CHAPTER 28

Non-Malayic languages of Sumatra and the Barrier Islands

Bradley McDonnell and Christina L. Truong

28.1 Introduction

This chapter covers the non-Malayic Malayo-Polynesian languages of Sumatra and the Barrier Islands (henceforth referred to as the non-Malayic languages of Sumatra or NMLS). This area is comprised of one apparent expansive subgroup (i.e. Barrier Islands-Batak, Nothofer 1986), several languages or language clusters that have not yet been found to form a larger subgroup with other Malayo-Polynesian languages (i.e. Lampung, Nasal, Rejang, Enggano, Gayo), and Acehnese, which subgroups with Chamic languages outside the region (see Brunelle, chapter 11, this volume).[1] NMLS are spoken in regions located on the periphery of Sumatra. This includes the north and south ends of Sumatra as well as the Barrier islands off the western coast of Sumatra (see Map 28.1). NMLS are extremely diverse. They share few typological features other than those due to their shared Austronesian inheritance, their general geographic position, and the contact they have all had with Malayic languages. The most divergent of the NMLS, Enggano, is perhaps the most aberrant Austronesian language altogether. Some scholars have even questioned its status as an Austronesian language (Capell 1982; see also Nothofer 1991: 394 for discussion), although now it is fairly widely accepted to be Austronesian (see Edwards 2015).

This chapter is organized as follows. §28.2 describes the consonant and vowel inventories, stress, and phonological processes. §28.3 presents an overview of common affixes and morphological processes in the languages. §28.4 covers basic syntactic properties including grammatical relations, case, agreement, word order, and noun phrase structure. §28.5 describes some aspects of tense, aspect, modality, and mood in NMLS. §28.6 summarises the chapter and describes directions for further research including the need for more documentation and description of NMLS. Unless mentioned otherwise, we draw on the following sources for our typological generalizations and examples: Acehnese (Durie 1985), Gayo (Eades 2005), Karo Batak (Woollams 1996), Toba Batak (Nababan 1981; Adelaar 1995d), Devayan (also called Simeulue; Kähler 1955), Sigulai (also called Sikule; Kähler 1955), Nias (Brown 2001), Mentawai (Morris 1900; Pampus 1989), Enggano (Kähler 1940; Crowley n.d.), Rejang (McGinn 1982, 2005), Nasal (McDonnell fieldnotes), and Lampung (Walker 1976).[2]

28.2 Phonology

NMLS present diverse phonological properties at almost every level. This section aims to describe these phonological systems, noting similarities and differences across languages in phonemic inventories and allophonic variation (§28.2.1), stress and prosodic prominence (§28.2.2). and phonological processes (§28.2.3).

28.2.1 Phoneme inventories

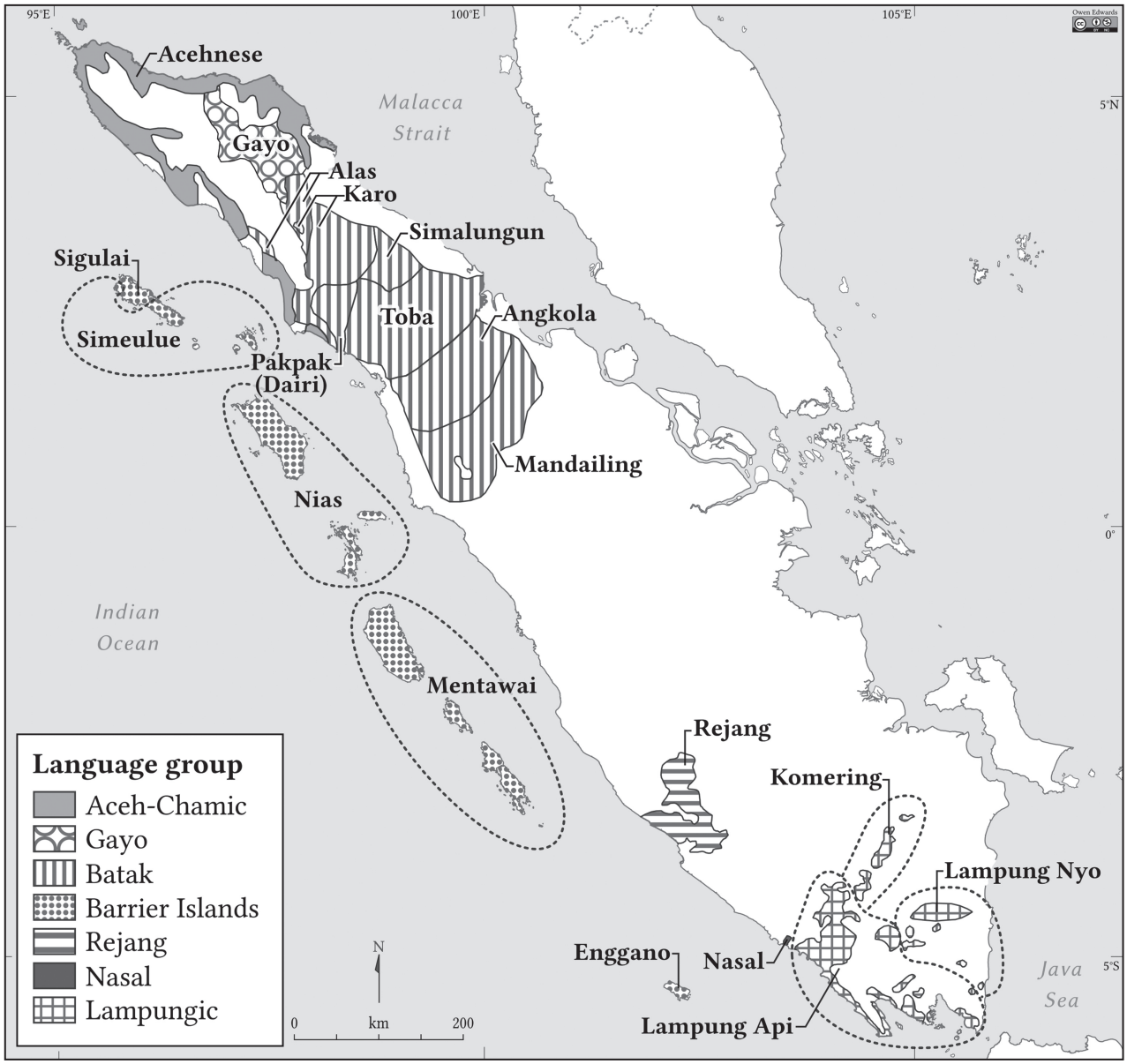
28.2.1.1 *Consonants*

The consonant inventories of the NMLS are comprised of anywhere from ten to twenty-two consonant phonemes. The median inventory size for NMLS is nineteen consonant phonemes, with the majority of languages contrasting somewhere between sixteen and nineteen consonant phonemes. One language, Enggano, has a much smaller inventory. The most comprehensive study lists twelve consonant phonemes, but of these, two—/r/ and /l/—are extremely rare outside of loanwords (Yoder 2011/2014: 21–4). Two other Barrier Islands languages, Devayan and Nias,

[1] In a recent paper, Billings and McDonnell (in revision) propose that Gayo, Enggano, Nasal, and Barrier Islands-Batak languages comprise a single subgroup, Sumatran (cf. also Smith 2017b).

[2] It was not possible to provide an exhaustive discussion of all NMLS. We chose to focus on the better described Batak languages. Leukon and Haloban are also not discussed because they lack descriptions see §28.6.

Bradley McDonnell and Christina L. Truong, *Non-Malayic languages of Sumatra and the Barrier Islands.* In: *The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia.* Edited by: Alexander Adelaar and Antoinette Schapper, Oxford University Press. DOI: 10.1093/oso/9780198807353.003.0028



Map 28.1 Location of NMLS in Sumatra and the Barrier Islands.

Sources: Language data from SIL International (R) (c) 2020. Includes geodata from Natural Earth (public domain) and Badan Pusat Statistik (licensed under CC BY-SA 4.0, https://creativecommons.org/licenses/by-sa/4.0).

have larger inventories with twenty-one and twenty-two consonants, respectively.[3] In some cases, the size of the inventory depends upon the analysis of a series of postploded nasals as either phonemes or allophones, which is discussed below.

Table 28.1 presents the consonant phoneme inventory of Gayo.[4] This inventory is representative of the majority of NMLS and exemplifies many common features of such consonant systems. The majority of these languages contrast voiced and voiceless stops at the labial, dental/alveolar, and

[3] These numbers do not include phonemes that only occur in loanwords. Such phonemes, however, are discussed below.

[4] In tables showing phonemic inventories, we have given orthographic representations in angled brackets.

Table 28.1 Gayo consonant inventory (Eades 2005)

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stop	p b	t d			k g	(ʔ) ⟨'⟩
Nasal	m	n		ɲ ⟨ny⟩	ŋ ⟨ng⟩	
Tap/Trill		r				
Fricative		s				h
Affricate			tʃ ⟨c⟩ dʒ ⟨j⟩			
Approximant	w			j ⟨y⟩		
Lateral		l				

velar places of articulation as well as alveopalatal affricates /tʃ/ and /dʒ/. There are several exceptions. Nias lacks a voiceless labial stop /p/ but instead has a labiodental fricative /f/. Enggano lacks a voiced velar stop /g/ and any affricates. Toba Batak lacks a phonemic voiceless alveopalatal affricate /tʃ/. According to Adelaar (1995d), it is only found in recent loanwords and was absent from van der Tuuk's (1971 [1864-1867]) grammar. Acehnese lacks any affricates but has voiced and voiceless alveopalatal stops /c/ and /ɟ/. The majority of NMLS have a phonemic glottal stop. Only the Batak languages, Sigulai, and the Kebanagung dialect of Rejang lack one. As for Gayo, the glottal stop is somewhat marked as it occurs only in recent loanwords.

In some languages, the voiceless apical stop is dental [t̪] (e.g. in Toba Batak and Karo Batak) or alveo-dental (e.g. in Gayo), while its voiced counterpart is alveolar [d]. This asymmetry is common throughout languages in western Indonesia (Blust 2013a: 172). Furthermore, in Acehnese, voiceless stops are aspirated [p^h, t^h, c^h, k^h] and voiced stops are 'murmured' with a 'whispery voice' phonation [b^h, d^h, $ɟ^h$, g^h] in onset position (Durie 1985: 26). As is the case for other NMLS, Acehnese stops are unreleased word-finally.

The majority of NMLS distinguish four nasal stops: labial, alveolar, palatal, and velar. Exceptions include Batak and Sigulai, which lack a palatal nasal and thereby distinguish only three nasals, and Enggano and Nias, which distinguish only two nasals, labial and alveolar. Acehnese, Rejang, and older speakers of Gayo additionally have a series of four postploded (or 'funny' nasals)—a consonant with nasal airflow through the closure followed by a brief plosive release—at labial, alveolar, palatal, and velar places of articulation.[5] Historically, these postploded nasals transparently come from nasal + voiced stop sequences. However, synchronic analyses of postploded nasals vary, which has consequences for the size of the consonant and even vowel inventories. Durie (1985), for example, analyses postploded nasals in Acehnese as allophones of their plain nasal counterparts. In his analysis, allophonic variation is triggered by the presence or absence of a immediately following nasal vowel in a stressed syllable: plain nasals occur before nasal vowels, while postploded nasals occur before oral vowels. As a result, Acehnese is analysed as contrasting nineteen rather than twenty-two consonant phonemes. Simplicity in the consonant inventory, however, results in complexity in the vowel inventory, where two sets of phonemes—nasal vowels and oral vowels—must be included under this analysis (see §28.2.1.2 for details). McGinn (2005) analyses postploded nasals in Rejang as phonemic. Nasalized vowels are the result of nasal spreading and are thus allophonic: vowels following plain nasals are nasalized (e.g. /maʔaʔ/ [mãʔaʔ] 'approach'), while vowels following postploded nasals are not (e.g. /im^{b}o/ [im^{b}o] 'forest') (McGinn 1982: 63). This results in a larger consonant inventory with twenty-two consonant phonemes. An alternate analysis of the postploded nasals as allophonic would mean that Rejang has only eighteen consonant phonemes. On a conceptual level, both Durie and McGinn's analyses are equally defensible for either language. There is no apparent advantage to analysing the postploded nasals as phonemic and vowel nasalization as allophonic versus analysing vowel nasalization as phonemic, with postplosion of nasals as allophonic.

Cohn and Riehl (2016), however, argue that postploded nasals in Acehnese and Sundanese ought to be analysed phonologically as nasal–stop clusters rather than unary segments. They show that the distribution of postploded nasals is the same as clusters of nasal + voiceless stop. They also provide phonetic evidence that the duration of Acehnese postploded nasals is more consistent with a cluster analysis

[5] These postploded nasals have also been referred to as "barred" nasals in Rejang (Coady and McGinn 1982).

than a unary one. Of course, the analysis of postploded nasals is language dependent, as Cohn and Riehl themselves point out, so without further research, the status of Rejang postploded nasals remains an open question. In Gayo postploded nasals are only contrastive for older speakers of the language (Eades 2005). Apparently, younger speakers of Gayo have collapsed plain and postploded nasals. In addition to postploded nasals in word-medial position, word-final nasals following oral vowels are optionally preploded in Rejang. Word-final preploded nasals are also apparently found in some dialects of Mentawai (Blust 1997b: 169, 2013a: 241).

The number of fricative phonemes in NMLS ranges from one to five. With the exception of Enggano, NMLS have at least the fricative /s/, and many languages, like Gayo, also have a glottal fricative /h/ (i.e. Acehnese, Kebanagung and Rawas dialects of Rejang, Batak languages). Among the Barrier Islands languages, Devayan and Sigulai have a series of four fricatives /f, s, x, h/ and Nias has a fifth phoneme /v/. In this regard, Mentawai diverges from the other Barrier Islands languages as it has only a single fricative /s/. The Pesisir, Lebong, and Musi dialects of Rejang also have a single fricative /s/.

Several languages have a back (i.e. velar or uvular) fricative or approximant, which alternates with an alveolar trill /r/ either in loanwords or between dialects. In Nasal, the alveolar tap/trill /r/ only occurs in some recent loanwords, while the uvular fricative /χ/ occurs in inherited vocabulary and older Malay loans where Malay has an alveolar trill /r/ (e.g. Malay *perang* 'war' is pronounced [pəχaŋ]). Lampungic languages have been described similarly (Walker 1976: 3). In Acehnese, Durie (1985) describes the orthographic <r> as a uvular approximate [ʁ] in the northern dialect, but in other dialects, it is realized as an alveolar trill [r]. The Barrier Islands languages have an alveolar trill /r/ and all but Mentawai also have a voiceless velar fricative /x/. However, these velar fricative phonemes do not appear to have the same relationship to the alveolar trill as seen in Nasal, Lampung, and Acehnese. Interestingly, Rejang is the only language in this region to lack both an alveolar trill and any sort of back fricative or approximant. All other languages have an alveolar trill, and all have a lateral approximant phoneme, however in Enggano, /l/ is found only rarely.

The phonemic status of labiovelar /w/ and palatal /j/ approximants is questionable in several languages. One common issue is that it is difficult to know how to analyse instances of [w] and [y]. In some cases, these glides are predictable between two vowels, where they break up vowel hiatus. In other cases, /w/ and /j/ are best analysed as vowels which fill an onset or coda position in the syllable, as in a tier-based analysis such as Hayes (1989). In Nasal, /w/ occurs in a number of words that are not predictable (e.g. *watu* 'stone'), while /j/ is much more rare. The most extreme case is Toba Batak. Adelaar (1995d) points out that /w/ and /y/ were not included in van der Tuuk's (1971 [1864–1867]) grammar and are only present in recent loanwords. Karo Batak, however, shows evidence of both labiovelar and palatal phonemic approximants. Finally, Sigulai lacks /w/ but apparently has /y/.

The Nias consonant inventory presented in Table 28.2 diverges from the other languages in significant ways. With twenty-two consonant phonemes, it is the largest inventory. It has several marked sounds. Most famously, Nias has the bilabial trill /ʙ/, which is especially remarkable because it can occur before any vowel; Ladefoged and Maddieson (1996) found that in other languages the bilabial trill only occurred before rounded vowels such as [u]. Nias also has an alveolar stop with trilled release /dʳ/ and a labiodental approximant /ʋ/. Yoder (2018) reports that Nias /ʙ/ and /dʳ/

Table 28.2 Nias consonant inventory from Brown (2001)

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stop	b t	d			k g	ʔ ⟨'⟩
Trilled release		dʳ ⟨ndr⟩				
Nasal	m	n				
Tap/Trill	ʙ ⟨mb⟩	r				
Fricative	f v	s			x ⟨kh⟩	h
Affricate			tʃ ⟨c⟩ dʒ ⟨j⟩			
Approximant	ʋ ⟨β⟩			j ⟨y⟩	w	
Lateral		l				

vary in phonetic realization between plain stop, stop with trilled release, and stop with fricative release.

28.2.1.2 *Vowels*

Most NMLS have average to moderately large sized vowel inventories (Maddieson 2013b) of five to eight vowel phonemes. A typical vowel inventory includes two front unrounded vowels /i/ and /e/, two back rounded vowels /u/ and /o/, the low central vowel /a/, and one mid or high central vowel. Table 28.3 shows the vowel phonemes of Lampung Api, which exemplifies a typical inventory. Walker (1976: 4) notes that realizations of the high vowels /i, u/ is varied across speakers, but they are frequently lowered to mid-high allophones [e, o] in unstressed open syllables and following certain consonants.

Table 28.3 Lampung Api vowel inventory

	Front	Central	Back
High	i		u
Mid	ɛ ⟨é⟩	ə ⟨e⟩	ɔ ⟨o⟩
Low		a	

A few NMLS contrast fewer than six vowel qualities. Nasal has five vowel phonemes (/a, i, u, ə, o/), lacking a mid front vowel /e/. Pampus (1989) lists Mentawai with only five vowel qualities (/a, i, u, e, o/). A number of languages, including Karo Batak, Toba Batak, Gayo, and Acehnese, have more than six vowel qualities because they distinguish two levels of mid vowels. Among NMLS, Acehnese, as shown in Table 28.4, has the greatest number of vowel qualities for monophthongs, with ten contrastive qualities for oral vowels (including four heights for central vowels) and seven contrastive qualities for nasal vowels. Enggano and Acehnese have also both been analysed as having phonemic nasal vowels (Durie 1985; Yoder 2011/2014), though not all scholars are in agreement on how to account for the nasalization patterns in these languages (see discussion of postploded nasals above and word-level nasalization below).

Some NMLS, such as Karo Batak, Toba Batak, and Nias, are reported not to contrast diphthongs phonemically, while other languages have a large number of contrastive diphthongs. For example, Rejang and Acehnese contrast five diphthongs (see Table 28.4 below), while Enggano contrasts six (/ai, aɨ, au, ei, əi, oi/) (Yoder 2011/2014: 32).[6] Others, like Lampung (Anderbeck 2007b) and Nasal, contrast just three diphthongs (/au, ai, ui/). Gayo contrasts diphthongs in some dialects and in poetic speech.

Table 28.4 Acehnese vowel inventory

	Front	Back	
	Unrounded	Unrounded	Rounded
Oral Monopthongs			
High	i	ɯ	u
Mid-high	e	ɤ	o
Mid-low	ɛ	ʌ	ɔ
Low		a	
Nasal monophthongs			
High	ĩ	ɯ̃	ũ
Mid-low	ɛ̃	ʌ̃	ɔ̃
Low		ã	
Oral Diphthongs			
High	iə	ɯə	uə
Mid-low	ɛə		ɔə
Nasal Diphthongs			
High	ĩə	ɯ̃ə	ũə
Mid-low	ɛ̃ə		ɔ̃ə

Enggano evinces an unusual word-level nasal harmony, in which a word has either all nasal vowels (e.g. [kãʔĩʔ] 'strong', [kũkũ] 'follow', [nõʔn̥õ̤] 'spilled') or all oral vowels (e.g. [kudi] 'belt', [ʔiaʔ] 'tie') (Yoder 2011/2014: 33). What is more, these nasal/oral vowel patterns do not appear to be the result of a single process. Yoder (2011/2014) proposes two nasal spreading processes that account for these patterns. First, in words containing nasal consonants, vowels are always nasalized, which can be attributed to spreading of nasalization from nasal consonants to underlyingly oral vowels in the same word. Second, in words with only oral consonants, either all vowels are oral, or all are nasalized. Nasalization in such words is attributed to phonemically nasalized vowels and subsequent spreading. Spreading of nasalization can be observed for oral vowels in prefixes when attached to a stem that contains a [+nasal] feature. For example, the /a/ in the Enggano adjectival prefix remains oral in words

[6] Diphthongs in Rejang vary significantly across different dialects, see McGinn (2005: 19).

like /kaʔ-pix/ [kaʔəpiç] 'sudsy', but is nasalized in words like /kaʔ-kĩh/ [kãʔəkĩh] 'dry' (in which the root vowel is phonemically nasal) and /kaʔ-man/ [kãʔəmãn] 'fragrant' (in which the root contains a nasal consonant). For a diachronic account, see Smith (2020).

28.2.2 Stress and prosodic prominence

NMLS represent a diversity of stress and prosodic prominence systems. Among these, Acehnese stands out, as it reflects the sesquisyllabic word structure of the Chamic subgroup (see Brunelle, chapter 11, this volume). Acehnese words consist of one or two syllables, of which the ultimate is most prominent. All Acehnese vowels, including diphthongs and nasal vowels, are contrastive in the final syllable, as opposed to the penult, which is characterized by fewer contrasts, and shorter and less distinct vowel pronunciation (Durie 1985). Only the final syllable of a word can bear stress, which is assigned at the phrase level to the penultimate or final word of the phrase.

The majority of NMLS have been described as possessing a predictable stress system. Of these Gayo, Sigulai, Enggano, Lampung, and Rejang (in addition to Acehnese as mentioned above) are reported to exhibit fixed final stress, while Karo Batak, Nias, and Mentawai are reported to show predictable penultimate stress with some exceptions. However, Toba Batak, has been described as a distinctive stress system. While stress in Toba Batak falls on the penultimate for most nouns and verbs, some words have final, lexical stress (van der Tuuk 1971 [1864–1867], Nababan 1981). Stress also has a derivational role: adjectives used predicatively bear final stress, while adjectives used attributively and related abstract nouns bear penultimate stress, as shown below in (1).

(1) Toba Batak

tibó	'high'	Predicative adjective
na tíbo	'which is high'	Attributive adjective
tíbo	'height'	Noun

(Roosman 2007: 93)

Many of the descriptive accounts of stress in these languages have relied on researcher perception, and further research is needed to identify the correlates of stress and to distinguish word stress from phrasal prominence (see Kaufman and Himmelmann, chapter 42, this volume). Roosman (2007) stands out as an exception to the general state of research on prosody for NMLS. Her study shows that word-level stress in Toba Batak is marked by rising pitch and that this pitch movement is influenced by prosodic boundaries and intonation. In general, however, further detailed phonetic study of prosody for these languages is needed.

28.2.3 Phonological processes

This section presents some characteristic phonological processes in NMLS, including nasal substitution, consonant gemination, and reduplication. In addition, we also include discussion of the distinctive morphophonemic process in Nias known as nominal mutation.

Like other western Indonesian languages, most NMLS exhibit a morphophonemic processes known as nasal substitution and nasal accretion. These processes have been reported for Gayo, Karo Batak, Toba Batak, Devayan, Sigulai, Nias, Mentawai, Rejang, Lampung, and Nasal. In most of these languages, the morpheme that triggers the process is a verbal prefix with the shape *mang-*, *maN-*, or even *N-* where *N* represents a nasal unspecified for place (see §28.3.1 for discussion of the function of this prefix). There is also frequently a corresponding nominal prefix, *pang-*, *paN-*, or similar, that also triggers the process. The initial stem consonants which are subject to substitution vary slightly from language to language; these generally include voiceless obstruents but can exclude /h/ and other back consonants such as /x/ and /k/ (e.g. Lampung, Nias). In the case of initial /s/ and alveopalatal consonants, the substituting nasal is usually palatal if the language has a palatal nasal phoneme, and alveolar otherwise, but some variation for /s/-initial stems has been reported for Toba Batak, Gayo, and Rejang (see Toba Batak examples in Table 28.5 below).

A number of NMLS show a process of consonant gemination, though the underlying motivations are not always similar. In Toba Batak, a nasal assimilates completely to a following stop, resulting in a geminate stop. This is an active synchronic process that occurs at a morpheme boundary with some exceptions, including environments which trigger nasal substitution as discussed above. Table 28.5 presents a paradigm for the Toba Batak prefix *mang-* that includes examples of both nasal substitution and consonant gemination using data from Percival (1981) and Nababan (1981). Some lexical items show exceptional behaviour compared to the general pattern for stems of the same shape; examples of these are indicated by asterisks in the table. Consonant gemination in Toba Batak can also be shown to have occurred diachronically for PMP nasal–stop sequences *mp, *nt, and *ŋk (Blust 1995c). In Karo Batak and Lampung, consonant gemination has a different trigger (Anderbeck 2007b). When a schwa in an open syllable appears in the stressed penult of a word, the initial consonant of the following syllable undergoes gemination, as in Karo Batak /təbu/ [təb.bu] 'sugarcane' (Woollams 1996: 30). This process appears to be motivated by a preference for longer duration in stressed syllables.

Table 28.5 Morphophonemic behaviour of Toba Batak *mang-*

Stem	/maŋ-/ + Stem	Gloss	Process
pukkul	*mamúkkul*	'to beat'	nasal substitution
bolos	*mamólos*	'to pass'	nasal substitution
boan	*mabbóan*	'to bring'	gemination*
taru-hən	*manaruhən*	'to convey'	nasal substitution
duda	*maddúda*	'to stamp rice'	gemination
cubo	*maccúbo*	'to try'	gemination
jalə	*majjálə*	'to receive'	gemination
karejo-hən	*makkarejóhən*	'to do'	gemination
gadis	*maŋgadis*	'to sell'	nasal accretion
mata-hən	*mamatáhən*	'to supervise'	nasal substitution
nipuran-i	*maŋanipuráni*	'to hand round betel nut'	V-insertion
suru	*manuru*	'to sell'	nasal substitution
siamun	*maŋiamun*	'to turn left'	nasal substitution*
həna	*maŋóna*	'to be efficacious'	nasal substitution
huliŋ	*makkuliŋ*	'to speak'	gemination*
lɛan	*maŋalɛán*	'to give'	V-insertion
rippu	*maŋarippu*	'to guess'	V-insertion
usuŋ	*maŋúsuŋ*	'to carry'	—

Reduplication is a very common morphological process in NMLS. In almost every one of them, two types of reduplication are active: single syllable reduplication and disyllabic or total reduplication. Examples of the two types from Toba Batak include *ta~tangis-an* (HAB~cry-NMLZ) 'that about which one cries' (van der Tuuk 1971 [1864–1867]: 395), and *poso~poso* (NMLZ~young) 'a newly born young (of animals)' (van der Tuuk 1971 [1864–1867]: 395).

While the reduplicant is commonly an exact copy of all or part of the base, sometimes phonological variation exists between the form of the base and the reduplicant. Certain reduplicated forms in Acehnese, for example, show variation in quality of the stressed vowel, as in *muda~mud**i*** 'young people', from base *muda* 'young' (Durie 1985: 43). The vowel in the reduplicant may also be systematically reduced. For instance, in western dialects of Karo Batak, initial-syllable reduplicants often are pronounced with a schwa instead of the full quality of base vowel, as in *p**ə**~pagi* 'tomorrow', from base *pagi* 'morning; tomorrow' (Woollams 1996: 92). Consonant alternation may also occur, as in Gayo *sakit~**m**akit* 'very sick; difficult', from the base *sakit* 'sick', where the substitution of /m/ for the initial consonant is reported to have an intensifying function (Eades 2005: 55).

As the examples in this section show, reduplication has a variety of functions. For nouns, it indicates plurality or variety. For verbs, it can indicate multiple actors, distributive action, iterative, habitual, or progressive aspect, or more intense state or action. Reduplication is also used to form words indicating manner and similitude or imitation, as in Karo Batak *pe-ganjang~ganjang-ken* (CAUS-SIM~tall-APPL) 'to put on airs and graces' (Woollams 1996: 95). Reduplication can also be used to form nouns, such as instruments associated with the base action, as in Nias *raʔu~raʔu* 'small net for catching fish', from the base *raʔu* 'catch'.

Nias exhibits a striking morphophonemic process known as nominal mutation. The case-marking functions of nominal mutation are discussed in §28.4.1.2. Nominal mutation occurs on nouns (including pronouns), classifiers, and the prefix *ira-* which indicates plurality. Phonologically, nominal mutation is marked by alternations in the initial consonant of consonant-initial words, the addition of initial /g/ or /n/ to vowel-initial words, or irregular changes for pronouns and *ira-*. Examples of nominal mutation are shown in Table 28.6 below. Note that a similar nominal mutation process has been reported for Sigulai (Kähler 1955: 11–12), but details about its application and function are much less clear (see discussion in §28.4.1.2). No other NMLS display a similar phenomenon. In examples below, nominal mutation is indicated by the glossing convention MUT.

28.3 Morphology

The NMLS are mildly synthetic with both prefixes and suffixes. Acehnese and Rejang are exceptions as they have no suffixes. In some languages, a base may only take a single prefix (e.g. Nasal), while in others several prefixes can stack (e.g. Nias). The use of infixes and circumfixes is also quite common, as is reduplication (see §28.2.3 above). The remainder of this section discusses some of the most common affixes and morphological processes in NMLS, but it is by no means exhaustive.

Table 28.6 Nias nominal mutation

Base item	Mutated form	Alternation
fakhe 'rice'	*vakhe*	f → v
kefe 'money'	*gefe*	k → g
ciʔaciʔa 'gecko'	*ziʔaciʔa*	c → z [dʒ]
baβi 'pig'	*mbaβi*	b → mb [ʙ]
doi 'thorn; fishbone'	*ndroi*	d → ndr [dʳ]
oβo 'boat'	*noβo*	∅ → n
oβoto 'small dike'	*goβoto*	∅ → g
yaʔo 1SG	*ndrao*	irregular
yaʔia 3SG	*ya*	irregular
ira- COLL	*ndra-*	irregular

28.3.1 Verbal morphology

28.3.1.1 *Agreement*

Three NMLS mark some form of agreement on verbs: Nias, Acehnese, and Enggano.[7] In all three languages, the bound affixes or clitics commonly occur without any cross-referenced lexical NP, meaning that they represent a fairly non-canonical form of agreement (Corbett 2006).

In Nias, agreement is marked by prefixes that distinguish number and person. These appear on verbs in transitive clauses, in which case they agree with the actor argument (A), and on irrealis verbs in intransitive clauses, in which case they agree with the single argument (S). Realis verbs in intransitive clauses take no agreement prefixes. Table 28.7 below shows the paradigm for Nias agreement prefixes.

Table 28.7 Pronominal prefixes in Nias by mood (Brown 2001: 124)

	Realis	Irrealis		Realis	Irrealis
1SG	*u-*	*gu-*	1PL.INCL	*ta-*	*da-*
			1PL.EXCL	*ma-*	*ga-*
2SG	*ö-*	*gö-*	2PL	*mi-*	*gi-*
3SG	*i-*	*ya-*	3PL	*la-*	*ndra-*

[7] Sigulai may also exhibit verbal agreement, but see §28.4.1.2 below for a discussion of complicating factors.

In Enggano, subject agreement is marked on the verb, but is limited to particular clauses (see §28.4.1.2 for details). It distinguishes number (singular, dual, plural) and person, and is indicated by means of a prefix or circumfix. Common forms used for subject agreement are presented in Table 28.8, however, other morphophonological changes may apply, resulting in surface forms that differ from those shown here (Crowley n.d.).

Table 28.8 Enggano subject agreement prefixes (Crowley n.d.: 21)

Singular		Dual or Plural	
1SG	*ʔu-*	1DU.INCL	*ka-*
		1PL.INCL	*ka- -aʔa*
		1DU/PL.EXCL	*ʔu- -ʔai*
2SG	*u-*	2DU/PL	*u- -aʔa*
3SG	*i-*	3DU/PL	*da-*

In Acehnese, agreement and/or argument cross-reference is marked on the verb by means of clitics, which, like other Acehnese pronominals, distinguish person, but not number (singular/plural) in most cases. Proclitics mark agreement for person with the Actor arguments in intransitive (S_A) or transitive (A) clauses, and enclitics optionally mark agreement with the Undergoer in intransitive (S_P) or transitive (P) clauses (see §28.4.1.3 for further discussion and examples). A paradigm is shown in Table 28.9.[8]

28.3.1.2 *Affixation on verbs*

This section describes some common affixes used on verbs in NMLS. These affixes mark transitivity, verb subclass (stative vs. dynamic), voice, valency, and other dimensions of the verbal action, such as aspect and volitionality.

Almost all NMLS make use of a verbal prefix that triggers the morphophonemic process of nasal substitution (see §28.2.3 above). These prefixes can be considered reflexes of PMP *maŋ- 'active verb' (Blust 2013a: 242). Only Acehnese and Enggano do not make use of such a verbal prefix. In Karo Batak, Lampung, and Nasal the prefix has been shortened to *N-* or *ng-*. In Gayo it is *mun-*, in Rejang it is *meng-*, and in all others the form is represented either as *mang-* or *maN-*, with phonological variation of the vowel quality in some cases. In languages with two basic transitive voices (see §28.4.1.1 below), this prefix typically marks what we are

[8] For simplicity, we exclude some pronominal forms listed as "very polite" or "reverential" by Durie (1985).

Table 28.9 Acehnese agreement clitics (Durie 1985: 117)

	Proclitic	Enclitic
1SG familiar	*ku=*	*=ku(h)*
1 polite	*lôn=, lông=*	*=lôn, =lông*
1PL.INCL neutral	*ta=*	*=teu(h)*
1PL.EXCL neutral	*meu=*	*=meu(h)*
2 familiar	*ka=*	*=keu(h)*
2 neutral	*ta=*	*=teu(h)*
2 polite	*neu=*	*=neu(h)*
3 familiar	*ji=, di=, i=*	*=ji(h), =i*
3 polite	*geu=*	*=geu(h)*

calling Actor Voice (AV); in other languages it commonly marks transitive and dynamic verbs. In a few languages, the function of the prefix is quite diminished, as with Nias *maN-/mo-*, which only marks intransitive dynamic verbs denoting an action associated with the base noun, as in *adulo* 'egg', *man-adulo* 'lay egg' (Brown 2001). In Rejang, *meng-* is restricted to intransitive verbs expressing durative action (McGinn 1982: 58).

In addition to reflexes of *maŋ, almost all NMLS have a reflex of the PMP infix *-um- which is associated with intransitive verbs (see Kroeger and Riesberg, chapter 47, this volume). Functions and forms of this affix vary across NMLS. One common function is to mark progressive, habitual, or durative aspect, as with Acehnese *meu-/-eum-*, Nias *m-/-um-*, and Mentawai *mu-*. In Devayan and Sigulai, the infix *-um-* marks verbs denoting processes. In Gayo, the prefix *mu-* marks intransitive verbs with various functions, including speaking verbs and inchoatives, while the infix *-em-* is used with verbs of motion. In Toba Batak, *(u)*m- or *-um-* forms intransitive dynamic verbs and verbs expressing comparison of state. In Rejang, *m(e)-/-em-* marks transitive (AV) verbs. In other languages, the function of this affix is narrow. In Karo Batak, *-um-* is not very productive and means to do an action 'erratically or unsteadily' (Woollams 1996: 77). In Enggano, *b(u)-/-ub-* is used in narratives to advance the action of the plot (Kähler 1940: 107). This affix appears to be used less frequently than the Enggano verbal prefix *ki-*, which simply marks the verbal predicate in a clause (see §28.4.1.2). An example is shown below in (2). In Nasal, the infix *-em-* occurs with a very limited number of roots and functions as a detransitivizing prefix that results in a meaning of 'do ROOT completely', as in (3).

Enggano

(2) *pãhũmãnã-hũmãnã* *ʔ<ub>ahadɔ,* *ʔ<ub> ahiudi*
morning-morning 1SG-BU-get.up, 1SG-BU-whistle.for
e-bɛɔ, *da-b-i* *e-bɛɔ* *ʔadiba.*
CORE-dog, 3DU/PL-PROG-come CORE-dog five
'I get up early in the morning, I whistle for the dogs, they come, five dogs.' (Kähler 1940: 107)

Nasal

(3) *iyo* *kak* *khadu* *suah* *di* *lahan=nyo.*
3 PFV finish <DETRANS>burn LOC field=3
'He burned his (entire) field.' (McDonnell fieldnotes)

Many NMLS contain a reflex of the Proto-Austronesian (PAn) infix *-in- (or its allomorph *ni-). In some NMLS, such affixes are used to indicate passive or Undergoer Voice (UV), as with Rejang *n(e)-/-en-*, Toba Batak *di-* and *ni-*, and Enggano *d(i)-* (see §28.4.1 below for a discussion of voice in these languages). Similar functions have also been reported for the prefix *ni-* in Sigulai and Devayan. In other cases, this affix forms nouns rather than verbs, especially deverbal nouns which result from the action indicated by the stem. For example, Gayo has *t-en-emak* 'rocks to channel water to a paddy' cf. *i-temak* 'to block irrigation' (Eades 2005: 69). In Acehnese, the infix *-eun-* is very productive and may form abstract properties, nouns expressing events, nouns denoting undergoers of an action, and instruments and locations associated with an action. In a few languages, such as Karo Batak, Sigulai, and Mentawai the affix appears only infrequently or is fossilized.

Many NMLS mark stative verbs with adjectival meanings by means of a verbal prefix derived from PAn *ma- 'stative prefix' (Blust 2013a: 376). This is the case for Batak languages and the Barrier Islands languages except for Enggano. Examples from Toba Batak are shown in (4). In some other NMLS, this affix is less productive, and may be associated with intransitives more widely, rather than stative verbs only. In Lampung Api, the prefix *ma-* is found on a subset of intransitive verbs which can have both dynamic and stative meanings as in (5).

Toba Batak

(4) a. *rara* 'red'
ma-rara 'to be red'

b. *bubu* 'fish trap (with an oblong bag)'
ma-bubu 'having a pot belly'
(van der Tuuk 1971 [1864–1867]: 87–8)

(5) Lampung Api

a.	*esaq*	'ripe'	root
	ma-esaq	'to be ripe'	INTR verb
	ng-esaq-ko	'to ripen s.t.'	TR verb
	(Walker 1976: 37, 41)		
b.	*impix*	'return'	root
	m-impix	'to return'	INTR verb
	(Walker 1976: 21, 40)		

Like other western Indonesian languages, NMLS make use of a number of verbal affixes that can change the valency of the verb. Reflexes of the PMP prefix *pa- have a causative function in Acehnese, Gayo, Batak languages, and the languages of the Barrier Islands.

Other applicative affixes commonly found in NMLS include reflexes of PMP verbal suffixes *-i, *-an, and *-aken (Blust 2003b). The functions of applicative suffixes derived from these forms vary across languages, with polyfunctionality of individual suffixes being very common. Common valency-increasing functions of these suffixes in NMLS include the introduction of a causer, patient, location, or beneficiary argument. However, non-valency increasing functions are also common, such as indicating iterative aspect (see §28.5 below), increased affectedness of the patient, or higher intensity, as well as lexicalized changes in meaning. For example, the Gayo suffix *-(n)en* signals a causative meaning when affixed to intransitive bases, signals a causative meaning or "an increase in volition or intention" when added to transitive bases, and allows addition of an affected P argument when affixed to nominal bases (Eades 2005: 186). Some examples of applicative affixes and their functions in Karo Batak are shown in Table 28.10.

28.3.2 Nominal morphology

For the most part, nouns in NMLS require no explicit marking for grammatical category such as number, gender, or animacy. Enggano is an exception; number, human vs. non-human, and proper vs. common noun are distinguished on nouns by means of prefixes. A paradigm is shown in Table 28.11. Some of these prefixes have another function in marking case. In general, case marking on nouns is rare among NMLS (see §28.4.1.2 for further discussion). The remainder of this section describes some common noun-forming affixes in NMLS.

Some NMLS make use of a nominal prefix derived from PMP *paŋ- which forms nouns that express an instrument or agent of the base action, or in some cases deverbal nouns in general. These include Karo Batak *peN-*, Toba Batak *paŋ-*, Gayo *pen-*, Lampung Api *paN-*, and Rejang *peng-*. Like reflexes of *maŋ, these prefixes trigger nasal substitution and accretion (e.g. Karo Batak *pengkawil* 'fisherman' from *kawil* 'to fish' (Woollams 1996: 77), and Gayo *penengkam* 'trap (n.)' from *tengkam* 'to catch' (Eades 2005: 38)).

Other nominal prefixes found in the NMLS include reflexes of the PMP nominalizing prefix *paR-. Examples include the Rejang deverbal noun prefix *pe-* (e.g. *pe-tulung* 'assistance', *tulung* 'to help') as well as noun-forming prefixes in Nias (*fa(ʔa)-*), Lampung Api (*pa-*), and Karo Batak (*per-*), which have more varied functions. Other affixes can also be used to form abstract nouns, especially those indicating abstract states. Examples include the circumfixes *ka- -an* in Lampung Api (e.g. *ka-xabay-an* 'fear', *xabay* 'afraid' (Woollams 1996: 95)), and *ke- -en* in Karo Batak (e.g. *ke-dung-en* 'conclusion', *dung* 'finished' (Woollams 1996: 86)). As discussed above, affixes derived from PMP *-in- may also play a role in noun formation.

A number of suffixes also function to form nouns, especially deverbal and locative nominalizations, in many NMLS. Some of these appear to be derived from PMP *-an, which as mentioned above also forms locative applicatives (§28.3.1). For example, Toba Batak *-an* may form nouns indicating the location of an action, as in *páŋan* 'eat', *paŋán-an* 'dish, plate' (Nababan 1981: 96). Lampung Api *-an* is very productive and often derives nouns indicating the result or object of an action, as in *aji* 'to chant', *aji-an* 'the chant' (Walker 1976: 26). In Nias, the suffix *-(C)a* forms nouns denoting the place or time of an action, as well as deverbal nouns and nouns denoting the object or result of an action. A similar suffix *-Ca* in Enggano forms "locational nouns" (Edwards 2015: 73).

28.4 Syntax

This section covers some basic syntactic properties of NMLS, including voice, case, agreement, and grammatical relations (§28.4.1), word order (§28.4.2), and the structure of the noun phrase (§28.4.3).

28.4.1 Voice, agreement, case, and grammatical relations

NMLS can be grouped into two basic types based upon whether the voice system is considered symmetrical or asymmetrical. Languages with symmetrical voice systems have multiple basic transitive constructions and are common throughout western Indonesia (see Kroeger and Riesberg, chapter 47, this volume). Languages with asymmetrical voice systems, on the other hand, have a single basic transitive construction with valency-changing alternations

Table 28.10 Causative and applicative affixes in Karo Batak

Affix	Function	Root		Affixed form	
pe-	causative	*galang*	'big'	*pe-galang*	'expand'
-ken	causative	*keri*	'depleted'	*keri-ken*	'deplete; use up'
-ken	theme applicative	*rukur*	'think'	*rukur-ken*	'think about'
-i	locative applicative	*kendul*	'sit'	*kendul-i*	'sit on; occupy'
-i	iterative aspect	*pekpek*	'hit'	*pekpek-i*	'hit repeatedly'
pe-...-i	causative, intensive	*kitik*	'small'	*pe-kitik-i*	'make smaller'
pe-...-ken	causative, intensive	*ganjang*	'high'	*pe-ganjang-ken*	'put up even higher'

Table 28.11 Enggano noun prefixes (Crowley n.d.: 11)

	Singular	Plural
human	*e-*	*ka-*
kinship term	∅-	*kahɔ-*
proper	∅-	—
common, non-human	*e-*	*e-*

(i.e. active–passive, ergative–antipassive) and are common among the world's languages. Symmetrical voice NMLS include all but Acehnese on mainland Sumatra (i.e. Nasal, Lampung, Rejang, Gayo, and Batak) and two languages of the Barrier Islands, Devayan and Mentawai.[9] Asymmetrical voice NMLS include most other Barrier Islands languages (Sigulai, Nias, and Enggano).[10] These two groups of languages have different strategies for marking grammatical relations. In symmetrical voice NMLS, grammatical relations are closely tied to voice, while in asymmetrical voice NMLS (i.e. Nias, Enggano), they are marked through case, agreement, and word order. Acehnese is in an uncertain position. It arguably has two transitive constructions, one in which the actor is unmarked and another in which it is marked by an agent case-marker *lé*. According to some studies, the latter is a passive, and Acehnese falls into the asymmetrical voice group (Legate 2012, 2014). However, because its position is still unclear, we provide a separate discussion of Acehnese.

[9] Based on descriptions of Devayan (Faridan 1981) and Mentawai (Morris 1900), these langauges appear to make use of AV and UV constructions that are consistent with those in the other symmetrical voice languages.

[10] However, we note that Nias maintains an alternation in relative clauses that is similar to symmetrical voice (see Brown 2001: 417 -21).

28.4.1.1 *Voice and grammatical relations in symmetrical voice languages*

Symmetrical voice NMLS have a two-way distinction between Actor Voice (AV), in which the transitive actor (A) argument is the subject, and Undergoer Voice (UV), in which the transitive undergoer (P) is the subject. In both constructions, the non-subject argument is not demoted, which distinguishes it from asymmetrical voice systems, which have a passive construction (Riesberg 2014; Chen and McDonnell 2019).[11] In NMLS the single (actor or undergoer) argument of an intransitive verb (S) is also considered a subject. The examples from Nasal in (6) below demonstrate these patterns of symmetrical voice and grammatical relations: (6a) is an intransitive construction where the verb is unmarked, (6b) is an AV construction in which the verb is prefixed with a homorganic nasal *ng-*, and (6c) is a UV construction in which the verb is prefixed with *di-*. In these constructions, the preverbal argument is the subject: S in (6a), A in (6b), and P in (6c). In transitive constructions, the post-verbal argument is considered a non-subject core argument, since they are unmarked and do not appear to be demoted. This includes P in (6b), and A in (6c).

Nasal

(6) a. *watang* *sijo* *masih* *hukhik.*
tree DEM.PROX still live
'This tree is still alive.'

b. *iyo* *kak* *ny-(s)uah* *lahan.*
3SG PFV AV-burn field
'He burned the field.'

[11] In some of the symmetrical voice languages, there is also a true passive construction in addition to AV and UV (see Kroeger and Riesberg, chapter 47, this volume).

c. *lahan* *ni* *kak* *khadu*
field that PFV finish
di-suah *Anton.*
UV-burn Anton
'Anton already burned the field.'
(McDonnell fieldnotes)

Evidence for a subject grammatical relation in symmetrical voice NMLS primarily comes from behavioural—as opposed to coding (Keenan 1976a)—properties. That is, subjects have the ability to raise, relativize, and/or float quantifiers (see Kroeger and Riesberg, chapter 47, this volume). Eades (2005) states that only subject arguments in relative clauses may be relativized. Examples are shown in (7) below: S is relativized in (7a), A is relativized in (7b), and P is relativized in (7c). The relative clause is within brackets, and the head noun, which is co-referential with the so-called 'gap' in the relative clause, is bolded.

Gayo
(7) a. ***jema*** *[si* *gintes* *kin* *aku]*
person REL surprised DAT 1SG
'the people who were surprised by me'

b. ***urang*** ***tue*** *[si* *mu-lahir-en*
person old REL AV-be.born-CAUS
tubuh=leu=ni]
body=1SG.POSS=this
'the parent who gave birth to me.'

c. ***kurik*** *[si* *i-geléh=è=a]*
chicken REL UV-slaughter=3SG.NSBJ=that
'the chicken that he slaughtered' (Eades 2005: 277–8)

Similar restrictions are found in other symmetrical voice NMLS such as Karo Batak (Woollams 1996; Norwood 2002), Toba Batak (Schachter 1984), and Nasal (McDonnell fieldnotes), but whether these hold in other symmetrical voice NMLS, such as Rejang, Devayan, and Lampung remains unclear because these languages lack descriptions of grammatical relations.

Non-subject arguments are more challenging to pin down. However, the discussion surrounding non-subject arguments in symmetrical voice NMLS concerns their core (or oblique) status (see Kroeger and Riesberg, chapter 47, this volume for a detailed discussion of the issues). The clearest evidence for the core status of non-subject arguments come from the strict position within the clause (see §28.4.2 below) and their lack of additional marking by prepositions. These properties are found in all of the symmetrical voice NMLS and demonstrated in the Nasal examples in (6b) and (6c), in which the non-subject arguments occur post-verbally and do not receive any additional marking.

28.4.1.2 *Case, agreement, and grammatical relations in asymmetrical voice NMLS*

Grammatical relations in asymmetrical voice NMLS are primarily marked by coding—as opposed to behavioural—properties, such as case marking and agreement but also word order (see §28.4.2). However, no clear patterns emerge for case and agreement across the two languages for which we have adequate descriptions, namely, Nias and Enggano. With limited descriptions of the remaining asymmetrical voice NMLS, Sigulai, our generalizations are limited.[12] Thus, we describe case and agreement in Nias and Enggano in this section.

Nias presents a particularly interesting situation because agreement and case-marking show two different patterns of grammatical relations, reinforcing the idea that grammatical relations are both language and construction specific (Bickel 2010). Agreement in Nias is obligatorily marked on the verb. It encodes person and number of S and A arguments in irrealis mode, as in (8) below. In realis mode, only A arguments show agreement, as in (9). Thus, agreement shows fairly straightforward evidence of a subject relation with nominative–accusative alignment.

Nias
(8) a. ***Ya****-ma-nana* ***nono-nia*** *ba*
3SG.IRR-DYN-hand child:MUT-3SG.POSS LOC
va-a-lio.
[NMLZ-STAT-quick]:MUT
'Her child will be crawling soon.'

b. ***Ya****-mbalö* *gefe* ***Ama*** ***Dali***
3SG.IRR-repay:IRR money:MUT Ama Dali
'Ama Dali wants to borrow (lit. 'repay') some money.' (Brown 2001: 502)

Nias
(9) ***I****-fa-tene* *ga* *ndrao*
3SG.RLS-DO-messenger here 1SG.MUT
khö-mi ***ama-gu****.*
DAT-2PL.POSS father-1SG.POSS
'My father sent me here to you.' (Brown 2001: 355)

Case in Nias is marked via a phonological process of nominal mutation (see §28.2.3). In general, A arguments do not undergo nominal mutation, as in (9), while S and P arguments do, as in (10). This apparent ergative–absolutive pattern in case-marking is exceedingly rare among the world's languages as the absolutive argument in Nias is marked (Comrie 2013). There are a some exceptions to this pattern. For example, clauses that contain one of several mental state

[12] Sigulai appears to have limited case marking via nominal mutation (see §28.2.3) and limited agreement (see discussion below).

verbs (e.g., 'like', 'fear') take two mutated arguments. There is also variation based on aspect (e.g. A is mutated when the verb is marked imperfective but not mutated when it is marked as perfective) or the (in)dependant status of a clause (e.g. A is mutated in some dependant clauses but unmutated in most independent clauses). For more details on these exceptions see Brown (2001: Ch. 7).

Nias

(10) a. *Aukhu* ***nidanö***
STAT:hot water:MUT
'The water is hot.' (Brown 2001: 342)

b. *La-bunu* ***mbaβi.***
3PL.RLS-kill pig:MUT
'They killed a pig.' (Brown 2001: 345)

Sigulai shows similar albeit more limited patterns of nominal mutation, for example, a subset of lexical nouns mutate but pronouns do not (Kähler 1955). However, these patterns do not clearly provide evidence for grammatical relations. While Kähler notes that nominal mutation occurs on subject arguments, which would presumably be S and A arguments, the only examples of noun mutation in Kähler (1955) involve S arguments (see (26a) in §28.4.2). A arguments are either realized as pronouns, as in (27a), or some sort of left-dislocated topicalization that apparently does not undergo noun mutation, as in (27b) below.[13] Given the limited description we have in Kähler (1955), patterns of nominal mutation in Sigulai are still unclear. Furthermore, Sigulai shows very limited evidence of agreement. It is only found in a very limited number of examples where the enclitic A argument cross-references a pre-verbal argument (see (27b) in §28.4.2).

Enggano marks case and agreement but such marking is much more limited than Nias. Person and number agreement only occur in a subset of clause types, and what might be considered case-marking prefixes are also noun classifiers, distinguishing human from non-human nouns (see §28.3.1.1). These hybrid case-marking/classifier prefixes do not distinguish grammatical relations but core arguments from several types of oblique or otherwise dependent arguments (Edwards 2015). That is, the case-marking prefixes *e-* (for non-human nouns, singular human nouns) and *ka-* (for plural human nouns) mark core arguments.[14] Other case-marking prefixes include the locative case-marker *i-* and a dependent prefix *u-*, which marks the possessor in nominal possession, nouns following the oblique marker *iʔiɔɔ/ʔɔ-* or the locative case-marked noun, among others (Crowley n.d.).

[13] Note that the example in (27b) is not evidence against nominal mutation of A. Brown (2001: 78–9) notes that when an argument is fronted before the verb in Nias, it is always unmutated.

[14] Proper names and kinship terms are unmarked for case.

The example in (11) demonstrates the functions of these three case-marking prefixes in Enggano.

Enggano

(11) *e-kɛʔɛpa* *eʔana* *ki-hɛkũ* *i-tɛbɛ* *u-kuɔ.*
CORE-bird that VERB-sit LOC-top DEP-tree
'The bird is sitting on top of the tree.'
(Kähler 1940: 182 cited in Crowley n.d.: 16)

Person and number agreement prefixes in Enggano only occur on verbs in a subset of clause types, including negative and subordinate clauses (Edwards 2015: 61). The examples in (12) demonstrate agreement with S and A arguments. In (12b), the third singular agreement prefix *i-* agrees with the subject *e-kaka eʔana* 'that person'. However, in (12a), the third singular agreement marker is prefixed to the verb, and here realized as *y-* before a vowel-initial root, but Kähler (1940) does not provide any examples where there is a lexical subject.

Enggano

(12) a. *kɛabaʔa* ***y****-ɛdɔ.*
NEG 3SG.AGR-cry
'He does not cry.' (Kähler 1940: 104)

b. ***e-kaka*** ***eʔana*** *kɛabaʔa* ***i****-pudu*
CORE-person that NEG 3SG.AGR-kill
e-kɔyɔ.
CORE-pig
'That person did not kill the pig.'
(Crowley n.d.: 41)

As case marking and agreement are limited, grammatical relations in Enggano are primarily evidenced in word order patterns. S and A arguments occur before the verb and P occurs after the verb (see §28.4.2 for more details).

28.4.1.3 *Agreement, case, and grammatical relations in Acehnese*

Acehnese grammatical relations have sparked controversy. Some have claimed that Acehnese has an active–passive alternation with subject and object grammatical relations (Lawler 1988; Legate 2012, 2014). Durie (1988) and others have argued that Acehnese does not have any grammatical relations since agreement patterns of arguments can be explained by semantic macroroles, Actor and Undergoer (Van Valin and LaPolla 1997: 255*ff*). According to Durie (1985, 1987), Acehnese distinguishes two macroroles, Actor (analogous to A) and Undergoer (analogous to P), in both transitive and intransitive clauses. Thus, S is split and could be represented as S_A for Actors and S_P for Undergoers (see Himmelmann 2005a: 133*ff*). These macroroles are primarily distinguished by the manner in which they are

cross-referenced on the verb: Actors are proclitics, while Undergoers are enclitics (see §28.3.1.1). Examples of this system are presented in (13).[15]

Acehnese

(13) a. *geu=jak* *gopnyan*
3.POL=go 3.POL
'(S)he goes.'

b. *gopnyan* *rhët(=geuh)*
3.POL fall=3.POL
'(S)he falls.'

c. *kèe* *h'an* *geu=paténh=kuh*
1SG.FAM NEG 3.POL=believe=1SG.FAM
gopnyan
3.POL
'(S)he doesn't believe me.' (Durie 1987: 370)

In most cases, the Actor proclitic is obligatory, while the Undergoer enclitic is optional. When a lexical Actor argument is present, it agrees with the proclitic, either S_A or A, as in (13a) or (13c), respectively. Likewise, when a lexical Undergoer argument is present, it agrees with the enclitic (if present), either S_P or P, as in (13b) or (13c), respectively.

Acehnese has a single case marker *lé*, which marks transitive Actor (A) arguments, as in (14). These case-marked Actors still agree with the obligatory proclitic on the verb. When an Undergoer enclitic is present, the case-marker is optional. Thus, the lexical Actor argument in (13c) above is not case-marked.

Acehnese

(14) *raja* *ji=kap* *lé* *uleue*
king 3.FAM=bite AGT snake
'A snake bit the king.' (Durie 1987: 371)

Durie (1987) shows that syntactic processes apply to either Actors (e.g. control), Undergoers (e.g. resultatives), or equally to Actors and Undergoers (i.e. core arguments). Based on this evidence, Acehnese does not appear to have any grammatical relations since the behaviour of arguments can be explained by these semantic macroroles. For alternative views on Acehnese, see Legate (2014).

28.4.2 Word order

NMLS are largely head-initial. They all have prepositions or case markers that occur before the noun (see discussion of Enggano case in §28.4.1.2 above). Standard negators occur before the predicates they negate and nouns occur before their modifiers with the exception of numerals and quantifiers, which commonly occur before the noun (see §28.4.3).

[15] Glossing has been updated to reflect the fact that these pronouns are clitics and not affixes.

The examples from Lampung Api below demonstrate head-initial word order in prepositional phrases (15) and negation (16).

Lampung Api

(15) *Holon* *hina* *lagi* *cecok* ***di*** *xangoq.*
person that continue stand at door
'That man is standing in the doorway.'
(Walker 1976: 11)

Lampung Api

(16) *Holon* *Lampung* *biasa=ni* ***maq*** *jadi*
person Lampung usual=3SG NEG become
padagang.
trader
'Lampung people usually don't become traders.'
(Walker 1976: 11)

In terms of clausal word order, NMLS are either verb-initial or verb-medial. However, in describing the particulars of clausal word order in these languages, it is again useful to treat asymmetrical voice NMLS separate from symmetrical voice NMLS because of the complications introduced by having two basic transitive constructions. We first discuss symmetrical voice NMLS.

28.4.2.1 *Word order in symmetrical voice NMLS*

Word order in symmetrical voice NMLS can be summarized as follows. The position of the non-subject argument is relatively constrained; it is, with very few exceptions, adjacent to the verb, forming a constituent that we refer to as the *predicate complex*. Subject arguments are less constrained and either occur before or after the predicate in intransitive clauses or the predicate complex in transitive constructions, resulting in the two patterns in (17). For our purposes, the notion of predicate complex subsumes predicates in intransitive constructions and the verbal predicate and non-subject argument in transitive clauses.

(17) a. Subject-initial order: Subject – Predicate complex
b. Subject-final order: Predicate complex – Subject

In symmetrical voice NMLS, both orders are possible. However, in two languages, Nasal and Lampung, the basic word order is subject-initial. To illustrate subject-initial word order, the Nasal examples in (18) are repeated from (6) in the previous section.

Nasal

(18) a. *watang* *sijo* *masih* *hukhik.*
tree DEM.PROX still live
'This tree is still alive.'

b. *iyo* *kak* *ny-(s)uah* *lahan.*
3SG PFV AV-burn field
'He burned the field.'

c. *lahan* *ni* *kak* *khadu* *di-suah*
field that PFV finish UV-burn
Anton.
Anton
'Anton already burned the field.'
(McDonnell fieldnotes)

In Batak languages, there is strong preference for subject-final word order (Cumming 1984), as in the Toba Batak examples in (19).

Toba Batak
(19) a. *Di-jahar* *si* *Poltak* *buku.*
UV-read PN Poltak book
'Poltak read the book.'

b. *Man-jahar* *buku* *si* *Poltak.*
AV-read book PN Poltak
'Poltak read the book.' (Erlewine 2016: 82)

In Gayo and Rejang, the basic word order differs based upon voice: intransitive and UV constructions prefer subject–final order, while AV constructions must be in subject-initial order, as in the Gayo examples in (20).

Gayo
(20) a. *I-jerang* *ine* *sine* *kerô=ni.*
UV-cook mother earlier cooked.rice=this
'Mother cooked this rice earlier.'

b. *Ine* *pora* *mi* *mu-jerang*
mother a.little more AV-cook
gule=ni.
edible.fish=this
'In a little while mother will cook this fish.'
(Eades 2005: 104)

As mentioned above, however, word order varies in all symmetrical voice NMLS. Subjects in Nasal and Lampung can appear after the predicate complex, and subjects in Batak can occur before the predicate complex. Only the AV construction in Gayo appear to be strictly subject initial, as demonstrated in the examples in (21). Crucially, the subject-final word order in (21b) is considered ungrammatical.[16]

Gayo
(21) a. *Aku* *mun-emah=è.*
1SG AV-make=3.NSBJ
'I made/am making it.'

b. * *Mun-emah=è* *aku.*
AV-make=3.NSBJ 1
'I made/am making it.' (Eades 2005: 174)

[16] Note that the predicate complex initial order in (21b) is possible when the AV construction is considered intransitive with a P argument that has a non-individuated reference, which Eades considers an incorporated noun. This may also be true for Rejang, but there is no evidence from ungrammatical examples. Interactions of voice and word order in Mentawai are unclear.

28.4.2.1.1 Order of non-subject arguments within the predicate complex

As mentioned above, non-subject arguments in all but a few cases are adjacent to the verb. While non-subject P arguments (in AV constructions) occur after the verb in canonical transitive clauses, non-subject A arguments (in UV constructions) in all symmetrical voice NMLS—with the exception of Rejang—are variable, occurring before or after the verb. The variable position is based on the person of the non-subject A argument. In Nasal, for example, first and second person non-subject A arguments occur before the verb while third person arguments occur after the verb, as in the examples in (22). However, in Gayo and Toba Batak, first person non-subject A arguments occur before the verb, while second and third person arguments occur after the verb.

Nasal
(22) a. *lahan* *ni* *kak* *khadu* *ku=suah.*
field that PFV finish 1SG=[UV]burn
'I already burned the field.'

b. *lahan* *ni* *kak* *khadu* *mu=suah.*
field that PFV finish 2SG=[UV]burn
'You already burned the field.'

c. *lahan* *ni* *kak* *khadu* *(di-)suah=nyo.*
field that PFV finish UV-burn=3SG
'He already burned the field.'
(McDonnell fieldnotes)

These non-subject A arguments are typically realized as clitics and in a number of cases differ in form from free pronouns (e.g. the Nasal first person singular free pronoun is *nyak* while the non-subject A form is *ku=*). If this argument is a proclitic it occurs without any UV prefix, but if it is an enclitic, it commonly combines with the UV prefix *di-*, as in (22c).

28.4.2.2 *Word order in asymmetrical voice NMLS*

In asymmetrical voice NMLS, the basic word order is either V(O)S (Nias, Sigulai), as in (23), or SV(O) (Enggano), as in (24).

Nias
(23) a. *Anakhö* *sibai* *ndrao.*
STAT.tired INT 1SG.MUT
'I'm very tired.' (Brown 2001: 190)

b. *I-rino* *vakhe* *ina-gu.*
3SG.RLS-cook rice:MUT mother-1SG.POSS
'My mother cooked rice.' (Brown 2001: 571)

Enggano
(24) a. *E-keʔepa* *ẽʔãnã* *kĩ-hãhãmõ.*
CORE-bird that VERB-fly
'That bird flies/Those birds fly.'
(Kähler 1940: 86 cited in Edwards 2015: 61)

b. *Kia* *ki-pudu* *e-koyo* *ẽʔãnã* *iʔioo*
3SG VERB-kill CORE-pig that PREP
u-bohe.
OBL-spear
'He kills that pig with a spear.'
(Kähler 1940: 196 cited in Edwards 2015: 61)

While word order appears to be fairly strict in asymmetrical voice languages, there is variation under certain conditions. For example, Brown (2001) notes that in Nias when the clause contains a clausal complement, it follows the verb and the subject argument.[17] In Enggano, variation from the basic SV(O) word order pattern occurs under two conditions. First, if the predicate is stative, the subject follows the verb, that is, it shows VS word order (Crowley n.d.: 32). The second condition is a bit more complicated. In Enggano, the verb is marked by either verbal agreement (see §28.4.1.2 above) or a verbal prefix *ki-* in basic SV(O) clauses, as in (24) above. However, there is a prefix *ka-* that when attached to the verb triggers a change in the word order to VS(O), as in (25). Other than triggering a change in word order, the semantic and/or pragmatic functions of *ka-* are not clear.[18]

Enggano
(25) a. *Kã-kõkõnã-hã* *e-koʔeʔe* *kude*
VERB-come.out-EMPH CORE-demon from
i-hoo *u-kuehi.*
LOC-inside DEP-forest
'The demon came out from inside the forest.'
(Kähler 1940: 203 cited in Edwards 2015: 61)

b. *Ka-nũkĩ* *e-kaka* *eʔana* *e-kihi.*
VERB-pull CORE-person that CORE-rattan
'That person pulled the rattan.' (Kähler 1940: 203)

Word order in Sigulai appears to be primarily verb-initial (Kähler 1955), as in (26b). However, in many examples in Kähler (1955), S appears before the predicate without any differences, as in (26a) below.

Sigulai
(26) a. *naitə̄* *mə=la* *mate.*
fire:MUT CMPL=EMPH die
'The fire has gone out.'

b. *löntuʔ* *bɔlöŋ-bölöŋ*
come beetle
'The beetle came.' (Kähler 1955: 17)

In transitive constructions, the vast majority of examples are verb-initial with A and P arguments occurring after the predicate. However, in these cases, the A argument appears to be a clitic pronoun, as in (27a). In the few examples where there are two lexical arguments, the lexical A argument appears before the verb but is set off with a comma, as in (27b), which likely means that it is prosodically marked. Whether these constructions represent word order variation or some sort of topicalization requires further research.

Sigulai[19]
(27) a. *maŋ-inu=do* *idanö.*
AV-drink=1SG water
'I drink water' (Kähler 1955: 28)

b. *baeliŋ,* *maŋ-ili=di* *bebiʔ-ni*
crab AV-molt=3SG shell-3SG.POSS
'The crab changes its shell.' (Kähler 1955: 28)

28.4.2.3 *Word order in Acehnese*

According to Durie (1985, 1988), Acehnese word order is relatively free, but the majority of clauses are predicate-initial and all arguments may follow the predicate in any possible order as long as A is marked by the case-marker *lé*, as in (28).

Acehnese
(28) a. *Geu=jak* *gopnyan*
3.POL=go 3.POL
'He goes.' (Durie 1988: 107)

b. *Ka* *geu=côm* *lôn* *lé* *gopnyan*
already 3.POL-kiss 1.POL AGT 3.POL
'She kissed me.'

c. *Ka* *lôn=pateh* *lé* *lôn* *aneuk*
already 1.POL=believe AGT 1.POL child
miet *nyan*
small that
'I believe that child.' (Durie 1988: 107)

The clearest constraint on word order is that only a single argument can precede the predicate, which Durie (1985) refers to as the 'Core Topic'. This argument can be S, A, or P, as in the examples in (29).

Acehnese
(29) a. *Gopnyan* *geu=jak*
3.POL 3.POL=go
'He goes.' (Durie 1988: 107)

[17] A similar pattern also occurs in Karo Batak AV constructions (Woollams 1996: 189).

[18] The *ka-* prefix presents a number of challenges. Edwards (2015) first notes that the *ka-* may be the combination of the verbal *ki-* prefix and a separate *a-* prefix, which triggers elision of the high vowel *i*. This means that *ka-* is not alternating with *ki-*, but represents an additional *a-* prefix. Furthermore, Edwards (2015: 61) summarises the only clue that Kähler provides to understand the function of *(k)a-*: "Verbs marked with the prefix *ki-* are described by Kähler (1940: 192–4) as having more nominal characteristics, while those with the prefix *ka-* are described as having more verbal characteristics."

[19] Since Sigulai is an asymmetrical voice language, *maŋ-* may be analysed as a marker of active voice.

b. *Gopnyan* *ka* *geu=côm* *lôn*
3.POL already 3.POL=kiss 1.POL
'She kissed me.'

c. *Lôn* *ka* *geu=côm* *lé* *gopnyan*
1.POL already 3.POL-kiss AGT 3.POL
'She kissed me.' (Durie 1988: 104–5)

28.4.3 Noun phrases

NMLS show a fair degree of similarity in regard to noun phrase (NP) structure. Across these languages, most elements of the NP occur after the head noun, including possessors, adjectival and nominal modifiers, relative clauses, and demonstrative determiners. Quantificational elements, which include numerals, numeral classifiers, and (non-numeral) quantifiers, are an exception and most commonly occur before the head noun. For some languages, such as Enggano and Sigulai, the position of some elements in the NP is not fully clear from available descriptions. The most typical NP structure in NMLS is shown in (30) with corresponding examples from Toba Batak (31) and Karo Batak (32).

(30) QUANT/NUM – Noun – PSR – Modifier – DEM

Toba Batak

(31) *sudé* *hálak* *húta* *(na)* *mɔra* *í*
all person village REL rich DEM.DIST
'all the rich villagers' (Nababan 1981: 109)

Karo Batak

(32) kenna kerbo bapa enda
all water.buffalo father DEM.PROX
'all these water-buffalo of Father's'
(Woollams 1996: 106)

Possessors most commonly occur immediately following the head noun. However, in some languages, including Acehnese, the possessor may either precede or follow an attributive modifier, as in (33a) and (33b) below, respectively.

Acehnese

(33) a. *sa-boh* *keubeue=**neuh*** *nyang* *ji=cu*
one-CLF buffalo=2.POL REL 3.FAM=steal
'one of his buffaloes that were stolen'

b. *bak* *pisang* *manyang=**lôn***
tree banana tall=1.POL
'my tall banana tree' (Durie 1985: 108)

Possessors may be either pronominal or lexical, with the former typically marked by a set of pronominal enclitics. Lexical possessor NPs often appear immediately following the head noun with no overt marking, but in some languages, possessors are grammatically marked. In Nias, possessor NPs and other nominal modifiers are marked with mutation, for example, *telau mbuaya* 'head of the crocodile', where the possessor noun *buaya* 'crocodile' is mutated (Brown 2001: 373), see §28.2.3. In Sigulai, this nominal mutation also marks nominal modifiers including possessors in a few nouns (Kähler 1955: 24). Particles may also be used to mark possessors as discussed below. Adjectival modifiers, nominal modifiers, and relative clauses typically occur after the possessor. In almost all NMLS, demonstrative determiners occur after the head noun and any possessors, modifiers, and relative clauses. However, Nias differs from this pattern. In Nias, a series of up to two demonstratives occurs after the head noun and any possessor as in (34), but Brown (2001: 372) notes that demonstratives can precede any number of other modifiers, which are analysed as relative clauses with verbal or numeral predicates.

Nias

(34) *Ba* *si?ulu=wa=e*
CONJ village.leader=that=Q
nama-da
ancestor:MUT-1PL.INCL.POSS
andre ***noema?e!?***
DEM.DIST DEM.RECOG
'And you mean that ancestor you've been talking about was a village leader!?' (Brown 2001: 411)

In some languages, the possession relationship, as well as other types of semantic relationships between the head noun and a following nominal modifier are marked with an intervening particle, which has been described as a 'ligature' or 'linker'. For example, in Gayo, full NP possessors are marked with the particle *ni=*, which often cliticises to the possessor NP, as in *kôrô ni ama* ~ *kôrô n=ama* 'father's buffalo.' The linker particle *ni=* is also said to mark kin relationships, part–whole relationships, and 'componential' relationships, in which the head noun is made up of the type of object expressed by the nominal modifier, as in *empus ni awal* (garden LNK banana) 'banana garden/plantation' (Eades 2005: 219). As mentioned above in §28.4.1, the prefix *u-* in Enggano similarly appears to mark dependent nouns including possessors and certain nominal modifiers (see (11) above). In Mentawai, *n=* (or allomorph *=t*) appears between a modifying nominal and the preceding head noun, as in *ūma n=abak* (house LNK=boat) 'boat house', and *mata=t ukui* (face=LNK father) 'face of father' (Morris 1900: 16).[20] Kähler (1955: 21) also mentions that attributive relationships can be marked by *-n* between the head noun and modifying nominal in Devayan as in *luma-n ana='u* (house-LNK child=1SG) 'my child's house'. In Toba Batak, some attributive modifiers are marked

[20] Morris (1900) does not distinguish clitics from affixes, but the behaviour of *n=* is more consistent with that of a clitic.

with *ni-* or *na-* (Nababan 1981: 108), while in Karo Batak, similar markers *ni ~ nu ~ u* are found before a possessor or nominal modifier only in archaic expressions (Woollams 1996: 137–8). As some of the preceding examples indicate, in many of these languages, there is no clear distinction between possessor NPs and nominal modifiers.

Of the NP elements discussed in this section, it appears that only quantificational elements canonically appear before the head noun. These include non-numeral quantifiers and their modifiers (such as degree words) or a numeral, sometimes together with a noun classifier. As an apparent exception, Enggano examples show that numerals and numeral classifiers follow the head noun (see (39) below). Additionally, descriptions of Nias, and Gayo show that some quantificational elements can 'float', that is, they may appear separated from the head noun, or entire NP, by another constituent (Brown 2001: 427; Eades 2005: 89). For example, in Gayo, a quantifier or numeral element most commonly appears immediately preceding the head noun, as in (35a) and (36a). These quantificational elements can also float, as shown in (35b) and (36b).

Gayo

(35) a. *Kahè i sien* ***delé*** *kule.*
later LOC here ***many*** tiger
'Later there will be many tigers here.'
(Eades 2005: 214)

b. ***Delé pedih*** *leu-engon kutu=mu.*
many very UV.1SG-see louse=2.POSS
'I see your many lice.' (Eades 2005: 89)

Gayo

(36) a. ***sara*** *belanga kôl jantar-dengké*
one pot big vegetable-meat
'a big cooking pot of vegetables' (Eades 2005: 215)

b. *Aku mun-osah emas kin tengku*
1SG AV-give gold DAT sir
sara bongkil.
one lump
'I will give you a lump of gold, sir.'
(Eades 2005: 216)

Besides cases of quantifier float mentioned above, variable position for quantificational elements has been described in a few other languages. For example, in both Acehnese and Mentawai, elements such as numerals and classifiers usually precede the head noun, but may also follow it (see Mentawai contrast between (38a) and (38b) below).

In NPs containing numerals, numeral classifiers are commonly used. Numeral classifiers are typically a small, closed class that are transparently derived from nouns. When classifiers are used, the common order is numeral–classifier, with the exception of Nias, where it is classifier–numeral. In a few languages, a ligature intervenes between numeral and classifier. In Lampung Api, this ligature is *nga*, as in (37). In Mentawai, it is *ṅa* (which is pronounced [ŋa] and is optional), as in (38). In Enggano, it is *h-* (< PMP **ŋa*), as in (38) (Crowley n.d.).

Lampung Api

(37) *telu ngam-biji manuq*
three LIG-COUNTER chicken
'three chickens' (Walker 1976: 17)

Mentawai

(38) a. *lima-ṅa munäṅ djō-djō*
five-LIG ANIM dog-PL
'five dogs'

b. *inu pulu šara tära bā̄*
bead ten one remainder COUNTER
'eleven beads' (Morris 1900: 23)

Enggano

(39) *e-ʔitɔ ʔakɔdu h-apɛ̃a*
CORE-banana three LIG-CLF
'three bananas' (Crowley n.d.)

28.5 Tense, aspect, and mood (TAM)

Tense, aspect, and modality (TAM) are marked in NMLS by a variety of strategies. These include the use of a special set of TAM markers, reduplication, and affixation directly on the verb—often with a portmanteau morpheme that signifies both aspect and voice. Most NMLS make use of more than one of these strategies. Additionally, Nias makes use of grammatical mood, which is described below. Aspectual meanings of reduplication are mentioned above (§28.2.3) and will not be discussed further here.

NMLS commonly make use of a closed set of morphemes that function as markers of tense, aspect, and/or modality. These markers are never obligatory, and their syntactic status varies across languages. Some TAM markers must be situated close to the verb and pattern like auxiliary verbs, while others are less integrated into the predicate complex and pattern like adverbs or other particles; many languages make use of more than one type. Some TAM markers are clitics (e.g. Nias, Sigulai), some are verbs (e.g. Acehnese, Toba Batak), and some are adverbials with a freer distribution (e.g. Gayo). Further discussion of these types of TAM markers follows.

The Acehnese preverbal TAM markers, which Durie (1985: 47) treats as complement-taking verbs, show close integration with the predicate. They can attach to the pronominal

clitic that appears on the verb, and are "pronounced as one phonological unit" with the predicate (Durie 1985: 248). Such preverbal TAM markers express tense and aspect including future, iamitive (meaning 'already'), and completive, as well as modality, such as probability and ability. The examples below show that the absence of a TAM marker, as in (40a), can indicate an intended action, while using the existential verb *na=* alongside a verbal predicate, as in (40b), indicates that the predicate is a fact (not merely an intention). In (40c) the iamitive marker *ka=* 'already' indicates that the event has happened; this morpheme also marks states that have already begun.

Acehnese

(40) a. *lön* *lön=jak* *u=kende*
1SG.POL 1SG.POL=go to=town
'I am going to town.' (expressing intention)

b. *lön* *na=lön=jak* *u=kende*
1SG.POL exist=1SG.POL=go to=town
'I go to town (habitually).' or 'I went to town.'

c. *ka=geu=jak*
already=3.POL=GO
'He has gone.' (Durie 1985: 248)

In Toba Batak, one set of TAM markers is treated by Nababan (1981) as a type of auxiliary verb. These are quite closely integrated with the predicate complex, as they must immediately precede or follow the head of the predicate. Toba Batak also makes use of particles which express aspectual or modal meaning. These are less closely integrated with the predicate; they occur either in clause-initial position or between the predicate and subject. The examples in (41) show the TAM marking auxiliaries *íkkɔn* 'must' and *musɛ* 'again' occurring in positions adjacent to the head verb *láo* 'go'.[21] The examples also show modal particles occurring between the predicate and subject, namely the affirmative particle *dɔ*, the narrative particle *ma* 'and so', and the distinct (but homophonous) hortative particle *ma*, respectively.

Toba Batak

(41) a. ***íkkɔn*** *láo* ***dɔ*** *hɔ* *tu* *húta*
must go AFF 2SG to village
'You must go to the village.'
(Nababan 1981: 85)

b. *láo* ***musɛ*** ***ma*** *ibána*
go again and.so 3SG
'(And so) he went again.' (Nababan 1981: 86)

c. *láo* ***ma*** *hɔ*
go HORT 2SG
'Please, go now!'
(van der Tuuk 1971 [1864–1867]: 360)

In Gayo, the set of aspectual markers include the progressive *tengah*, perfect *nge*, and immediate perfect *ben* or *teku*. These are treated as a type of adverb; they must precede the predicate they modify, but it is possible for a subject NP to intervene, suggesting that these are less tightly integrated with the predicate complex than the Acehnese preverbal TAM markers and Toba Batak auxiliaries.

In Nias, the proclitic *ma=* expresses perfective aspect. It occurs in realis clauses, either on the verb, as in (42a), or on one of a set of auxiliary verbs that express aspectual meanings, as in (42b) below. Sigulai appears to make use of a similar completive marker *mɔ=* (Kähler 1955: 17).

Nias

(42) a. *Ma=a-buso* *ndrao.*
PFV=STAT-replete 1SG.MUT
'I'm full. (I have become replete.)'

b. *Ma=aβai* *i-fazökhi* *zagö.*
PFV=finished 3SG.RLS-fix roof:MUT
'He has finished fixing the roof.' (Brown 2001: 478)

In addition to the use of special TAM markers, aspectual meaning in NMLS can be marked by affixation directly on the verb, and in particular has been tied to portmanteau functions of voice and applicative affixes (see §28.3.1). In

Table 28.12 Partial paradigm for voice and aspect in Toba Batak (Nababan 1981: 70–1)

Form	Gloss	Description
jóu	'call!'	imperative
joú-i	'call repeatedly!'	iterative, imperative
mad-jóu	'to call'	simple aspect, AV
mad-joú-i	'to call repeatedly'	iterative, simple aspect, AV
j<um>óu	'have called'	completive, AV
j<um>oú-i	'have called repeatedly'	iterative, completive, AV
hu-jóu	'called by me'	simple aspect, UV (inflected for 1SG actor)
j<in>óu	'have been called'	completive, UV
joú-ɔn	'will be called'	future, UV

[21] Spelling in these examples has been altered so that all conform to the same orthographic conventions.

Toba Batak, a rich system of verbal affixes are used to indicate TAM. There are two AV prefixes: *mang-* marks AV and simple aspect, while *-um-* marks AV and completive aspect. In UV clauses, simple aspect is indicated by *di-* and completive aspect by *ni-/-in-*, except for first person singular and first person plural inclusive actors, which take pronominal proclitics instead of these verbal prefixes. Nababan (1981: 70) also notes that the suffix *-ɔn* can be used to mark future action or to express a promise or duty. Batak languages also make use of the suffix *-i* to indicate iterative aspect. A partial paradigm showing voice and aspect in Toba Batak is given in Table 28.12. Despite the fact that they are correlated in some NMLS, few studies on the relationship between voice and aspect have been carried out for these languages, so this relationship is not well understood.

Nias is unique among NMLS for its use of grammatical mood. Nias verbs are obligatorily marked for realis or irrealis mood. Examples of irrealis mood in Brown (2001) show that it is used to express future, potential, hypothetical, and desiderative actions, among others. Grammatically, the distinction between irrealis and realis mood in Nias is indicated by the set of pronominal agreement prefixes that appear on the verb, as shown in Table 28.7 in §28.3.1.1 above. As mentioned earlier, realis verbs take no such prefix in intransitive clauses. This is shown in (43) where the realis verb is unmarked in (43a), but the irrealis verb is marked with *ya-* (3SG.IRR) in (43b).

Nias

(43) a. *Göna* *ya* *teu.*
be.struck 3SG:MUT rain
'She got caught in the rain.'

b. *Na* *mofanö* *ya* *mana*
if leave 3SG:MUT at.this.time
***ya**-göna* *teu.*
3SG.IRR-be.struck rain
'If she leaves now she will get caught in the rain.' (Brown 2001: 498)

A secondary indicator of mood is found in the presence of the infix *-um-*.[22] An example is shown in (44) where the realis verb is marked with the prefix *la-* (3PL.RLS) in (44a), and the irrealis verb is marked with both the prefix *ndra-* (3PL.IRR) and the infix *-um-* (IRR) in (44b).

Nias

(44) a. ***La**-ohe*
3PL.RLS-carry
'They carried (it).'

b. *Lö* *tola* *löʔö* ***ndra**-**m**-ohe.*
NEG can NEG 3.PL.IRR-IRR-carry
'They will have to carry it.' (Brown 2001: 504)

28.6 Conclusion and future directions for research on NMLS

This chapter represents the first of its kind to describe the typological features of NMLS. In doing so, we demonstrate that NMLS comprise a diverse group of languages with very little that unites them. Many NMLS exhibit features found in western Indonesian languages, such as an articulated voice system marked by several verbal affixes, applicative suffixes, reduplication, among other properties. Some NMLS contain more marked features, such as postploded nasals in Acehnese and Rejang, agreement in Acehnese, Nias, and Enggano, and case-marking in Enggano, Nias, and Sigulai. One typological parameter that is particularly informative is the distinction between asymmetrical and symmetrical voice languages. While Acehnese appears to fall somewhere between the two groups, the other NMLS can be classified in one group or the other. This division is significant because the symmetrical voice languages do not mark case or agreement, while many of the asymmetrical voice languages do. There are likely other relevant typological features that unite (some portion of) NMLS, but much more research into individual languages is needed in order to develop further cross-linguistic generalizations.

Despite a relatively long history of linguistic research in the region (e.g. Hazeu 1907; van der Tuuk 1971 [1864–1867]), many NMLS lack any in-depth description, including Nasal, Mentawai, Devayan, Leukon, Haloban, and Sigulai. Aside from Nasal, which the first author has a project to document and describe, the other languages have limited descriptions that vary in quality, reliability, and comprehensiveness. For the remaining NMLS much more description is still needed, especially Enggano, Lampungic languages, Rejang, and Batak languages such as Alas, Simalungun, and Dairi-Pakpak. These languages have little in terms of detailed, accessible descriptions. Furthermore, documentation in the form of archived collections of audiovisual recordings, transcriptions, and annotations of various speech events is even more rare.[23] In some sense, this is expected since much of the research done in Sumatra and the Barrier Islands was done before modern documentary linguistics was widely

[22] The infix *-um-* is obligatory for one verb class when the clause is irrealis. In this case, *-um-* does not indicate progressive aspect (as it may in realis clauses).

[23] Currently, there are closed collections of Mark Durie's materials on Acehnese (Durie 1980, 1982, n.d.), an open collection of Gayo audio recordings (Eades 1998), a collection of transcribed conversations of Mentawai (Gil 2015b), and a collection of Leukon audio and video recordings (Lubis and Williams 2019). The first author is also working on a project to document Nasal, which is archived with the Pacific and Regional Archive for Digital Sources in Endangered Cultures (McDonnell 2019).

practised. Thus, basic documentation and description of the NMLS listed above is the highest priority for this region.

Aside from the documentation and description of individual languages, there are many outstanding questions that we were unable to answer in this chapter. We list a few of these questions in (45).

(45) a. Are postploded nasals in Rejang and Acehnese unitary segments or nasal–stop sequences?
b. What is the nature of agreement and nominal mutation in Sigulai? How similar are they to Nias?
c. What is the nature of voice and grammatical relations in Mentawai, Sigulai and Devayan? How do they fit into the typology of voice systems (i.e. symmetrical vs. asymmetrical voice) in NMLS?
d. Other than triggering differences in word order, what are the functions of Enggano verbal prefixes *ki-* and *ka-*?

These questions represent just some of the specific outstanding issues in NMLS that deserve further detailed study.

Finally, we are also able to identify three broad areas of linguistic research on NMLS that ought to be prioritized. First, as descriptions of stress do not reliably separate word-level stress from phrase-level prominence(s), detailed phonetic studies on the interaction of word- and phrase-level prominence of individual languages are important for understanding prosody in NMLS. Second, grammatical voice and its relationship to grammatical relations, case, agreement, and word order is crucial to unlocking the grammars of NMLS. More studies that illuminate the defining properties of grammatical subjects and distinguish core arguments from obliques are needed. Third, as grammars of many NMLS (especially those of mainland Sumatra) appear to have undergone significant changes as a result of contact with Malayic languages (see e.g. Blust 1984 for Rejang; Adelaar 1995d for Toba Batak; and Anderbeck 2007b for Lampungic languages), further research on language contact may explain a number of the typological patterns we observed in NMLS.

Acknowledgements

We would like to thank Sander Adelaar, Antoinette Schapper, and Blaine Billings for valuable feedback on an earlier version of this chapter. We would also like to thank Owen Edwards, Mary Dalrymple, and I Wayan Arka for providing access to unpublished Enggano materials. Map 28.1 is used by permission and redistribution is not permitted. Special thanks to SIL International for permission to use language data displayed therein. The first author would also like to acknowledge Johan Safri, Wawan Sahrozi, and Anton Supriyadi, all of whom are collaborating on the documentation and description of Nasal. He is also grateful to his research counterpart in Indonesia, Yanti (Atma Jaya Catholic University of Indonesia), and to the Ministry of Research and Technology in Indonesia for allowing him to conduct research on Nasal. Discussion of the Nasal data is based upon work supported by the National Science Foundation under Grant BCS–1911641. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.