those officiating the events, although it is not unusual for code-switching to occur between Besemah and Standard Indonesian.[24] In all other "everyday" circumstances, Besemah is strictly used (i.e., when talking to friends while playing a game, bargaining at the market, or arguing with a neighbor). This disparity is clear at wedding celebrations, where the wedding ceremony is conducted almost entirely in Standard Indonesian, but any other conversation during or after the ceremony is most certainly conducted in Besemah. When speaking to

22. See Sakai (2006:39) for a similar characterization of the villages of the Gumai, another highland group of South Sumatra. For a more complete description of the Besemah highlands, including contact and attitudes with other ethnic groups of South Sumatra, see Collins (1979:55–61). However, one should note that much has changed in these villages since the fieldwork of the anthropologist William Collins from 1971–73, including greater accessibility to other villages and cities due to paved roads and access to technology.
23. Pagar Alam was only recently recognized as a city, in 1996.
24. Apparently, this is true of education as well, but I have not seen a "natural" school environment.

outsiders, speakers of Besemah normally resort to the mesolect Palembang Indonesian (marked with the word-final [o]), and in some cases of usually younger speakers, Standard Indonesian or Jakarta Indonesian is used.[25]

Although the level of bilingualism in Standard Indonesian is not clear, there is a sharp contrast between older speakers of Besemah (approximately 50–60 years of age and older), who are largely monolingual, and younger speakers of Besemah who show varying abilities in Standard Indonesian, Jakarta Indonesian, and Palembang Indonesian, utilizing an increasingly large amount of vocabulary from these varieties of Indonesian.[26] Speakers who represent the latter group also show some variation, but this appears to be dependent on other factors, such as level of education, economic status, and contact with outsiders. Nevertheless, age is the most striking of these sociolinguistic variables and the simplest to pinpoint. For this reason, when referring to sociolinguistic factors in this section, age is the primary factor under consideration. However, the other factors listed above do appear to be important indicators of mid vowel use in regards to younger Besemah speakers.

Even though the lack of mid vowels is rarely if ever noted by speakers of Besemah, it is apparent to outsiders, especially in loan words from Standard Indonesian. As discussed in 2.1.1.3, mid vowels in loan words show some variation depending on the sociolinguistic situation, although it generally appears that when speakers of Besemah are conversing, a mid vowel in a loan word will surface as a high vowel in open syllables. However, loan words also show variation in the age of the speaker. There are two examples that demonstrate this variation clearly. The first is that of so-called minimal pairs between a loan word and an inherited lexical item. Such is the case of the derived forms [ŋəbʊn] 'to make a garden' from [kəbʊn] 'garden' and [ŋəbon] 'to pay a debt' from the Dutch borrowing *bon* 'bill', as well as in the case of [kalʊŋ] 'necklace' and [kaloŋ] 'fruit bat' from Javanese.[27] For older speakers, there is no phonetic difference between these pairs; all examples of /u/ above surface as [ʊ], meaning these pairs are actually homophones. On the other hand, there is greater variation among younger speakers, where some speakers contrast these based on the phonetic difference between the [ʊ] and [o] and others do not. The second example of possibly emerging mid vowels comes from the pronunciation of proper names. Although proper names are not usually considered loan words, in many cases, Besemah borrows various proper names, such as *Amanda* or *Ahmed*. In the case of *Hendi,* younger speakers most commonly pronounce the name [hendi], while older speakers pronounce [ɪndi], which also lacks the initial [h].[28] These examples indicate a shift in younger speakers who now demonstrate an ability to differentiate mid and high vowels in loan words, which is not the case among older speakers.

25. The use of Palembang Indonesian is not restricted to use when speaking with outsiders from Indonesia, but many (usually older speakers) commonly resort to Palembang Indonesian, even when addressing a foreigner from outside Indonesia who uses Standard Indonesian or even Besemah.
26. Children under the age of six who have not yet had formal schooling do appear to have a limited ability in Standard Indonesian. However, they appear to show increasing ability in colloquial Indonesian, i.e., Jakarta Indonesian, mostly learned from watching television.
27. The prefix [ŋə] is appended to single syllable words, such as *bon* 'bill'.
28. It is important to note that this is variable depending on the speaker's level of education and contact with outsiders, where some younger speakers also cannot produce the lower [e] or [o] in the penultimate syllable.

A possible exception to the emergence of mid vowels being a recent phenomenon comes from Helfrich (1904), which lists ten roots for Besemah and Seraway with a mid back vowel in the word-initial (or the penultimate syllable) position. The presence of the mid back vowel points to the possibility of an earlier emergence of mid vowels in loan words before intense contact with Standard Indonesian. The eleven roots from Helfrich (1904) are as follows (with the Standard Indonesian equivalents in italics):

(5)

[obah]	'change'	*ubah*	[omoŋ]	'talk'	*omong*
[oban]	'grey hair'	*uban*	[oɣaŋ]	'person'	*orang*
[obat]	'medicine'	*obat*	[oroŋ]	'cancel'	*urung*
[otaʔ]	'brain, mind'	*otak*	[ombaʔ]	'wave'	*ombak*
[oloʔ]	'ridicule'	*olok*	[ombar]	'let loose'	*umbar*
[oŋgol]	'superior, distinguished'	*unggul*			

Helfrich only notes one of these roots as having a variant, [ombaʔ] as [umbaʔ] 'wave'.[29] Interestingly, five of these roots have a corresponding *u* in the penultimate syllable of the Standard Indonesian lexical item, which could be explained in two cases by vowel lowering in closed syllables, [ombar] 'let loose' and [oŋgol] 'superior, distinguished'. In the three remaining cases, the reason for Helfrich's transcriptions is unclear. However, in the Besemah highlands a century later, the above roots are still pronounced with a [u] in an open syllable or [ʊ] in a closed syllable. The only exception to this is [oloʔ] 'ridicule', which was not found at all. The most plausible explanation for this disparity between the Besemah of the nineteenth century and the Besemah currently spoken is dialectal difference, as Helfrich appears to have collected Besemah data closer to the coast in the lowlands of West Sumatra, in the city of Manna (see map 1).

Even though this is in no way a comprehensive discussion of the emergence of mid vowels in Besemah and/or the factors that correspond to this change, it does appear from a cursory analysis that more recent intensive contact with Standard Indonesian, Palembang Indonesian, and Jakarta Indonesian through universal education, increasing access to language through media, and increasing mobility are affecting mainly younger, usually more educated speakers of Besemah in their use of mid vowels. This, in turn, creates a sharp divide between older and younger speakers, the prior of which had limited exposure to the external influences listed above.

3.3 IMPLICATIONS FOR OTHER MALAYIC LANGUAGES OF SUMATRA. The Besemah vowel phoneme inventory (re-)raises some important issues regarding the vowels of Malayic languages of Sumatra. These issues are split into the following two questions for further study:

a. What is the status of various Proto-Malayic reflexes of *-a in the Malayic languages of Sumatra? That is, for any given language, are vowel final segments allophones (as in Besemah) or have these final segments become phonemes (as in Palembang Malay)?

29. In Helfrich's orthography, he distinguishes <oe>, <*o*>, and <o>, which correspond to [u], [ʊ], and [o], respectively. There is no instance of the mid-vowel [e], but he does distinguish <i> and <*i*>, which correspond to [i] and [ɪ], respectively.

b. Is there a correlation between linguistic contact and the spread of mid vowels in the Malayic languages of Sumatra? That is, is there any evidence that mid vowels have been diffused through contact with other Malayic languages of Sumatra?

These questions are briefly discussed with the hope that the value of Besemah and possibly other undescribed Malayic varieties is appreciated for its ability to shed light on the development and status of vowels in Malayic languages of Sumatra. However, this discussion is nowhere near a full treatment of the questions above, nor is it meant to be.

Although the issue of the Proto-Malayic *-a was not explicitly taken up in the above description, Besemah shows important implications for the treatment of the reflexes of Proto-Malayic *-a in the Malayic languages of Sumatra. That is to say that Besemah evinces what appears to be a final reflex -[ə], which in reality turns out to be an allophonic variant of what is still -[a] (as was argued in 2.2.2). Adelaar (2005:206) briefly states that the reflex of Proto-Malayic *-a may perhaps be an areal trend that surfaces in western Indonesian languages. Throughout much of the eastern lowlands and western coast of central and parts of southern Sumatra including Palembang, Jambi, and Minangkabau, the *-a commonly surfaces as -[o] or in some cases -[ə] (cf. Tadmor 2003, Anderbeck 2003:44–45).

In Central Malay languages, this is slightly more varied where final vowels that correspond to *-a are -[aw] in southern Seraway, -[o] in northern Seraway and Lintang, and -[ə] in Besemah and Semende. In the case of Lintang, the phonemic status of final -[o] appears to be the same as Besemah: -[o] is an allophone of a phonemic -/a/. From an initial investigation, Lintang shows similar characteristics in the mid vowels of loan words (i.e., *motor* is [mutɔr] 'motorcycle'), and in the case of numeral compounds the final vowel remains a low-central vowel, (i.e., 'two' is [duo] from /dua/, but 'twelve' is [duabəlas] from /dua/ 'two' and /blas/ 'teen'). Besemah and possibly Lintang, then, raise questions about the reality of the phonemic status of the word-final *-a reflex in other Malayic languages of Sumatra. As mentioned in 2.2.2, Palembang Malay, which has mid vowel phonemes, shows a different pattern for the word final -[o] in compounds: [duo] 'two' from /duo/ still surfaces as [duobəlas] 'twelve' from /duo/ 'two' and /blas/ 'teen'. Based on these isolects, it appears that some Malayic isolects have phonemicized the *-a reflex, while others have not. However, this issue needs to be more fully described, tested, and weighed against other phonological factors such as symmetry in the vowel phoneme inventory in other Malayic varieties (cf. Anderbeck 2003:39–40), as has been done here in the case of Besemah.

As was briefly pointed out in section 1, Malayic isolects differ in the status of their mid vowels, in which mid vowels range from "well-developed" to nonexistent. The Malayic languages of the eastern lowlands of Sumatra appear to have "well-developed" mid vowels that occur in the penultimate syllable (i.e., dialects of Jambi and Palembang Malay: see fn. 4 above). On the other hand, the Malayic languages on the west coast of Sumatra, including dialects of Minangkabau and Seraway, show "less-developed" mid-vowel phonemes in open syllables. Still another case exists with Kerinci, which maintains a restriction on mid vowels in the penultimate syllable, but shows significant variation in final vowels (Prentice and Usman 1978). From these still preliminary findings and those from Besemah above, Malayic isolects of Sumatra can be thought of as a continuum that highlights the vowel phonemes of fully developed mid vowels (i.e., dialects of

Jambi Malay and Palembang Malay), to partially developed mid vowel phonemes (i.e., Minangkabau and Seraway), to not at all developed mid vowel phonemes (i.e., Besemah, Kerinci, and possibly other Malayic varieties that are yet to be described).

Not surprisingly, this continuum corresponds to an areal feature by which the more isolated languages in the Barisan mountain range running vertically along the western side of Sumatra show features of more conservative vowel phoneme inventories, while those on the eastern coast show the most innovative, followed by less innovative languages of the western coast. However, this feature may not be restricted to the highlands, and may correlate also with more isolated locations in the Sumatran lowlands. From an initial investigation, Anderbeck (2003:40) posits that Malayic varieties of Jambi appear to be split into varieties located along the Batanghari River that appear to have mid vowels, and Malayic varieties not along the Batanghari River that appear without mid vowels.[30] These small shreds of evidence point to the possibility that mid vowels may have diffused along the coasts of Sumatra and along the rivers of the lowlands, and are only, in the present, making their way to the more isolated highland varieties, such as Besemah. However, due to a lack of reliable data on both coastal and highland varieties of Malayic languages in Sumatra, this question is left unanswered until more is known about some of the more isolated Malayic languages of Sumatra.

4. CONCLUSION. This paper has outlined the following: (1) a phonological description of Besemah vowels, which asserts that Besemah shows no evidence of mid vowel phonemes or a separate mid central vowel phoneme; (2) the current sociolinguistic situation in the Besemah highlands that shows how Besemah appears to be assimilating with other Malayic varieties that have already seen the emergence of mid vowels; and (3) the landscape of vowel phoneme inventories in some Malayic languages of Sumatra that point to a possible areal trend concerning the appearance of mid vowels in vowel phoneme inventories.

APPENDIX 1. CONSONANT PHONEMES IN BESEMAH

Besemah has 20 consonant phonemes as seen below.

		LABIAL	DENTAL	PALATAL	VELAR
STOPS	VOICELESS	/p/	/t/	/ʧ/	/k/
	VOICED	/b/	/d/	/ʤ/	/g/
FRICATIVES	VOICELESS		/s/		
	VOICED				/ɣ/
NASALS		/m/	/n/	/ɲ/	/ŋ/
TRILL			/r/		
LATERAL			/l/		

- Of the plain stops, only voiceless stops /p, t, k/ occur in the syllable coda, while /k/ is almost always realized as [ʔ] in the coda.

30. The Batanghari River is an important trade route that runs east from the Minangkabau highlands to the Jambi lowlands.

- The /ɣ/ contrasts with /r/, but appears in complementary distribution in some cases; /ɣ/ does not occur after the high back vowel /u/ (see 2.1.1 above). In some dialects of Besemah /ɣ/ surfaces as [h] intervocalically.
- /h/ never occurs word-initially; it occurs occasionally word-medially but commonly word-finally.
- Palatal sounds do not occur in the coda position.

APPENDIX 2. 200-ITEM WORDLIST FOR BESEMAH

The 200-item basic wordlist is based on Adelaar (1992:133–37). The Proto-Malayic reconstructions appear in the original form, but some of English and Standard Indonesian words have been slightly altered. The Besemah wordlist was collected by the author in the village of Karang Tanding, ten kilometers northwest of the city of Pagar Alam (see fn. 2). Adelaar (1992:98) notes that "The function of *ø is to indicate that in its place there was no PM [Proto-Malayic] phoneme, although it is likely that there was a (nonphonemic) glottal stop … ." The Besemah words are written in phonetic form in IPA format, expect that [y] represents the palatal glide. All Proto-Malayic words containing a high vowel are shaded, while only the Besemah lexemes that have a corresponding high vowel form in Proto-Malayic are shaded.

	ENGLISH	INDONESIAN	PROTO-MALAYIC	BESEMAH
1.	hand	tangan	*taŋan	taŋan
2.	left	kiri	*kA-iri/*kibaʔ	kidaw
3.	right	kanan	*k/anan	kanan
4.	leg/foot	kaki	*kaki	kətɪŋ
5.	walk	(ber-)jalan	*((mb)Ar-)jalan	(bə-)ʤalan
6.	road/path	jalan	*jalan	ʤalan
7.	come	dating	*datəŋ	dataŋ
8.	turn (v.)	belok	*biluk	biloʔ
9.	swim	(be-)renang	(*(mb)A-rənaŋ)	(bə-)kayoh
10.	dirty	kotor	(*kamah/*kumuh)	kulat
11.	dust	debu	*dəbu	dəbu
12.	skin	kulit	*kulit	kulɪt
13.	back	belakang, punggung	*bAlakaŋ	bəlakaŋ, pɪŋgaŋ
14.	belly	perut	*pərut	pəɣʊt, busʊŋ
15.	bone	tulang	*tulaŋ	tulaŋ
16.	intestines (guts)	usus	*pərut	usʊs, pəɣʊt
17.	liver	hati	*hati	ati
18.	breast	susu	*susu(ʔ)	kipaʔ
19.	shoulder	bahu, pundak	—	bakɪɣ
20.	know (things)	tahu	*tahu(ʔ)	kəruwan, tawu
21.	think	(ber-)pikir	—	(bə-)pikir
22.	fear	takut	*takut	takʊt
23.	blood	darah	*darah	daɣah
24.	head	kepala	*hulu(ʔ)	palaʔ
25.	neck	léhér	*lihər	liyaɣ
26.	hair	rambut	*buø(uə)k, *rambut	gʊmbaʔ
27.	nose	hidung	*hiduŋ	idʊŋ

	ENGLISH	INDONESIAN	PROTO-MALAYIC	BESEMAH
28.	breathe	(ber-)nafas	*ñawa	(bə-)napas
29.	sniff, smell	cium	—	əmbaw
30.	mouth	mulut	*mulut	mulʊt
31.	tooth	gigi	*gigi	gigi
32.	tongue	lidah	*dilah	lidah
33.	laugh (v.)	(ter-)tawa	*tawaʔ	(tə-)tawə
34.	cry (v.)	tangis	*taŋis	taŋɪs
35.	vomit (v.)	muntah	*m/u(n)tah	mutah, (tə-)mutah
36.	spit	(ber-)ludah	*ludah	(bə-)liyʊw, liyʊ(ɣ)
37.	eat	makan	*ma/kan	makan
38.	chew	kunyah, mamah	*kuñah, *mamah	pəpaʔ
39.	cook (v.)	masak, tanak	*m/asak, *tanak	masak, tanaʔ
40.	drink (v.)	minum	*inum	minʊm
41.	bite (v.)	gigit	*gigit	gigɪt
42.	suck	hisap	*hi(ŋ)səp	isap
43.	ear	telinga	*tAliŋa(ʔ)	təliŋə, ʧupɪŋ
44.	hear	dengar	*dəŋər	dəŋaɣ
45.	eye	mata	*mata	matə
46.	see	lihat	*lihat	kinaʔ
47.	yawn (v.)	kuap	*uap	uwap
48.	sleep (v.)	tidur	*tidur	tidoʔ
49.	lie down	(ber-)baring	—	mulɪk
50.	dream (v.)	(ber-)mimpi	*m/impi/*impi	(bə-)mɪmpi
51.	sit (v.)	duduk	*duduk	dudoʔ
52.	stand (v.)	(ber-)diri	*diri	(bə-)təgoʔ
53.	human being	orang	(*uraŋ)	ʤəmə, uɣaŋ
54.	man/male	laki-laki	*laki(-laki)	lanaŋ
55.	woman/female	perempuan	—	bətinə
56.	child	anak	*anak	anaʔ, budaʔ
57.	husband	suami	*laki	laki
58.	wife	istri	*bini	bini
59.	mother	ibu	*(ə)ma(ʔ), *induʔ, *ina	umaʔ, əndʊŋ, ibʊŋ
60.	father	bapak	*apa(ʔ)	bapaŋ
61.	house	rumah	*rumah	ɣumah
62.	roof/thatch	atap	*hatəp	atap
63.	name	nama	—	damə
64.	say	(ber-)kata	*tutur	(bə-)katə
65.	rope	tali	*tali	tali
66.	tie up, fasten	ikat	*ikət	kəbat
67.	sew	jahit	*jahit	ʤayɪt
68.	needle	jarum	*jarum	ʤaɣʊm
69.	hunt (v.)	buru	*buru	buru
70.	shoot (arrow)	tembak	*timbak, *panah	tɪmbaʔ
71.	stab, pierce	tikam	*tikəm, *tusuk	tusoʔ
72.	hit, slap	pukul	*pukul, *paluʔ	badoʔ, pukʊl
73.	steal	curi	*maliŋ	malɪŋ

	ENGLISH	INDONESIAN	PROTO-MALAYIC	BESEMAH
74.	kill	bunuh	*bunuh	bunoh
75.	die, be dead	mati	*m/ati	mateʔ
76.	live, be alive	hidup	*hidup	idʊp
77.	scratch (an itch)	garuk	*garut, *garuk, *garuʔ	gaɣʊt
78.	cut, hack (v.)	potong	*tətək, *tarAs	tətaʔ, pʊntʊŋ
79.	stick (wood)	kayu	*kayuʔ	kayu, pʊntʊŋ
80.	split	belah	*bəlah	bəlah
81.	sharp	tajam	*tajəm	landap
82.	dull, blunt	tumpul	*tumpul	tʊmpʊl
83.	work (in field)	bekerja	*(mb)Ar-huma(ʔ)	(bə-)gaweh
84.	plant (v.)	tanam	*tanəm	tanam
85.	choose	pilih	*pilih	pileh
86.	grow	tumbuh	*t/um/buh	tʊmboh
87.	swell (abscess)	bengkak	*bəŋkak	nidap
88.	squeeze	peras	*pərəs, *pərah	pəɣah
89.	hold (in fist)	pegang	*pəgaŋ	pəgaŋ
90.	dig	gali	*kali	kadʒah
91.	buy	beli	*bəli	bəli
92.	open, uncover	buka	*bukaʔ	bukaʔ
93.	pound, beat	tumbuk	*tumbuk, *tutuk	tutoʔ
94.	throw	lempar	—	banah
95.	fall (v.)	jatuh	*jatuh, *labuh	umban
96.	dog	anjing	*asuʔ	aɲdʒɪŋ
97.	bird	burung	*buruŋ	burʊŋ
98.	egg	telur	*təlur	təloʔ
99.	feather	bulu	*bulu	bulu
100.	wing	sayap	*sayap	sayap
101	fly (v.)	terbang	*tAr(ə)baŋ	tərbaŋ
102.	rat	tikus	*tikus	tikʊs
103.	meat, flesh	daging	*dagiŋ, *isiʔ	dagɪŋ
104.	fat, grease	lemak	*ləmək	ləmaʔ
105.	tail	ekor	*ikur	ikoʔ
106.	snake	ular	*ulər	ulaɣ
107.	earthworm	cacing	*hulət, caciŋ	əŋgəlaŋ
108.	louse (head)	kutu	*kutu	kutu
109.	mosquito	nyamuk	*ñamuk	agas
110.	spider	labah-labah	*lawaʔ, laba[ʔ]	ləlawah
111.	fish	ikan	*ikan	ikan
112.	rotten	busuk	*busuk	busoʔ,† baŋay
113.	branch	dahan	*dahan	ɣantɪŋ
114.	leaf	daun	*daun	dawʊn
115.	root	akar	*akar	akaɣ
116.	flower	bunga	*buŋa(ʔ)	buŋə
117.	fruit	buah	*buah	buwah
118.	grass	rumput	*rumput	ɣʊmpʊt
119.	earth, soil	tanah	*tanah	tanah
120.	stone	batu	*batu	batu

	ENGLISH	INDONESIAN	PROTO-MALAYIC	BESEMAH
121.	sand	pasir	*pasir	pasır
122.	(fresh) water	air	*air	ayeʔ
123.	flow	alir	*alir	aɲʊt
124.	sea	laut	*tasik	lawʊt
125.	salt	garam	*sira	ɡaram
126.	lake	danau	*danaw	danaw
127.	woods, forest	hutan	*hutan, *rimbaʔ	utan
128.	sky	langit	*laŋit	laŋıt
129.	moon	bulan	*bulan	bulan
130.	star	bintang	*bintaŋ	bıntaŋ
131.	cloud	awan	*a(bw)an	aban
132.	fog	kabut	*kabut	kabʊt
133.	rain	hujan	*hujan	ujan
134.	thunder	guntur	*guntur, *guruh	guroh
135.	lightning	kilat	*kilat	kilat
136.	wind	angin	*aŋin	aŋın
137.	blow	(ber-)tiup	*t/iup	(bə-)tiyʊp
138.	warm	panas	*panas	panas, aŋat
139.	cold	dingin	*diŋin	diŋın
140.	dry	kering	*kəriŋ	kəɣıŋ
141.	wet	basah	*basah	basah
142.	heavy	berat	*bərat	beɣat
143.	fire	api	*api	api
144.	burn (v.)	bakar	*tunu, *bakar	ʤudʒʊl
145.	smoke	asap	*asəp	asap
146.	ash	abu	*habu	abu
147.	black	hitam	*hitəm	itam
148.	white	putih	*putih	puteh
149.	red	merah	*(ma-)irah	abaŋ
150.	yellow	kuning	*kunit	kunıŋ
151.	green	hijau	*hijaw	iʤaŋ
152.	small	kecil	*kəcil, *kətik	kəʃeʔ
153.	big	besar	*bəsar, *raya	bəsaʔ
154.	short	pendek	*pandak, *pindik	pandaʔ
155.	long	panjang	*pañjaŋ	paŋnʤaŋ
156.	thin	tipis	*nipis, *m/ipis	nipıs
157.	thick	tebal	*təbəl	təbal
158.	narrow	sempit	*səmpit	səmpıt, sʊmpaʔ
159.	wide	lebar	*libar	libaɣ
160.	painful, sick	sakit	*sakit	sakıt, gərıŋ
161.	shy, ashamed	malu	*malu	malu
162.	old	tua	*tuha(ʔ)	tuwə
163.	new	baru	*bAharu	baru, aŋar, əmpay
164.	good	baik	*baik	iloʔ
165.	bad, evil	jahat	*jahət	karoʔ
166.	correct, true	benar	*bənər	benaɣ
167.	night	malam	*ma-lə(hØ)əm	malam

	ENGLISH	INDONESIAN	PROTO-MALAYIC	BESEMAH
168.	day	hari	*hari	aɣi
169.	year	tahun	*tahun	tawʊn
170.	when?	kapan	—	kəbilə
171.	hide	sembunyi	*buni	(bə-)sımbʊn
172.	climb	naik	*naik	nayeʔ
173.	at	di	*di	di
174.	in, inside	(di) dalam	*(d-)aləm	di dalam
175.	on (top of)	(di) atas	*atas	di puʧoʔ
176.	below	(di) bawah	*bah, *babah	di bawah
177.	this	ini	*(i)ni(ʔ)	tini, ini, ni
178.	that	itu	*(i)tu(ʔ), *(i)na(n), *(a)na(ʔ)	titu, itu, tu
179.	near	dekat	*dəkət, *(h)ampiŋ	paraʔ, dampıŋ
180.	far	jauh	*jauh	jawoh
181.	where?	di mana	*-mana(ʔ)	di manə
182.	I	saya, aku	*aku	aku
183.	you (SG)	engkau, kamu	*kau(ʔ)	kabah
184.	he/she	dia, ia	*ia	diyə
185.	we (EXCL)	kami	*kitaʔ, *kami	kami
185.	we (INCL)	kita	—	kitə
186.	you (PL)	kamu, kalian	*kamu(ʔ)	kamu
187.	they	mereka	*sidaʔ	diyə
188.	what?	apa	*apa	tuwapə, apə
189.	who?	siapa	*si-apa, *sai	siyapə, sapə
190.	other	lain	*bukən	layın
191.	all	semua	*habis	ɢalə
192.	and	dan	—	ŋah
193.	if	kalau, jika	—	kalu, amu, amə
194.	how?	bagaimana	—	loʔ manə
195.	no, not	tidak	*-daʔ	didə
196.	count (v.)	hitung	*hituŋ	(mə-)rikın
197.	one	satu	*əsaʔ	sə-, suteh, sikoʔ
198.	two	dua	*dua(ʔ)	duwə
199.	three	tiga	*təlu	tigə
200.	four	empat	*əmpat	əmpat

† This [see item no. 112] is no longer used for ‘rotten’, but means ‘foul smell’.

REFERENCES

Adelaar, Alexander. 1992. *Proto-Malayic: The reconstruction of its phonology and parts of its lexicon and morphology.* Canberra: Pacific Linguistics.

———. 2005. Structural diversity in the Malayic subgroup. In *The Austronesian language of Asia and Madagascar*, ed. by Alexander Adelaar and Nikolaus P. Himmelmann, 202–26. London: Routledge.

Adelaar, Alexander, and D. J. Prentice. 1996. Malay: Its history, role, and spread. In *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*, ed. by S. A. Wurm, P. Mühlhäusler, and D. Tryon, 673–93. Berlin and New York: Mouton de Gruyter.

Aliana, S. A., S. Nursato, S. Arifin, S. Soetopo, and M. Waif. 1987. *Morfologi dan sintaksis Bahasa Melayu Palembang.* Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Anderbeck, Karl. 2003. *Malay dialects of the Batanghari river basin (Jambi, Sumatra).* Master's thesis. Bangi: Universiti Kebangsaan Malaysia.

Brandes, J. L. 1884. *Bijdrage tot de vergelijkende klankleer der westersche afdeeling van de Maleisch-Polynesische taalfamilie.* Utrecht: Stoomdrukkerij P. W. van de Weijer.

Cohn, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7:167–216.

Collins, James T. 1986b. Kajian dialek daerah dan rekonstruksi Bahasa Purba. *Dewan Bahasa* 30:344–65.

Collins, William. A. 1979. *Besemah concepts, a study of the culture of the people of South Sumatra.* PhD diss. Berkeley: University of California.

———.1998. *The guritan of Radin Suane: A study of the Besemah oral epic from South Sumatra.* Leiden: KITLV Press.

Ferguson, C. A. 1959. Diglossia. *Word* 15:325–40.

Gil, David. 1994. The structure of Riau Indonesian. *Nordic Journal of Linguistics* 17:179–200.

Gil, David, and Uri Tadmor. 1997. Towards a typology of Malay/Indonesian dialects. Paper presented at the *First Annual International Symposium on Malay and Indonesian Linguistics.* Penang: Universiti Sains Malaysia, January 14.

Helfrich, O. L. 1904. Bijdragen tot de kennis van het Midden Maleisch (Besemahsch en Serawaisch dialect). *Verhandelingen van het Bataviaasch Genootschap van kunsten en wetenschappen* 53.

———. 1915. Nadere aanvullingen en verbeteringen op de Bijdragen tot de kennis van het Midden Maleisch. *Verhandelingen van het Bataviaasch Genootschap van kunsten en wetenschappen* 61(3–4): 1–28.

———. 1921. Supplement op de Nadere aanvullingen en verbeteringen op de Bijdragen tot de kennis van het Midden Maleisch. *Verhandelingen van het Bataviaasch Genootschap van kunsten en wetenschappen* 63(3).

———. 1927. Nadere bijdragen tot de kennis van het Midden Maleisch. *Verhandelingen van het Bataviaasch Genootschap van kunsten en wetenschappen* 68(3).

———. 1933. Bijdragen tot de kennis van het Midden Maleisch (supplement to Helfrich 1927). The Hague: Nijhoff.

McDowell, Jonthan E. 2007. Bhinnêka tunggal ika: Unity in diversity, Malayic varieties of Southern Sumatra. *10th Annual International Symposium on Malay/Indonesian Linguistics.* Manokwari: Universitas Negeri Papua, August 6.

Mitani, Yasuyuki. 1980. Languages of South Sumatra. In *South Sumatra: Man and agriculture,* ed. by Yoshihiro Tsubouchi, 1–16. Kyoto: The Center for Southeast Asian Studies, Kyoto University.

Prentice, D. J. 1990. Malay (Indonesian and Malaysian). In *The major languages of East and Southeast Asia,* ed. by Bernard Comrie, 185–207. London: Routledge.

———. 1994. Manado Malay: Product and agent of language change. In *Language contact and change in the Austronesian world,* ed. by Tom Dutton and Darrell T. Tyron, 411–41. Berlin: Mouton de Gruyter.

Prentice, D. J., and Hakim Usman. 1978. Kerinci sound-changes and phonotactics. In *Second International Conference on Austronesian Linguistics: Proceedings,* fascicle 1, *Western Austronesian*, ed. by S. A. Wurm and Lois Carrington, 121–63. Canberra: Pacific Linguistics.

Sakai, Minako. 2006. The origin structure of Kute among the Gumai: An analysis of an indigenous territorial institution in the highlands of South Sumatra. In *Sharing the earth, dividing the land,* ed. by Thomas E. Reuter, 39–63. Canberra: Australia National University E- Press.

Sneddon, James. 2003. Diglossia in Indonesian. *Bijdragen tot de Taal-, Land- en Volkenkunde* 159(4): 519–49.

Tadmor, Uri. 2003. Final /a/ mutation: A borrowed areal feature in western Austronesian languages. In *Issues in Austronesian historical phonology,* ed. by John Lynch, 15–35. Canberra: Pacific Linguistics.

Teoh, Boon Seong. 1988. *Aspects of Malay phonology revisited—A non-linear approach.* PhD diss. Urbana-Champaign: University of Illinois.

van der Hoop, A. N. 1932. *Megalithic remains in South Sumatra*. Zutphen: W. J. Thieme & CIE.

van Heuven, Vincent J., Lilie Roosman, and Ellen van Zanten. 2008. Betawi Malay word prosody. *Lingua* 118: 1217–87.

van Zanten, Ellen, and Vincent J. van Heuven. 1998. Word stress in Indonesian: Its communicative relevance. *Bijdragen tot de Taal-, Land- en Volkenkunde.* 154(1):129–49

Voorhoeve, P. 1955. *Critical survey of studies on the languages of Sumatra.* The Hague: Nijhoff.

bradley.mcdonnell@gmail.com