The Development of Voicing Rules in Standard Burmese
The Development of Voicing Rules in Standard Burmese

Yoshio Nishi

In Modern Standard Burmese (Myanmar) are observed two distinct types of voicing sandhi: (1) in the environment $C_1C_2$, and if both $C_1$ and $C_2$ are any one of /p t t k s/, both $C_1$ and $C_2$ become voiced, and (2) in the environment where $C$ occurs after nonstop rhymes except atonic ones within a word or phrase with a postposition or postpositions, $C$ becomes voiced if it has the voiced counterpart, hence $C = /p ph t t th c ch k kl s sh/$. However, a fair number of varied exceptions are found for the first rule. The aim of the present paper is to explain the development of these two sandhi rules and the exceptions by assuming four stages for the development of three phonological rules, either obligatory or optional, or the orders of their application on the evidence from the regional dialects of Burmese.

Key Words: Burmese (Myanmar), historical phonology, Sandhi Rules, atonicization, dialects

キーワード：ビルマ語（ミャンマー語）、歴史的音韻論、サンディ規則、軽声化、方言
There are basically two types of voicing sandhi observed in Standard Burmese (SB)\(^1\), or what is often called the Rangoon-Mandaly dialect. The first sandhi may be stated as follows: when an atonic syllable C\(^{1}\)-a- is followed by an initial consonant C\(^{2}\) in the next syllable, and both C\(^{1}\) and C\(^{2}\) are either voiceless unaspirated stops, affricates or sibilants which have voiced counterparts, i.e. /p t c k s/, C\(^{1}\) and C\(^{2}\) become voiced, e.g. /gaba/ < /ka/ (< /kan\(^{2}\)/) + /pa/ {kam\(^{2}\) + ('a)p\(^{2}\)} ‘bank’, /dagaun/ < /ta/ (< /ti\(^{2}\)/) + /kaun/ {ta+ ('a)kon} ‘one animal’, /daji/ < /ta/ (< /t\(^{2}\)/) + /ci/ {s\(^{2}\)+kri\(^{2}\)} ‘headman’. However, though this sandhi rule is quite generally observed, it seems to be not obligatory throughout SB, and there are a number of variables and exceptions to this rule\(^2\).

The second type of voicing sandhi involves both the unaspirated and aspirated stops, affricates and sibilants with voiced counterparts, i.e. /p ph t th c ch k kh; s sh/. The rule is: if the initial consonant of a syllable is any one of them, and preceded by a nonstop tonal syllable in close juncture, it becomes voiced, e.g. /khayiga/ < /khayi/ + /kh\(^{2}\}/ {khar\(^{2}\)+kha} ‘charge for a journey’, /wu\(^{2}\)th\(^{2}\)udo/ < /wu\(^{2}\)th\(^{2}\)/ + /to/ {wat\(^{2}\)+tu} ‘short story’, /g\(^{2}\)ja2/ < /n\(^{2}\)+ca?/ {n\(^{2}\)+kyap} ‘five rupees’, /yej\(^{2}\)an/ < /ye/ + /ch\(^{2}\)/ {re+khyam\(^{2}\)} ‘cold water’, /takh\(^{2}\)zi/ < /takh\(^{2}\)/ (< /ta/ + /kh\(^{2}\)/) + /si/ {tackhu+ci} ‘one each’, /oun-zi/ < /o\(^{2}\)+shi/ {un\(^{2}\)+chi} ‘coconut oil’. These two rules are more fully explained in (Okell 1969: Sections 17–20), but with the first type of sandhi from a different point of view from mine\(^3\). My opinion is that since weakening or atonicization\(^4\) is unpredictable and may sometimes be optional, it should not be stated along with the voicing sandhi.

Since Standard Burmese has a series of voiced stops and a voiced sibilant in its phonemic inventory, the voiced consonants derived by these sandhi rules, especially, the second one, used to be transcribed as such. However, since its voicing is predictable by the rules, nowadays voicing is often left unrepresented in phonemic representation.

The development of these sandhi rules, together with atonicization, may be explained comprehensively by assuming four successive stages in the historical...
development of three voicing rules and atonicization. In the first stage, the voicing rule was applicable only to the nonaspirated stops and affricate and the sibilant (the first voicing rule)\(^5\). In the second, some of non-initial syllables became atonic (atonicization), and we may also assume, especially from alternative spellings in Old Burmese (OB) inscriptions, that there had already been many words with weakened and/or atonic syllables though their vowels may not have been restricted to /a/ as in SB. In the third, the second voicing rule followed it, which was only applicable to C\(^1\)- (the second voicing rule). Finally, in the fourth, a new voicing rule appeared which applied only to the aspirated stops and affricates (the third voicing rule). It should be noted that atonicization could be applied repeatedly to more than one nonfinal syllable, hence /kalabye/ \(<\)/kalà/ + /pye/ {kulà\(^2\)+pre} ‘Indian country’. Presumably, it had also been working even after the third voicing rule came into effect, which may have brought into being forms like /shàdò/ \(<\)/shà/(\(<\)/shan/\)+thò/ {cham+thui\(^2\)} ‘hair-pin’.

As far as the development of voicing in Burmese is concerned, no reliable evidence is offered by inscriptions. However, we observe the alternative and regular use of voiced letters in -gri -gri -grih and -bri -bri as variable spellings for two morphemes, kri ‘great’ and pri ‘to finish’, in the Myazedi inscription, both of which are spelled with voiced letters in two other early OB inscriptions (Lokahteikpan (ink writings) and Ngathilattin)\(^6\) as well, when they occur in the second syllable\(^7\). This fact suggests that the first rule of voicing was already emergent at the turn of the twelfth century. Unfortunately, this orthographic representation of voicing later seems to have been abolished in the process of the orthographic standardization of OB, which eventually eliminated morphophonemic alternations. To continue my speculation further, assuming that this holds true, it also suggests that voicing would not be distinctive in early OB\(^8\). Whatever the truth may have been, little evidence is supplied by inscriptions alone, the development or nondevelopment of voicing sandhi in regional dialects may be called on to support the above assumption to some extent.

To date we have been informed that, roughly speaking, there are seven regional dialects distinct from Standard Burmese, or the Rangoon-Mandalay dialect in Burmese, that is, Yaw, Arakanese, Tavoy, Mergui, Intha, Danu and Taungyo. Now, some reliable data on voicing sandhi, though hardly sufficient for some, have been provided for all these dialects: Yaw (Yabu 1980), Arakanese (Okell 1995; Bradley 1979), Tavoy (Nishida 1972; Okell 1995), Mergui (Nishida 1972), Intha (Okell 1995), Taungyo (Yabu 1981a) and Danu (Yabu 1981b)\(^9\).

Of these seven dialects, no voicing sandhi is observed in Intha, Taungyo and Danu, at the level of phonemic representation. They have no voiced stops or sibilants as phonemes, except Taungyo, where /z/ occurs\(^10\). However, Yabu reports that the voiceless unaspirated stops and affricate of Taungyo
have half-voiced variants, while the voiceless unaspirated stops and affricate, and the voiceless sibilant of Danu have voiced variants. The half-voiced or voiced variant occurs exactly where the second type of voicing sandhi is observed in SB. On the other hand, Arakanese, Tavoy and Mergui all are reported to have only the first voicing rule in the same environments, while Yaw has the first and second voicing rules, but they may not apply there in a number of cases where voicing occurs in SB. In Intha, Okell mentions that no voicing occurs.

The different degrees of voicing which developed in these dialects may represent the stages of development of the first voicing sandhi in SB. Intha represents the earliest stage in which no voicing occurs yet; Taungyo and Danu the next stage, in which the voicing or tendency to voicing of the voiceless unaspirated stops and affricate and the voiceless sibilant is in the emergent state, but it is still a redundant feature; and Arakanese, Tavoy and Mergui the first stage, in which the first voicing rule occurs, while Yaw has reached the third stage. It may be surmised that the development of voicing in early OB, supposing that my interpretation of inscriptional evidence is right, was probably in this second stage or in a transitional stage from second to third.

If the above assumptions hold true, we may claim that the voicing of the unaspirated and aspirated series of consonants occurred independently of each other, and the voicing of the unaspirated series (the first voicing rule) preceded that of the aspirated one (the third rule). It is also assumed that the two types of voicing sandhi in SB proceeded by the following steps. First, (1) voiceless unaspirated stops and affricates and the sibilant became voiced due to the first voicing rule, then (2) the preceding syllables were atonicized (atonicization), and finally (3) the voicing of $C^1$- (the second voicing rule) took place. (4) The voicing of the voiceless aspirated stops and affricate (the third voicing rule) must have been the last to appear. However, as atonicization was optional and still working at the time when the third voicing rule was introduced, the order between (3) and (4) may have been reversed, differing from subvariety to subvariety. Also, the second voicing rule may have never been introduced in some subvarieties, hence /kabà/ ‘bank, cliff’, /tągaun/ ‘one animal’ and /tajì/ ‘headman’ (cf. /gabà/, /dągaun/ and /dađi/ above) and the first voicing rule never applied to the earlier s which eventually changed to /t/ in SB in some subvarieties, hence /natò/ /nà/ (< /nwa̰/ ) + /tò/ {nwa$^2$ + (‘a)sui$^2$} ‘bull’, varying with /nadjò/. By so assuming, we can offer a possible solution to many of the variables and exceptions cited in Sections 19 and 20 in Okell (1969).

There remains still a fair number of exceptions and variables that must be explained among those cited in Okell (1969). Thus, we may have to consider the effect of Written Burmese on spoken varieties in some cases where no voicing rule applies to $C^1$, as well, to set up an arbitrary unordered optional rule such as the de-aspiration rule applicable only to the initial of atonic syllables to explain variable forms like /kabaʔ/ {khà-pat} “belt”, /tabın/ {than$^2$-pań} “tod-

256
dy Palm", /sabin/ {sham-pan} “hair”, /sɔdɔ/ {sham-thui²} “hair-pin”, or to attribute such examples as /njapi/ {nà²-pi} “fish paste” and /ŋacaunša/ {nà²-kron-lhyā} “kind of fish” to earlier borrowing from a distinct dialect or dialects, or from a different language or languages with Burmanization. However, there still remain a few exceptions, such as /hnakaun/ `two animals’, /khunnacaʔ/ ‘seven kyats (rupees)’ and /bɛhnape/ ‘how many feet’, which stubbornly resist any attempt at explication by phonological rules. My guess is that it may possibly be caused by analogy with the proclitic negative adverb /ma-/ with a nasal initial, the weakening of whose nuclear vowel, later atonicized, presumably dates back earlier than they did.

It is now also clear that though the first sandhi voicing or the tendency for it is a very conspicuous feature in the development of SB and some regional dialects, it is not a feature traceable to a common origin, but must be regarded as only a result of drift or convergence.

Notes

1) Standard Burmese here is not a well-defined variety. I follow the general use of this term, referring to both spoken and written forms which are used in a very wide, but not clearly defined, area from Upper Burma down to Lower Burma, usually alluded to as the Rangoon-Mandalay dialect. But it is quite certain that it contains a wide range of variations or subvarieties. The variations found for the types of sandhi here discussed may represent the differences among the subvarieties, both social and regional. However, they may be considered to belong to the same dialect as against all the regional dialects mentioned in the text. Though Burmese is one of the two rare languages which enjoy the status of the state official language (together with Dzongkha of Bhutan) among more than 250 Tibeto-Burman languages, it can hardly be said to have been studied worth its status. Thus, there is virtually no systematically recorded data on variations in SB. Fortunately, we have reliable and fairly extensive data of all the known regional dialects, provided by a number of scholars, and among all, we owe most to two scholars, J. Okell and S. Yabu, though there remains much to work on for them yet.

2) I noticed the voicing of C’ as something strange some time in 1964 or 1965 when I had daily contact with my Burmese friends in Japan, because it yielded forms different from those recorded in the glossary compiled by Cornyn and Musgrave (1958), the only Burmese-English lexicon with phonemic transcriptions at the time, and later I thought of this sandhi rule. When I wrote a Burmese writing system (in Japanese), I added it as another type of sandhi observable in SB in a schematic form there. This small, though exhaustive, writing system was never published, and only a few copies were distributed some time in 1968 or 1969. (It is, however, referred to in (Nishida 1972) and in several works of S. Yabu.) It is also interesting to note that the voicing of C’ is apparently officially recognized as the standard in contemporary SB since the Myanmar-English Dictionary (1993), compiled by the Myanmar Language Commission, in which each Burmese entry is followed by a phonemic transcription (at the level of traditional phonemics) of the Burmese form, adopt forms derived from this sandhi rule.

3) These sandhi rules are more fully described in Sections 17–20 by Okell in his Reference Grammar of Colloquial Burmese. As I think that the description of types of voicing sandhi is probably exhaustive, covering all the variables and their exceptions, and his Grammar is easily accessible to most readers, I have quoted all the examples in this paper from that work.

4) Okell (1967: 14) uses the term ‘weakening’ to cover both the simplification of syllable structure (to å [=a here]) and atonicization. Thus, he says: ‘When this [weakening] occurs, the first of two joined syllables loses its tone; i.e. the vowel (and final consonant if there is one) is re-
placed by the toneless vowel'. However, in the context of the Burmish languages in general, weakening (of syllables) and atonicization must be distinguished. In Achang (Dai and Cui 1985: 22–23) and Zaiwa (Xu and Xu 1984: 17–19), weakening affects only the nuclear vowel and the length of the syllable, with a number of exceptions in which the initial consonant also undergoes a change, while the tone always remains intact. For this reason, I preferred to use atonicization for his ‘weakening’ in this paper. Though it is conceivable that weakening theoretically may precede atonicization as an intermediate stage, no such examples have been reported in Burmese.

In this connection, we note that except for /myama/ {mran-ma} ‘Burma, Burmese; Myanmar’ and /byama/ {brahma} ‘Brahman’ (<< Pāli brahma), with three derivatives of the latter, both of which are registered in the *Myanmar-English Dictionary* (1993), no other atonic syllables have initial clusters. However, there are atonic syllables with written clusters with velar initials. This may suggest that atonicization occurred after what may be called the Great Consonant Shift in SB (probably between Early/Mid-18C to Early 19C) which includes the change of those clusters to affricates.

5) It is sure that SB /s/, /sh/ and /t/ derived from the earlier affricates tc (WB c), tch (WB ch) and s (WB s) through the stage of depalatalization, hence ts and tsh, and that of fronting, hence ð, and we may infer that these changes occurred during the Early Modern Burmese period (Nishi 1976; 1997), which eventually restructured the consonantal system of Middle Burmese as a whole. However, since no evidence for when the first voicing rule was introduced in SB has been found so far, the statement here remains partly ambiguous. On the basis of some contemporaneous records by Europeans, Bradley suggests that the change s>ð took place around 1780 in Burmese and in Arakanese after 1798, while the changes ts, tsh, dz (Burmese) and tc, tch, dz (Arakanese) to s, sh, z occurred in both Burmese, Arakanese, and others after 1798 (Bradley 1979). This, I believe, is a possibility, but, assuming that OB and the language of the *Mian-dian-guan* Zazi represented the earlier forms of SB, about which I have no doubt, c, ch must have been alveolo-palatals; cf. (Nishida 1972: 51) and (Nishi 1997: 991, fn. 6).

6) The date of the Lokaheteikpan inscription is not known, but it is estimated as nearly contemporaneous with the Myazedi inscription. The former is actually not a single inscription, but a collection of ink writings on the walls of the Lokaheteipan pagoda, but, for convenience, I call them here collectively as an inscription. The date found on the Ngathilattin inscription is B.E. 482 (A.D. 1120).

7) It should be noted, however, that no example of the use of voiced letters is found in [Maung] Chitsagon { [Mon] Khyac-ca-kon²} plaques, votive tablets excavated from a mound in a field whose owner’s name was Maung Chitsa/Khyitsa, on the east side of the famous Ananda Pagoda, Pagan, from 1926 to 1927. These plaques are presumed to be prior to the reigns of Kings Alaungsithu {‘Alon-caña-sú} (1113–1165) and Narapathisithu {Narapasi-caña-sú} (1174–1211) on various grounds, and it is conceivable that they may more faithfully reflect the popular pronunciation of the pre-Standardization period. All examples of kri there are spelled as such, and not with the voiced letter g. Apart from Ba Shin (1962), there are two other readings of the same plaques, by U Mya (1961) and Than Tun (1964). Unfortunately, not all their readings agree. For instance, it is only Ba Shin who reads -h in them.

Of the two instances of voiced letters given here, bri ~ bri is used as either a full verb or a subordinating marker, apparently derived from the verb, which corresponds to SB/pi/ ‘after, and’, cf. (Okell 1969: 173ff.). The use of a voiced letter for a full verb in such an environment is not known in SB. Thus, we may suspect that it affords evidence for an argument for the existence of a voiced series in OB, but no other evidence can be invoked for it. There are a few more words, including one place name, which are spelled with a voiced letter in the environments where voicing sandhi is expected in the Myazedi and Lokahetteipan inscriptions. However, since they are found only once each in one of the inscriptions mentioned here, I left them out here.

8) One of the most controversial subjects in the history of SB is the development of initial voiced phonemes. Some of them are simply loans of various origins, but some probably resulted from the loss of the first syllable after the sandhi voicing of the second syllables. Three of the ex-
amples of such origin suggested by Benedict (1972: 22, fn. 75) are persuasive. We may add at least one more similar example to them, that is, **SB/jó/ {WB khyui²}** (but cf. **OB 'u-khruiw**) ‘crust (of cooked rice)’. Nishida also suggests a similar view, but gives no example (Nishida 1972: 291, fn. 29). We may add some more analogous examples, but there are far more examples which cannot be explained as such. Burling (1967: 38), on the other hand, proposes that the voicing of initial consonants resulted from assimilation to voiced word-medial consonants, but, apart from his mistakes in two of his quoted examples, this view is quite unlikely. In this connection, it is interesting to note that Okell (1995: 105–106) gives some twelve examples which show a change of the earlier **m/-hm-** (one instance) to **b**- among the morphemes with their initial clusters corresponding to **WB mr-/-mhr-**. This change, if it is not an ongoing change, is a sporadic, or uncompleted change in the sense that not all **m**’s of this type of initial cluster show the same change, remaining as **m**-. The change of **m**- to **b** is not an isolated example in Tibet-Burman. The example of Bisu is well-known, and, interestingly, it seems to me that here again the change is sporadic though Bradley considers it as a conditioned change. It may be surmised that some of the initial voiced consonants in **SB** were introduced by such changes, and it is interesting to note that many voiced initials in **SB** derive from the earlier initial clusters.

Another important contributor to the voiced initials is Pāli loans, with Mon as the mediator. It is conceivable that the voiced series in fact existed in the subsystem of **SB** throughout the **OB** and Middle Burmese (**MB**) period, as Pāli, and possibly Sanskrit for a short time, may have held an exalted position as an extra-High, or at least a Dummy High, Variety among Buddhist monks and the literate class. Mon, which was the mediator of both writing and Buddhism, presumably occupied the status of High Variety for a short time in the early **OB** period. However, the derivation of **SB** forms from the corresponding **OB** and **MB** loans is quite often a highly complicated matter. I have already shown two such examples in (Nishi 1975: 8–9, fn. 8).

9) Some of these dialects are reported to have subdialects as well as regional variations. Arakanese is said to have three subdialects, Rakhain (=Akyab-Mrohaung), Kyaukphru and Rambre, and Marma is another branch of Arakanese (Bradley 1979: 2). Yabu proposes to divide the Taungyo dialect into two subdialects, Major Taungyo (the variety spoken by the Nīman {Nan-man} branch of the Taungyo people) and Minor Taungyo (the variety spoken by the two other branches, Yathu {Ra-sū} and Pyinhu {Pran-sū}), listing their differences in some detail (Yabu 1981). The dialect referred to here is Minor Taungyo. He also mentions some regional lexical variations in Yaw (Yabu 1980). Similarly, Okell observes some dialectal differences within the Intha (‘An²-sā²) area and ‘much greater differences’ outside (1995: 55). Nishida refers to a variety of the Tavoy dialect spoken in a small island, not far from Mergui, but, apart from lexical variations from the major variety spoken in Tavoy, the phoneme inventory of both varieties is the same (Nishida 1972: 295).

10) In fact, a slight complication is brought in by the introduction of the voiced sibilant /z/ in Minor Taungyo. It has three sibilants /s/, /zh/, /z/. The voiceless ones /s/ and /sh/ correspond to **WB c** and **ch**, respectively. However, word-initially, /za/ ([za: ~ dzə:]) occurs only in interrogatives; /za/ ‘what’, /zah/ ‘who’, /zah/ ‘which’, /zal/ ‘how’, /zama/ ~ /tama/ ‘where (place)’, /zahal/, /zahal/ ‘why’, /zal (/zal/) ʔ ‘how much’, /zak/ ‘where (goal)’, /zakha/ ~ /takha/ ‘when’, /zamah, /zamakh/ ~ /tamakh/ ‘how many’ and /zamani/ ~ /tamaani/ ‘what time’, which all include /za/ ([za: ~ dzə:] in his Taungyo data, and all other occurrences of this sibilant is due to the sandhi alternation that occurs word-medially. It is probably related to Arakanese /za/ ‘what’ and is apparently a cognate with Danu /sa/ ‘what, which, where’. Thus, the voicing of its initial may be secondary in this dialect. There is no distinction that parallels the use of /ba/ and /be/ in **SB**, a distinction generally observed in most dialects except Taungyo and Danu. As for Mergui, interrogatives are not given in (Nishida 1972). Interestingly, /s/ [sh] in Major Taungyo corresponds to **SB/t/ (/: WB s), /s/ (/: WB ch), and /s/ (/: WB rh, lhy), while /ts/ corresponds to **SB/s/ (/: WB e). Minor Taungyo shows changes which exactly follow the pattern of **SB** here; **WB** s, c, ch, rh/lhy : /t/, /s/, /sh/ and /s/. The merger of the earlier s and tsh (<ch=tsh) parallels that which occurred in Intha.
Okell (1995: 63) maintains that Intha once had /t/, derived from the earlier s, but later it changed to /sh/ under Shan influence. The same theory may apply to Major Taungyo. However, since Minor Taungyo appears to be more innovative and similarly under Shan influence, the picture may not be as simple as his theory suggests.

**BIBLIOGRAPHY**

Ba Shin  

Benedict, P. K.  

Bradley, D.  
1979 *Arakanese vowels*. (12th Sino-Tibetan Conference.)

Burling, R.  

Dai, Qingxia and Cui Zhichao  

Nishi, Y.  
1975 *OB ry- ni-tsuite* (Old Burmes ry-: a remark on PLB resonant initials). *Shiroku* (Kagoshima University, Japan) 8, 1–13.


Nishida, T.  

Okell, J.  


Than Tun, U  

U Mya  

Xu, Xijian and Xu Guizhen  

Yabu, S.  
