

# ONSET-CODA ASYMMETRIES IN MISHMI AND OTHER SOUTHEAST ASIAN LANGUAGES

Seok-Chae Rhee

University of Illinois at Urbana-Champaign

## 1. Introduction

The purpose of this study is to provide a formal analysis of the onset-coda asymmetry found in Mishmi and many other Southeast Asian languages. The onset-coda asymmetry refers to different distributions of consonants in the syllable onset and coda positions. As is well-known, it is often found in many languages that only a subset of the consonant inventory appears in the coda of a given language, whereas in the onset position the distribution of consonants is relatively free. Coda neutralization also reduces possible consonant types that can appear in the surface coda. According to my investigation, Mishmi and many Southeast Asian languages show a remarkably similar onset-coda asymmetry. For an account of this, I propose an aperture node-based coda constraint, and show that high ranking of the proposed constraint covers the overall pattern of the observed onset-coda asymmetry.

## 2. Mishmi

Mishmi is one of the languages spoken in Arunachal Pradesh, the North East tip of India. It belongs to the Tibeto-Burman family of languages. According to Sastry (1984a, b), it has three principal dialects: Idu, Digaru, and Miju. Data on Mishmi in this paper are exclusively based on Sastry's fieldwork, which was conducted in 1975 and 1976 on the Digaru dialect. He mentioned that his work was cross-checked with different dialects of Mishmi.

Table (1) shows the consonant inventory of Mishmi. Mishmi has stops, affricates, fricatives, nasals, a lateral, and a trill. Aspirated stops (i.e., ph, th, kh) and aspirated affricate (ch) are analyzed as clusters in Sastry, thus, omitted from the inventory. (Without any investigation, I just follow him, since it is not critical to my analysis whether an aspirated consonant is a cluster or a singleton consonant.)

### (1) Mishmi consonant inventory

Stop:				
voiceless	p	t		k
voiced	b	d		g
Affricate				
voiceless		c		
voiced		j		
Fricative:				
voiceless		s		h
voiced		z		
Nasal	m	n		ɲ
Lateral		l		
Trill			r	

There are four types of syllables in Mishmi: V, C<sub>X</sub>V, VC, and C<sub>X</sub>VC.

### (2) Mishmi syllable types

- Nucleus type (V) /e/ 'this', /o/ 'to shoot'
- Onset-Nucleus (C<sub>X</sub>V) /gu/ 'ash', /pla/ 'salt'
- Nucleus-Coda (VC) /aŋ/ 'house', /im/ 'feather'
- Onset-Nucleus-Coda (C<sub>X</sub>VC) /sag/ 'nest', /kwag/ 'dog'

The subscript X in the onset indicates that consonant clusters can occur in that position. Trill, lateral, glottal fricative, and glides can be a second member of the clusters as seen in (3a) and (3b). Hence, aspirated stops, palatalized, and labialized consonants can occur in the onset. However, consonant clusters do not occur in the coda position (3c). This is the first onset-coda asymmetry.

### (3) a. /r, l, h/ can be a part of consonant cluster in the onset

(<sub>o</sub>[pr-, tr-, kr-, pl-, kl-, bl- gl-, ml-, ph-, th-, kh-, ch-, ...])

### b. /y, w/ can be a part of consonant cluster in the onset.

(<sub>o</sub>[py-, ty-, ky-, cy-, my-, ny-, sy-, kw-, gw-, rw-, yw-, ...])

### c. Consonant clusters never occur in the coda

(\*-pr, \*-tr, \*-kr, \*-pl, \*-ph, \*-kw, \*-sy, ...]<sub>o</sub>)

The distribution of singleton consonants also illustrates an onset-coda asymmetry. That is to say, possible singleton consonants that can appear in the coda position are strictly limited: only bilabial and dorsal voiceless oral stops /p, k/ and nasal stops /m, n, ŋ/ can occur in the surface coda positions. Fricatives, affricates, trill and lateral occur only in the onset, but not in the coda (Sastry 1984a: 58).<sup>1</sup>

(4) a. non-occurrence of fricatives and affricates in the coda

\*-s, \*-z, \*-h, \*-c, \*-j ]<sub>o</sub>

(They occur only in the onset: *su.wa* 'cap', *zap* 'red',  
*hu.lu* 'baboon', *ca.ba* 'clam', *ju.wa* 'cold', ...)

b. non-occurrence of trill and lateral in the coda

\*-r, \*-l ]<sub>o</sub>

(They occur only in the onset: *rap* 'horn', *ro* 'siege',  
*pa.ro* 'pigeon', *lim* 'fruittone', *a.lo* 'garlic', ...)

c. nasals are allowed in the coda as well as in the onset

-m, -n, -ŋ ]<sub>o</sub>

(*mi.ya* 'wife', *tīm* 'sky', *no.sa* 'picture', *phun.du* 'brinjal'  
*niŋ* 'we', *ta.ŋa* 'fish', ...)

Another noteworthy property is that final stops are always strictly unreleased in Mishmi. According to Sastry, non-release of final stops goes together with devoicing of underlying voiced stops. Due to devoicing and non-release, voiced stops and voiceless stops are neutralized into voiceless unreleased stops in the coda position. Consequently, voiced stops are not allowed in the surface coda. For example, as seen in the data set (5), final /p/ in /mabap/ 'fox' and final /b/ in /macab/ 'cow' are pronounced as same unreleased voiceless stop [p̚]. Likewise, final /k/ in /dayk/ 'language' and final /g/ in /kwag/ 'dog' are neutralized as voiceless unreleased stop [k̚].<sup>2</sup>

(5) Non-release & Devoicing

/dyap/	[dyap̚]	'greengram'
/mabap/	[ma.bap̚]	'fox'
/kab/	[kap̚]	'bulbul'
/macab/	[ma.cap̚]	'cow'
/tabab/	[ta.bap̚]	'snake'
/glebab/	[gle.bap̚]	'duck'
/dayk/	[dayk̚]	'language'
/pareyk/	[pa.reyk̚]	'frog'

/sag/	[sakʔ]	‘nest’
/hagrayg/	[ha.graykʔ]	‘flying-squirrel’
/kwag/	[kwakʔ]	‘dog’
/kacyig/	[ka.cyikʔ]	‘rat’

Non-release of final stops of Mishmi could be compared with that of English. In English, as is well-known, it is basically optional whether a final stop is released or unreleased.<sup>3</sup> So that, a final stop in an English word like ‘cat’, for example, could be released like [kʰæt̚] or unreleased like [kʰæt̚̚]. However, in other languages like French and Arabic, final stops are always released. According to Tranel (1987), a final /k/, for example, in a French word like [sak] ‘bag’ is always released. Thus, French is opposite to Mishmi in this respect. A summary of Mishmi onset-coda asymmetry is given in (6).

(6) Summary: Mishmi onset-coda asymmetry

possible onsets: p, b, t, d, k, g, c, j, s, z, h, m, n, ŋ, l, r (and  
consonant clusters described above)

possible codas: p, k, m, n, ŋ

### 3. Similar Pattern in other Southeast Asian Languages

Interestingly, a similar onset-coda asymmetry pattern is widespread in Southeast Asian language groups. Some of them are given in (7): Thadou, Cantonese, Thai, Ao, Garo, Lotha, Bantawa, Sherpa, Zhanglu Kam are good examples. Though not presented in this paper, Tagalog and some Mon-Khmer languages could be also included in this type. Even Korean, which cannot be included in the purview of Southeast Asian languages, exhibits a similar pattern (Kim-Renaud 1974). Again, final stops in the coda of the given languages are all unreleased.

(7) Onset-Coda Asymmetries in other languages

• Thadou (Krishan 1980)

Onsets: p, t, k, p<sup>h</sup>, t<sup>h</sup>, b, d, g, c, v, s, z, x, h, m, n, ŋ, mʔ,  
nʔ, ŋʔ, l

Codas: p, t, m, n, ŋ, mʔ, nʔ, ŋʔ, l

• Cantonese (Kao 1971)

Onsets: p, t, k, p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>, c, c<sup>h</sup>, f, s, h, m, n, ŋ, l

Codas: p, t, k, m, n, ŋ