CHAPTER 23

MAINLAND SOUTH EAST ASIA

MARC BRUNELLE, JAMES KIRBY, ALEXIS MICHAUD, AND JUSTIN WATKINS

23.1 SCOPE OF THE CHAPTER

MAINLAND South East Asia (MSEA), often defined as a *Sprachbund*, is a linguistic area where languages from five different phyla (Austroasiatic, Austronesian, Hmong-Mien, Kra-Dai, and Sino-Tibetan) converge and have developed similar structures (Matisoff 1973; Alieva 1984; Enfield 2003, 2005).[1] While convergence processes are easy to identify in the region, its geographical boundaries are ill-defined, and one should not understate its typological diversity (Henderson 1965; Brunelle and Kirby 2016; Kirby and Brunelle 2017). In this chapter, we cover the area encompassing the Indochinese Peninsula (Vietnam, Cambodia, Laos, Thailand, Myanmar, and Malaysia) but also include Guangxi and Yunnan in southern China (excluding Chinese varieties, which are covered in chapter 22) and northeast India. As Austronesian languages are covered in chapter 25, our discussion of this phylum is limited to Chamic languages spoken in Vietnam and Cambodia and to Austronesian languages of the Malay Peninsula.

Our main goal is to give an overview of representative types of word-level (§23.2) and phrase-level (§23.3) prosody, highlighting areas of convergence between families without understating their diversity.

23.2 WORD-LEVEL PROSODY

In this section, we first discuss the most common word shapes and stress patterns found in MSEA (§23.2.1). As these two properties are largely dependent, they are discussed together. We then give an overview of the diverse tonation systems of the region (§23.2.2).

---

[1] Indo-European and Dravidian languages are also spoken by sizeable language communities in Burma, Malaysia, and Northeast India, but are not covered in this chapter.

23.2.1 Word shapes and stress

The basic vocabulary of many MSEA languages is monosyllabic. This is the case in most Kra-Dai and Hmong-Mien languages, but also in Vietnamese, an Austroasiatic language. However, in most of these languages, a significant part of the lexicon is made up of compounds, and most languages also have some polysyllabic loanwords. This can be illustrated with Vietnamese. The Vietnamese basic lexicon is largely monosyllabic, as illustrated in (1). Our transcriptions follow the conventions in Kirby (2011), except for the tone notation.

(1)   Vietnamese monosyllables
       đi   [ɗi⁴⁴]   'to go'   *nghiêng*   [ŋiəŋ⁴⁴]   'to be leaning'
       tuyết   [tɥiət⁴⁵]   'snow'   *ngoan*   [ŋwaːn⁴⁴]   'to be well-behaved'

However, Vietnamese has a significant proportion of non-monosyllabic words. According to Trân and Vallée (2009), 49% of its lexicon is disyllabic and 1% is trisyllabic. Native compounds (2) and reduplicants (3) make up most of the disyllabic vocabulary.

(2)   Native Vietnamese compounds
       nhà nghỉ   [ɲa²¹ ɲi²¹]   house+rest   'inn, low-end hotel'
       kiếm ăn   [kiəm⁴⁵ an⁴⁴]   search+eat   'to make a living'
       bố mẹ   [ɓo⁴⁵ mɛ³¹ˀ]   father+mother   'parents'
       vui tính   [vuj⁴⁴ ti̪⁴⁵]   happy+temper   'to be good-tempered'

(3)   Vietnamese reduplicants
       bạn bè   [ɓaːn³¹ˀ ɓɛ²¹]   friends + RED   'friends'
       tim tím [tim⁴⁴ tim⁴⁵]   RED + purple   'purplish'

Vietnamese also has a large number of compounds whose morphemes are borrowed from Chinese. These often have opaque semantics that, as such, seem better analysed as polysyllables (4). A significant number of loanwords from other languages are also polysyllabic, even if monomorphemic (5). Besides, although this is rarely pointed out in the literature, a number of native Austroasiatic words such as *tắc kè* [tak⁴⁵ kɛ²¹] 'gecko' and *thọc lét* [tʰɔk͡p³¹ let⁴⁵] 'to tickle' seem to constitute polysyllabic morphemes.

(4)   Opaque Sino-Vietnamese compounds
       tuần lộc   [twɤ̌n²¹ lok͡p³¹]   docile + deer   'elk, reindeer' 馴鹿
       thái độ   [tʰaːj²⁴ ɗo³¹ˀ]   appearance + degree   'behaviour' 態度

(5)   Vietnamese monomorphemic polysyllables (loanwords)
       ban công   [ɓaːn⁴⁴ koŋ͡m⁴⁴]   balcony (< French *balcon*)
       phô tô cóp pi [fo⁴⁴ to⁴⁴ kɔk͡p⁴⁵ pi⁴⁴]   photocopy (< French *photocopie*)
       Trà Vinh   [tɕa²¹ viŋ⁴⁴]   place name (< Khmer *[Preah] Trapeang*
                                                  ព្រះត្រពាំង)

By definition, monosyllabic words cannot bear paradigmatic or syntagmatic word stress. However, even in languages whose core lexicon is monosyllabic, polysyllabic words can have fairly complicated stress patterns. While Vietnamese polysyllables do not seem to

show any type of word-level prominence (Brunelle 2017), the Indic loanwords of many South East Asian languages have alternating stress systems that are not necessarily attested in their native lexicon (Luangthongkum 1977; Potisuk et al. 1994, 1996; Green 2005). For instance, polysyllabic Thai words show a tendency to alternating iambic stress, stress clash avoidance, and the application of the stress-to-weight principle, as illustrated in (6).

(6) Stress in Thai polysyllabic words
     (examples adapted from Luangthongkum 1977:199)

โทรทัศน์	ˌtʰoːrəˈtʰát	'television'
มะเร็งในเม็ดโลหิต	məˌreŋnəiˌmétloˈhìt	'leukaemia'
ไวยากรณ์ปริวัตร	ˌwaijəˌkɔːnpəriˈwát	'transformational grammar'

In these Thai polysyllables, stress is realized primarily through longer duration. The tones of stressed syllables are also realized more fully, while those of unstressed syllables are raised and partially neutralized (Potisuk et al. 1996).

Many MSEA languages also have a canonical 'sesquisyllabic' word shape, a structure typical of the region. The concept of the sesquisyllable seems to be attributable to Henderson (1952), but the term was coined by James Matisoff (1973) to designate words containing 'one syllable and a half'. Generally speaking, a sesquisyllable is a disyllable with an iambic stress pattern. Its unstressed first syllable is called the 'minor syllable' or the 'presyllable', and has a reduced phonological inventory and a limited array of possible syllable structures. Its stressed second syllable has the full array of possible contrasts of the language and can have a more complex syllable structure.

Sesquisyllables show variation across and sometimes even within languages. Thomas (1992) argues that there are four types of sesquisyllable. In the first type, a fully predictable schwa is inserted in some clusters, as in the Khmer word ក្បាល [kəɓaːl] 'head', which is underlyingly /kɓaːl/. Most authors consider such cases as monosyllables rather than sesquisyllables and treat their schwa as an excrescent vowel (Thomas 1992; Butler 2014). The second type of sesquisyllable consists of iambic disyllables in which the first vowel is a schwa and where the CʔC- sequence contrasts with corresponding CC- clusters. Examples from Jeh, an Austroasiatic language of the Central Vietnamese Highlands, are given in (7).

(7) Jeh sesquisyllables (Gradin 1966)

     trah 'to chop out'   təˈrah 'to squawk (of chicken)'
     khej 'month'   kəˈhej 'moon'

The third and fourth types of sesquisyllable distinguished by Thomas (1992) are qualitatively similar; they consist of sesquisyllables whose minor syllables can only contain a subset of the vowels that can appear in the main syllables. Examples from Northern Raglai, an Austronesian language of south-central Vietnam, are given in (8). While Northern Raglai has six phonemic vowels that contrast in length and nasality, only three are allowed in minor syllables.

(8) Northern Raglai (Nguyễn 2007)

     piˈtuk 'cough'   paˈtih 'thigh'   buˈmaw 'mushroom'

Interestingly, the trochaic mirror image of sesquisyllables, namely disyllables with an initially stressed syllable and a reduced second syllable, does not seem to be attested in MSEA.

Many languages of the area also have a non-sesquisyllabic polysyllabic structure as their canonical word shape. One example is Malay, a language that tends to have disyllabic roots but can have much longer grammatical words because of affixation or loans from Indic or Western languages. Careful analysis strongly suggests that Peninsular Malay does not have word stress (Mohd Don et al. 2008). Many Sino-Tibetan languages can also be shown to be polysyllabic because segmental or tonal processes affect their prosodic words. In Qiang and Shixing, for instance, the lenition of word-medial consonants provides positive evidence for polysyllabic prosodic words (LaPolla and Huang 2003: 31–32; K. Chirkova 2009: 12–13).

23.2.2 Tonation

Many South East Asian languages employ one or more contrastive laryngeal properties that we term 'tonation' (following Bradley 1982). This includes not only the use of pitch but also properties such as vowel quality, voice quality, intensity, and/or duration. The extent to which it is useful to sub-typologize languages according to exactly which property or properties they (canonically) employ remains a matter of some debate (Abramson and Luangthongkum 2009; DiCanio 2009; Enfield 2011; Gruber 2011; Brunelle and Kirby 2016); despite this, we have broadly organized the following sections by phonetic property in order to emphasize the diversity and phonetic variability of the region's word-level prosodic systems.

23.2.2.1 *Inventories*

Around 20% of the languages spoken in MSEA are completely atonal (Brunelle and Kirby 2015). These languages are virtually all either of Austronesian or Austroasiatic stock. Diversity is greater in Austroasiatic languages, while the Austronesian languages of MSEA are either atonal or have simple tonation-type properties.[2]

Many languages of the area, especially in the Austroasiatic and Austronesian phyla, have been described as having 'registers'. Henderson (1952) was the first author to employ the term 'register' to refer to a 'bundle' of (broadly suprasegmental) features, such as phonation type, pitch, vowel quality, intensity, and vowel duration, leading to the designation of (voice-)register languages in the South East Asian linguistic literature (Henderson 1952; Gregerson 1973; Ferlus 1979; Diffloth 1982). Register is normally understood to arise from the neutralization of voicing in onsets and subsequent phonologization of phonetic properties originally associated with voicing.

A hallmark of register systems is redundancy, in the sense that one can identify multiple co-occurring properties. The Austroasiatic language Mon is an example of a canonical register system relying on pitch and phonation, but also on vowel quality and duration (Lee 1983; Diffloth 1985; Luangthongkum 1987; Abramson et al. 2015). Another example is Wa, a Mon-Khmer language spoken in northeastern Myanmar and in the southwest of Yunnan province in China that distinguishes two lexical registers termed 'clear' and 'breathy' (Watkins 2002).

---

[2] Tsat, a Chamic (Austronesian) language spoken in Hainan, has a fully fledged tone system (Maddieson and Pang 1993; Thurgood et al. 2015).
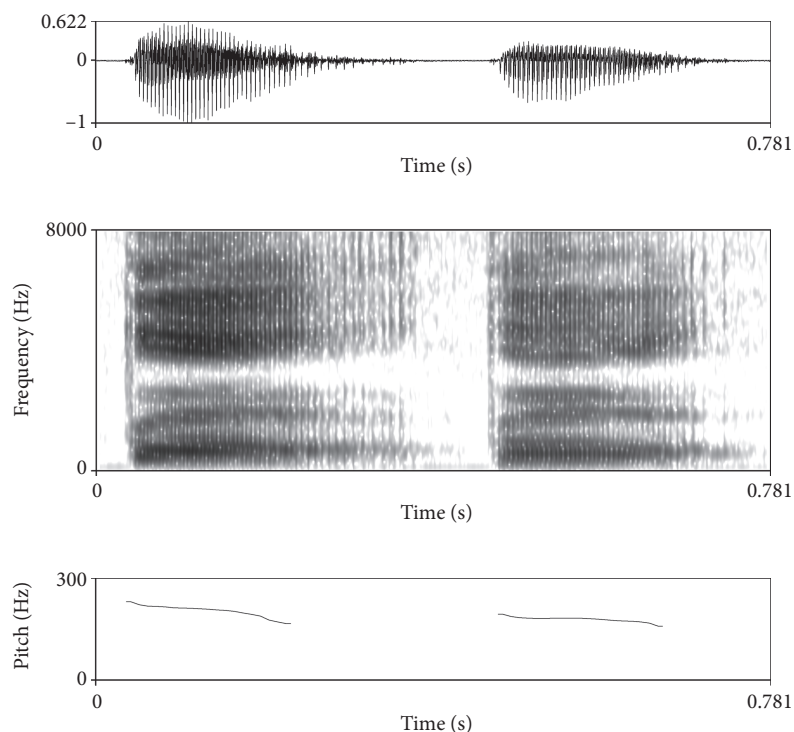
FIGURE 23.1 Waveforms, spectrograms, and pitch tracks of the Wa words tɛʔ 'land' (clear register, left) and tɛ̤ʔ 'wager' (breathy register, right). The clear register is characterized by sharper, more clearly defined formants; the breathy register has relatively more energy at very low frequencies.

In Wa, vowels in breathy register are characterized principally by their relatively breathier phonation type rather than the modal phonation of clear register vowels, illustrated in Figure 23.1. In addition, there are typically differences such that clear register vowels have slightly higher pitch than breathy register vowels. Vowel duration and vowel quality are mostly insignificant with respect to Wa register, though for some speakers there may be contrasts in these quality differences. The Wa register contrast applies independently of syllable-final /h/ and /ʔ/, making possible the set of distinct syllables in (9).

(9)   Vowel register independent of laryngeal consonants in Wa

    tɛ   'sweet'   tɛ̤   'peach'

    tɛʔ   'land'   tɛ̤ʔ   'swear'

    tɛh   'reduce'   tɛ̤h   'turn over'

An outstanding question concerns the stability of register systems, which have frequently been seen to 'restructure' (Huffman 1976) or move to realize a contrast by means of a single acoustic property. An apparently recent shift from register to a primarily pitch-based system has been documented for several dialects of Khmu (Suwilai 2004; Svantesson and House 2006; Abramson et al. 2007). Restructuring can also lead to the development of a

large vowel inventory, as apparently occurred in the history of Khmer (Huffman 1976) or Haroi (Lee 1977; Mundhenk and Goschnick 1977).

We can contrast registers with tone inventories based on pure pitch. By 'pure' pitch, we mean to refer to a system in which pitch is the only phonetic exponent of a suprasegmental tonation contrast. A good example of such a language in MSEA might be Southern Vietnamese (Gsell 1980; Vũ 1982; Brunelle 2009b). However, setting aside restructured register languages such as Khmu, it is not clear whether such systems actually exist, and, if they do, they may in fact be rather rare: it seems reasonable to assume that there are always at least low-level spectral effects present in 'pure' pitch systems. In any case, it is probably still possible to differentiate between tone systems where these spectral effects are redundant and those systems where they are a necessary element of patterns of tone contrasts, as detailed in §23.2.2.2.

A related issue here concerns the phonological analysis of primarily pitch-based tone systems. The languages of sub-Saharan Africa provide compelling evidence for an analysis based on sequences of level tones (from two, High and Low, to as many as five levels; see chapter 12; see also chapter 4). In Asia, such systems appear to be significantly less common (see Evans 2008 for an overview), though cases do exist, such as Pumi (Jacques 2011; Daudey 2014; Ding 2014) and Yongning Na (Michaud 2017: 87–101). Evidence for this type of decompositional analysis comes primarily from morphotonological alternations (see §23.2.2.2). To our knowledge, these systems are restricted to Sino-Tibetan languages of the Himalayas, on the northern periphery of the area under consideration here. Analyses of other languages of South East Asia in terms of level tones have also been proposed (e.g. Morén and Zsiga 2006 on Thai), but such proposals are challenging to evaluate in the absence of language-internal (morpho)phonological evidence (Clements et al. 2010).

Finally, MSEA is home to a number of languages with complex tonation systems involving multiple phonetic properties. While there may be a certain amount of variation, a hallmark of such systems is the canonical co-occurrence of two or more phonetic properties. For example, three of the six tones in Northern Vietnamese are systematically realized with a laryngealized voice quality in sonorant-final syllables (Vũ 1982; Nguyễn and Edmondson 1997; Michaud 2004), and perceptual research has shown that the strong glottalization of the low glottalized tone is normally sufficient for identification, to the point of largely overriding pitch cues (Brunelle 2009b). Hmong-Mien languages also tend to exhibit systems of this type (Huffman 1987; Andruski and Ratliff 2000; Esposito 2012; Garellek et al. 2013, 2014). For example, Black Miao, a Hmong-Mien language spoken in Guizhou province, China, contrasts five level tones, but three of these tones are also respectively characterized by laryngealized, tense, or breathy phonation, all of which are important cues for accurate native-speaker discrimination (Kuang 2013b). Although strictly speaking outside MSEA proper, a number of Wu languages spoken in China also have mixed phonation and pitch tonation systems (Rose 1989). These languages are perhaps especially notable for employing 'whisper' and/or 'growl' phonation types, probably involving oscillation of epilaryngeal structures (Edmondson et al. 2001).

23.2.2.2 *Tonal phonology, tone sandhi, and morphotonology*

Tone serves a wide range of functions in the world's languages: in addition to its phonemic function, it can mark grammatical categories, it can be assigned according to paradigm-specific rules, and it can even constitute the sole phonological form of a morpheme (grammatical tone; see chapter 4). In MSEA, the vast majority of Austroasiatic, Austronesian, and Tai-Kadai tone

languages have 'inert' tones (tones that are not active in phonology or morphology), whereas productive tonal processes are more commonly found in some Hmong-Mien and Sino-Tibetan languages.

The first type of tone process found in the area is tone sandhi in its narrow sense: a tone turns into another contrastive tone in a specific tonal environment. For instance, White Hmong has seven tones, out of which five undergo the permutations in (10) in most compounds and some phrases. This tone sandhi seems partly fossilized in contemporary White Hmong, but there is little doubt that it was productive at an earlier stage of the language (Ratliff 1987; Mortensen 2004).

(10)   White Hmong tone sandhi (Ratliff 1987)
        52, 22, 31ʔ   →   42
        24   →   33   / 55, 53 _____
        33   →   22

Tone sandhi must be distinguished from tonal coarticulation, which could be characterized as phonetic accommodation between adjacent tones. Studies of tonal coarticulation in Central Thai and Vietnamese suggest that progressive coarticulation is much stronger than regressive, and that assimilatory effects are more common than dissimilatory ones in these languages (Han and Kim 1974; Gandour et al. 1992a, 1992b, 1994; Brunelle 2009a). Tone sandhi could develop from the misinterpretation of some forms of tone coarticulation, but this seems to require more than simple phonologization (Brunelle et al. 2016).

The most complex sandhi-like processes in the region are doubtless found in the Kuki-Chin languages of Burma, Mizoram, and Nagaland. In these languages, combinations of tone spreading and positional tone sandhi sensitive to the boundaries of prosodic domains are commonplace (Hyman and VanBik 2002, 2004; Watkins 2013). In the Tibeto-Burman southern Chin language Sumtu, of which the Myebon dialect is described by Watkins (2013), a morpheme may have lexically high or low tone. Functional morphemes attached to a noun or verb stem may have no lexically specified tone, in which case their tone is derived by a process whereby high and low tones alternate such that adjacent highs or lows are avoided where possible (i.e. unless a lexically specified tone makes adjacent highs or lows inevitable). Examples of sentences with a lexically high tone verb stem [pék] 'give' and a low tone verb [hŋà] 'borrow' are given in (11). To the right a string of verbal auxiliaries and particles are attached, and to the left of the stem a subject/object prefix is attached. Only the verb stem has lexical tone: the attached morphemes are assigned alternating high and low tones so no adjacent tones are the same.

(11)   a.   ʔə̀-m-pék-bà-láʔ-hnì   L-H-L-H-L
              3-TR-give-again-must-PRF
              'He has had to give back.'

        b.   ʔə̀-m-pék-làʔ-hní   L-H-L-H
              3-TR-give-must-PRF
              'He has had to give.'

        c.   ʔə́-hŋà-láʔ-hnì   H-L-H-L
              3-borrow-must-PRF
              'He has had to borrow.'

In Sumtu, the dual number in verb paradigms is indicated by tone, as shown in (12). The lexically low tone verb [sì?] 'go' has minor-syllable pronominal prefixes attached. In the singular and the plural forms, these prefixes have a high tone: having lexically assigned tone, they assume the tone that is the polar opposite of the stem to which they are attached. However, the dual number is indicated by a tone change in the pronoun prefix; the low tone dual pronominal prefix provokes a dissimilatory tone change in the verb stem, so that in the dual forms the verb stem has a high tone.

(12)   Tone change in Myebon Sumtu dual number verb forms

	SINGULAR		DUAL	PLURAL
1	ká-sì?	INCL	mà-sí?	má-sì?
		EXCL	kàn-sí?	kán-sì?
2	ná-sì?		nàn-sí?	nán-sì?
3	?á-sì?		?àn-sí?	?án-sì?

A second type of tone alternation is tone spreading, a process observed in some level-tone systems: for instance, in Yongning Na (Sino-Tibetan), L tone spreads progressively ('left to right') onto syllables that are unspecified for tone (Michaud 2017: 324). 'Spreading' of level tones is a process of phonological copying; this needs to be distinguished from cases where the domain of phonetic realization of a lexical tone category is the entire phonological word, as in Tamang (Sino-Tibetan). The four tones of Tamang 'unfold' over an entire phonological word: non-initial syllables of words, whether they be a suffix or part of a single morpheme, never carry their own tone, so that their fundamental frequency (f0) curve can be considered an expression of the tone lexically carried by the initial lexeme, which is allowed to unfold over the available space—the entire phonological word (Mazaudon and Michaud 2008). This can usefully be distinguished from 'tonal coarticulation' on toneless syllables, as illustrated by Northern Mandarin, where the phonetic realization of a toneless suffix is heavily influenced by the tone of the preceding syllable but where the latter can still be considered to be realized phonetically on the syllable to which it is lexically associated (Chen and Xu 2006).

Tone can also be used to mark morphological alternations. In MSEA, this is relatively rare, except in Sino-Tibetan, where morphological alternations involving tone are most abundant in Kuki-Chin (see Ozerov 2018 for an overview and a case study of Anal) and in Na-Qiangic (Evans 2008; Jacques and Michaud 2011; Daudey 2014). Cases of morphology conveyed solely by tone (i.e. tonal morphology proper) are much rarer than cases of conditioning of tone assignment by morphosyntax (i.e. morphotonology). In Anal (Ozerov 2018), omission of grammatical suffixes leads to a grammatical distinction being marked only by tonal alternations on the last syllable of the stem. Interestingly, traces of the reduced suffix can consist of (i) changed tone, (ii) vowel lengthening, or (iii) both tone change and vowel lengthening. Another example is the Burmese creaky tone, which can express possession on a restricted number of lexemes (pronouns, kinship terms, and a few more) in place of the full possessive marker, also carrying creaky tone (Okell and Allott 2001: 273). Naxi (Sino-Tibetan) has cases of reduction of H tone grammatical words to a floating H tone, whereas M and L tone syllables that become coalescent are reported to retain a vowel target of their own—that is, the reduction process stops short of complete segmental ellipsis (Michaud and He 2007).

23.3 PHRASAL PROSODY

The phrasal prosody of MSEA languages has attracted far less systematic attention than their word-level prosody. In this section, we first review research on prosodic domains (§23.3.1). We then go over descriptions of intonational patterns and their interaction with final particles (§23.3.2) and explore the role of information structure in the languages of the area (§23.3.3).

23.3.1 Prosodic phrasing

The study of prosodic phrasing in MSEA has developed steadily over the past decade. Research has focused on the difficulty of applying the standard Prosodic Hierarchy (Selkirk 1984; Nespor and Vogel 1986) to the languages of the region. While some languages, such as Boro, faithfully conform to the hierarchy (Das 2017), a number of researchers question the very existence of a universal hierarchy, especially in the Sino-Tibetan domain, and argue for emergent domains (Hildebrandt 2007: 353–376; Bickel et al. 2009; Post 2009; Schiering et al. 2010; Michaud 2017). Lhasa Tibetan, for instance, is argued to have no phonological phrase but two word-size prosodic domains (Lim 2018).

   Most studies adopt a narrower scope and focus on evidence (or lack thereof) for specific prosodic domains (Phạm 2008; E. Chirkova and Michaud 2009; Karlsson et al. 2012; Brunelle 2016). For instance, the absence of segmental or suprasegmental processes in grammatical words argues against the existence of a prosodic word in Vietnamese (Schiering et al. 2010; Brunelle 2017; but cf. Phạm 2008). The lack of phonetic difference between homophonous compounds and phrases, such as *hoa hồng* [hwa⁴⁴ hoŋ͡m²¹] (flower + pink) 'rose' or 'pink flower', reinforces this conclusion (Ingram and Nguyễn 2006).

   To our knowledge, the issue of prosodic recursion, the embedding of a prosodic constituent within a constituent of the same type, has not yet been explored systematically in MSEA. A notable exception is Boro, a language in which a tone-spreading process suggests that enclitics are parsed into a recursive prosodic word that also encompasses the prosodic word formed around its host (Das and Mahanta 2016; Das 2017).

23.3.2 Intonation

Intonation, and more specifically the interaction between tone and intonation, has been studied in a number of MSEA languages. Although it is still too early to reach strong conclusions, it seems that boundary tones[3] can play an important role in the intonational phonology of languages with small tone inventories (Blood 1977; House et al. 2009; Karlsson et al. 2010; Karlsson et al. 2012; Phạm and Brunelle 2014). In Northern Khmu, a two-way

---

[3] The term 'boundary tone' is used as a convenient label for intonational effects that are mostly realized at the edge of intonational domains. We recognize a divergent range of views on whether these effects should be formalized as tones or as a different type of primitive (on this topic, see Rialland, to appear).

tone contrast does not prevent the realization of a phrasal H tone on the rightmost edge of every prosodic phrase; the tone curves are adjusted accordingly (Karlsson et al. 2012). A simpler example is Eastern Cham, a language in which sentence-final boundary tones concatenate with register on the final syllable, as illustrated in (13).

(13)   Final boundary tones realized on the final syllable in Eastern Cham (Phạm and Brunelle 2014): registers are autosegmentally represented as H/L for convenience

a. L   H   H   L   L L%   b. L H H   L   L H%
    | | | | | / | | | | | /
    ça ka naw paj̊ʔ çɤ   ça ka naw paj̊ʔ jɤ
    boy *name* go study already   boy *name* go study already
    'Ka has gone to school.'   'Has Ka gone to school?'

The effect of boundary tones can also be seen in languages with large tone inventories. The clearest cases are languages in which the pitch contour of toneless particles can be predicted based on intonation, such as Thai (Pittayaporn 2007), or in which an intonational contour overrides the lexical tone of discourse markers, such as backchannels and repair utterances in Northern Vietnamese (Hạ 2010, 2012). However, the typical scenario in such languages is that intonational effects are realized through a combination of various cues, such as the global pitch height and slope of the utterance, phrase-final pitch contour, and duration (Trần 1967; Đỗ et al. 1998; Luksaneeyanawin 1998; Nguyễn and Boulakia 1999; Michaud 2005; Vũ et al. 2006; Brunelle et al. 2012; Mạc 2012). It is unclear whether these intonational cues, which show great speaker variability, can be analysed as categorical boundary tones in the autosegmental-metrical sense (Michaud 2005; Brunelle et al. 2012; Brunelle 2016).

The lack of categorical realization of intonation in languages with large tone inventories could be facilitated by sentence-final particles, which are a pervasive feature of most MSEA languages. These often have the same function as intonation, arguably making it redundant. In fact, Hyman and Monaka (2011) have proposed treating such particles as a part of the intonational system. However, the existence of final particles alone does not imply that intonation is not employed, either redundantly or primarily (e.g. Dryer 2013); much more work is needed in this area.

23.3.3 Information structure

In many MSEA languages, information structure is primarily marked by means of syntactic restructuring and overt morphological markers. See Michaud and Brunelle (2016) for an overview of such markers in Yongning Na and Vietnamese. More relevant to this chapter is the prosodic marking of information structure. Although these structures have not received much attention in the languages of MSEA, they seem to mainly include prosodic phrasing and overt focus.

A Yongning Na example of information structure realized through prosodic phrasing is given in (14). In this example, 'dog meat' is topicalized and thus forms a tone group separate from the rest of the sentence, a phrasing that is marked by the bolded tone changes (see Michaud 2017: 324–327 for detailed tone rules).

(14) /kʰvˌmiˌ-ʂeˌ **dz**ɯ˥ mə˧ ɖo˧   pi˥ zo/
      dog-meat   eat NEG ought_to say ADVB
      kʰvˌmiˌ-ʂe/, **dz**ɯ˧-mə˧-ɖo˧-pi˧-zo˥
      'It is said that one must not eat dog meat! / It is said that dog meat is something one must not eat!' (Michaud and Brunelle 2016: 783)

Vietnamese is the MSEA language in which overt focus has been studied the most systematically. Studies have been conducted on corrective focus (Michaud 2005; Vũ et al. 2005; Brunelle 2017) and pragmatic focus (Jannedy 2007). Results reveal that speakers can realize focus through a number of correlates of vocal effort, such as raised f0 and intensity, increased duration, and a fuller realization of tone contours and phonation types associated to tones. However, speakers do not need to use all these cues simultaneously, and they exhibit significant individual variation. In spontaneous speech, prosodic focus is normally accompanied by morphosyntactic focus-marking strategies.

23.4 Conclusion

In this chapter, we have attempted to give an overview of the diverse prosodic systems of MSEA. We have argued that it is difficult to characterize the languages of the region in terms of a few stereotypical prosodic properties. The chapter also reflects the state of our current knowledge on the prosodic structures of MSEA: while their word-level prosody is well understood, it is imperative that more work be conducted on their phrasal prosody, which is still ill-understood.