

# Khasi Clusters and Greenberg's Universals

Eugénie HENDERSON<sup>†</sup>  
University of London

In his well-known 1965 paper on initial and final consonant sequences, Greenberg, in what was clearly a preliminary search for 'universals' or 'quasi-universals' in this field, proposes a number of what he cautiously terms 'generalizations' regarding initial and final consonant clusters based upon a sample of some 100 languages. Khasi is cited on two occasions as seeming to offer counter-evidence to certain rules of cluster formation which are otherwise almost universally applicable to his sample. In both instances, Greenberg notes that the phonetic description in his sources is not sufficiently detailed for him to be able to decide whether the exceptions are valid or not. It is the purpose of this paper to provide further phonetic information which demonstrates that Khasi does indeed supply an exception to the general rule in one case, but that in the other, there is no such exception.

To take the latter case first: Greenberg states that an earlier observation of Trnka's concerning the non-occurrence of sequences which differ in only one feature is valid on the whole for sequences involving 'differences in laryngeal adjustment such as voicing, voicelessness, and glottalization.'<sup>1</sup> He goes on to note that Khasi initial *td* appears to be an exception here. Not so. Khasi /t/ and /d/ differ in two features: both in laryngeal adjustment, voice vs. voicelessness, and in the place of articulation, dental vs. alveolar. Khasi /t/ is dental whereas /d/ is alveolar, as in some other languages of the South East Asian area, such as Vietnamese, Khmer, Malay, N. Sulawesi. It is our Western habits that incline us automatically to assume that /t/ and /d/ are a pair, differentiated only by the presence or absence of voice. Careful phonetic observation demonstrates that this is not the case in Khasi, which has a strong tendency to avoid homorganic clusters.<sup>2</sup> Khasi therefore does not afford an exception to the general rule in this instance. Examples such as *tdong* 'tail' and *tnat* 'twig' (where *n* is alveolar) exhibit heterorganic initial clusters, consisting of segments differentiated from each other by *two* or *three* features respectively.

The second instance quoted by Greenberg concerns his generalization No. 21 which states:

Except for voiced nasal followed by homorganic unvoiced obstruent,  
an unvoiced consonant or sequence of unvoiced consonants in initial

---

<sup>1</sup>A wider generalization made by Trnka about the non-occurrence of sequences of phonemes differentiated by a mark of correlation was earlier observed by Trubetzkoy to be invalid as far as nasal and voiced homorganic stops were concerned.

<sup>2</sup>For a fuller account of Khasi clustering patterns see Henderson 1976.

systems immediately preceding a vowel is not itself preceded by one or more voiced consonants.

Khasi is once more cited as an exception. The generalization excludes initial clusters such as *bs-*, *bt-*, *dk-*, *mt-* etc., whilst permitting clusters like *ps-*, *nt-*, *tk-*, *mp-*, *tg-* etc.

Underlying this generalization would seem to be the idea that there is a tendency to have the voiced consonants of a cluster closest to the vowel, i.e. adjacent to the voiced nucleus of the syllable. This may indeed be a tendency, but is not a universal. Khasi has initial clusters *bt-*, *bth-*, *bs-*, *bsh-* [bʃ], *dk-*, *dkh-*, *dp-*. In fact, there seems in Khasi to be a deliberate *dissimilation* of voicing in initial clusters, especially when the cluster is of two stops. Thus, alongside *bt-*, *dk-* and *dp-*, we can have *tb-*, *kd-* and *pd-*, but *pt-* and *tk-* are reserved for a few onomatopoeic and what one might call 'expressive' words only. The general word stock of the language prefers voice dissimilation, viz.:

bti	'to lead by the hand'	pdot	'throat'
bthi	'sticky'	pdeng	'middle'
dkar	'tortoise'	tbian	'floor'
dkhar	'plainsman'	tba	'to feel'
dpei	'ashes'	pjah	'cold'
bsa	'to feed'		
bshad [bʃa:t]	'civet'		

Contrast:

tkor-tkor	'plump and tender'
ptej-ptej	'(used of something very sticky)'

As for the suggestion in generalization 21 that the only permitted sequences of voiced followed by a voiceless consonant consist of a voiced nasal followed by a homorganic unvoiced obstruent, Khasi once again appears to provide the exception with such clusters as *mt-* and *mth-* in *mtian* 'to drift', *mthin* 'strong and stout', *mtung* 'in a large heap', though it might be argued that these words are 'expressives'.

Readers familiar with Khasi will think of apparent exceptions to the alleged Khasi preference for voice dissimilation in obstruent clusters, such as *kthi* 'hand', *kpa* 'father', and other similar forms. Cluster initial *k-* is an exception because Khasi has no [g] sound. The stop system is the familiar South East Asian one:

k	t	p
kh	th	ph
-	d	b

We thus get *k* in clusters where we would otherwise expect *g-*, and there are as a consequence occasional contrastive pairs such as *kpa* 'father', *kba* 'wheat'. This is clearly not the place to embark upon a discussion of the phonological history of

Khasi; suffice it to say that some understanding of historical processes may help us to make sense of apparent anomalies in the synchronic system.

Greenberg speculates that the Khasi 'exceptions' may perhaps be explained away by a fuller phonetic description; there may perhaps, he suggests, be a *svarabhakti* vowel between the two stops, or the first of the pair may be an 'unvoiced lenis', although written with *b* or *d*. As regards the first suggestion, it is true that the Welsh missionaries who first romanized the language wrote *kypa*, *kyba*, *dypei*<sup>3</sup> etc., implying the presence of a short unstressed vowel between the two consonants, and it is indeed difficult for Westerners to pronounce such clusters without an intervening vowel, just as it is difficult for Japanese or Hindi speakers to cope with English clusters like *sp-*, *st-* without the support of a vowel. Khasi speakers, however, manage pretty well, although it is of course not possible immediately to halt, for example, the vocal cord vibration of a voiced consonant without the chance of some overlap with a following voiceless one, but this overlap is minimal with Khasi speakers in most instances, and is certainly less than would be found with a 'true' short vowel. I have tried out some Khasi clusters on the Frøkjær-Jensen combined oscilloscope and mingograph, with attachments.

At Figure 1 are the results for *kpa* and *kba*.

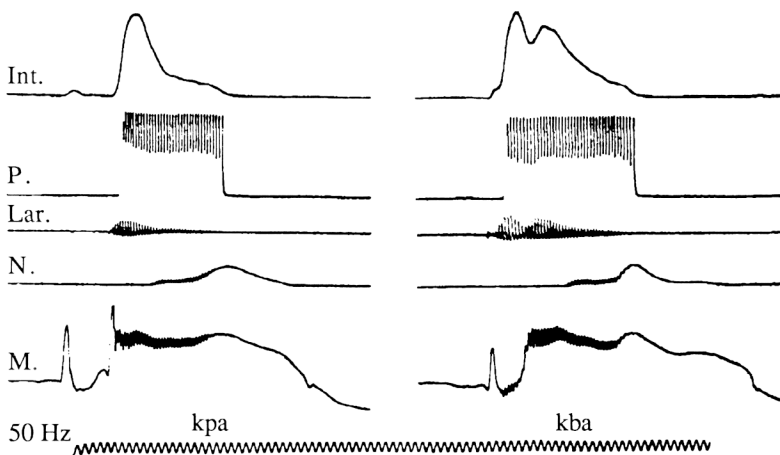


Figure 1

<sup>3</sup>The generally accepted practice nowadays appears to be to mark no intervening vowel between a 2-consonant cluster, but to insert a *y* in a 3-consonant cluster, usually after the first consonant.

For those unfamiliar with the equipment, there are six channels whose functions will be very briefly indicated here. The top line (Int.) measures *intensity*. The second line (P) shows *pitch*, with the fundamental frequency following the bottom of the striations. Then comes the larynx tracing (Lar.), showing the vibrations of the vocal cords. The fourth line (N.) shows the nasal tracing; an olive is held in the nose to record vibration in the nasal passage. The mouth tracing (M) indicates variations in the air pressure of the stream of air issuing from the mouth. The bottom line is the time marker, running at 50 Hz.

The mouth tracing for *kpa* shows a sharp excursion for the *k* followed by a stronger one for the *p*-, with no voice vibration on the larynx tracing until the onset of the vowel, i.e. *no* intervening vowel. The tracing for *kba* shows the same sort of overall pattern, but with voicing beginning before the release of the *b*, i.e. during the stop itself.

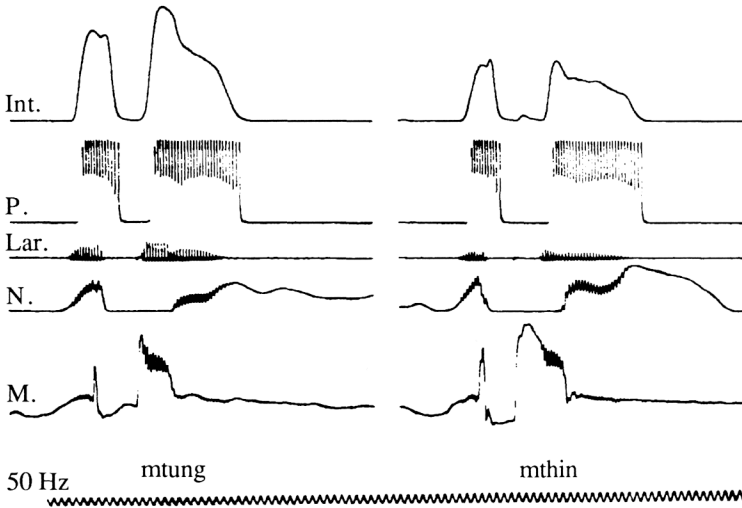


Figure 2

Compare *mtung* and *mthin* on Figure 2. In the tracings of both words there are large excursions for the *m* and the plosive, with voicing and nasalization clearly visible for the *m*, followed by an interval with neither, during the closure phase of the plosive. Then in *mtung* voicing (for the vowel) follows immediately after the release of the *t*, while with *mthin* there is a short interval with no voicing, which corresponds to the aspiration of the *th*, before vowel voicing begins.

At the top of Figure 3 we have the tracings for *bti* and *bthi*. In both cases we have two excursions for the plosives, with voicing shown clearly on the larynx tracing for the *b* (no voiceless lenis here!) There follows an interval of no voicing,

(i.e. no vowel), until after the release of the *t*. In *bthi* there is a short period of aspiration between the release of the second plosive and the following vowel.

At the bottom of Figure 3 is the tracing for *bdi*, a rare word for 'twenty', which is the only word in the language with two voiced initial plosives. Notice how the larynx tracing shows voicing throughout the word, as contrasted with the interrupted lines for *bti* and *bthi*.

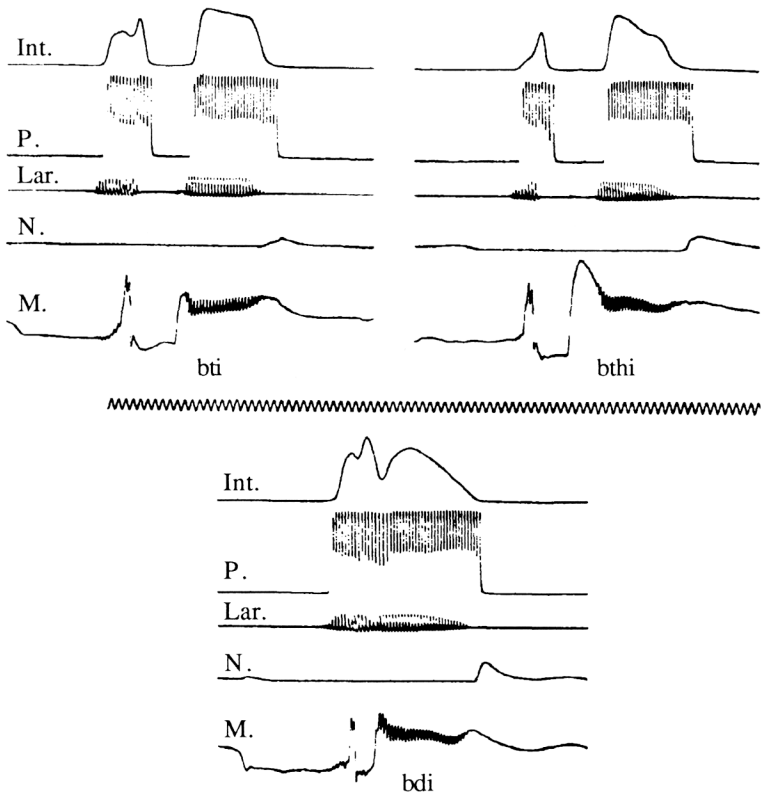


Figure 3

I have no explanation for this isolated breach of the Khasi voice dissimilation rule by the word *bdi!* What is interesting, of course, is that it should be the form which *fits* Greenberg's generalization which calls for explanation in Khasi, rather than the exceptions. Dissimilation of voicing and of place of articulation is the general rule in Khasi clusters, unusual as this may be in the languages of the world.<sup>4</sup>

## REFERENCES

- Greenberg, Joseph. 1965. 'Some generalizations concerning initial and final consonant sequences'. *Linguistics* 18, 5-34.
- Henderson, Eugénie J.A. 1976. 'Khasi initial clusters'. In Philip Jenner, Laurence Thompson and Stanley Starosta, ed., *Austroasiatic Studies*, 1, 523-38. Honolulu: University of Hawaii Press.

Received: Helsingør revised

---

<sup>4</sup>It is worth noting that such language-specific rules may also be relevant to historical studies. To take one simple example, the Khasi heterorganic rule is important for the understanding of the reflexes of the Austroasiatic prefixes in the language. We may expect to find the common velar prefix before kinship terms, body parts, animal names, etc. *only* if the root itself does not begin with a velar. Cf. forms *kpa* 'father', *kmie* 'mother', *ksew* 'dog', *kti* 'hand', etc., but *tnga* 'spouse', *dngiem* 'bear', *dkar* 'tortoise', etc.