

A PHONETIC DESCRIPTION OF TIBETAN
with a review of the literature.

Olle Kjellin

The purpose of this paper is to give an outline of the phonetics and phonology of an idiolect representative of Lhasa Tibetan, as the author has interpreted it. The author does not belong to any specific phonological school, but his treatment is profoundly influenced by the theories of generative and natural phonology. The treatment proposed here differs widely from all previous descriptions and treatments proposed by other investigators of Tibetan. Moreover, these other treatments also differ very much from each other. Therefore, it was felt that it would be useful for the reader first to have summaries of a few earlier treatments.

Since the author's basic aim of investigating Tibetan was to observe experimentally the possible connections between consonant types and tone in a purported tone language (Kjellin 1976), the review and the description will necessarily be centered about the said aspects of spoken Tibetan.

1. Review.

As concerns the "standard" dialect of Lhasa, there is unanimous agreement in the literature that this is a tone language in which every lexical item has to have a suprasegmental tone indication, feature or even "phoneme", as for instance in Chinese. The number and shapes of the tones vary with investigator and informant, but many descriptions partly agree with each other. The same can be said about the descriptions of the segmental sounds, consonants and vowels. However, very few have paid any particular attention to the fact that there appears to be some connection between the consonant voicing and clustering on the one hand, and the "tones" on the other.

Most descriptions of Tibetan concentrate on the writing and the "pronunciation of letters", and they have to be reviewed with that fact in mind. In many if not most cases, however, it turns out that the orthography, which has hardly changed at all since it was constructed after Indian models in the 7th or 8th century A. D., to a surprising extent resembles the lexical, or "underlying", forms of words and syllables which have to be posited in order for the correct speech forms to be derived (or "generated") not only in regard to the "tones" but also in other aspects of the phonology, such as Umlaut, devoicing, etc. Incidentally, similar justifications of English orthographic forms can be found via the phonological theory of Chomsky and Halle (1968).

The greatest influence on Tibetan linguistics during the last 85 years has been exerted by Jäschke and his dictionary of 1881. In the introduction he claims that: "A system of tones has been introduced under manifestly Chinese auspices" in the central dialects, with the modification that "only the first principles of what are known as the high and low Tones, have made their way into Tibetan." However, two lines below he actually does mention that "the Tone is determined by the initial consonant of the word", although he might be referring to the letter. Thus, using the orthographic forms in

the examples, he claims that the distinction between ཨ' (sa) and ཨ' (za) is found "not in the consonant, but in the Tone, pronouncing ... ཨ' [sa] with a high note, and ... ཨ' [za], on the contrary with a low note, and, as it appeared to me, more breathed and floating", the last clause possibly referring to some breathiness.

There is no reason to believe the "Chinese auspices", which, if at all, would rather affect the Eastern dialects bordering on and mingled with Chinese dialects. Jäschke evidently did not notice that the distribution of dialects having "Tone" and devoiced initials coincides with the distribution of dialects which have lost their consonant clusters, although he is careful to describe this latter distribution also, and the relations to written Tibetan. Thus he gives the pronunciation wsgrags in both Balti (Western periphery) and Khams (Eastern periphery) for the written form bsgrags (where Lhasa pronounces drAa, i. e., [qaa] with low-high-low pitch).

Then, with a four-page "phonetic table" to compare written forms and the pronunciation of the corresponding words in six dialects, Jäschke concludes "that for scientific objects the pronunciation (sic!), as it is given in Column 1 [the writing], is the most suitable. ... its system is regular enough to render it unnecessary to give the pronunciation of every individual word throughout the work [dictionary]." A very early generative phonologist, indeed! Why have the implications escaped attention up to this day?

P. M. Miller (1951) recognizes "two suprasegmental phonemes of tone", (high and low), and claims that every syllable "may be assigned one or the other of these ... inherent tones" with the exceptions that syllables beginning with the initial consonants h, r, voiceless r, and voiceless l "appear to be atonic", and that those beginning with the voiceless aspirated affricate tsh "appear to be monotonemic: ... high tone". These "exceptions" of course, are very strange. The present author has not found any evidence for such claims.

P. M. Miller correctly notes that: "In syllables with final m or ŋ and in those lengthened by the presence of 'zero' variants of final consonants the tone may exhibit a glide at the end of the syllable," but he believes that "This glide may be slightly up or down in anticipation of the higher or lower pitch of the succeeding syllable." In the Tibetan observed by the present author the final "glide" is up for sonorants, and down for obstruents, whether they are represented by "'zero' variants" or not and irrespective of the pitch of the following word.

As concerns consonants and consonant types, P. M. Miller's "phonemes" do not differentiate between voiceless unaspirated stops (which are followed by a high-pitch vowel) and voiced stops (which are followed by an initially low-pitched vowel), but believes this difference in voicing is due to the difference in tone, that is, low tone is supposed to cause voicing, or, conversely, high tone is supposed to cause voicelessness. However, neither possibility is very likely: Apart from the phonetical evidence (fundamental frequency high after voiceless, low after voiced consonants; cf. e. g. Lehiste and Peterson 1961), Hyman and Schuh, drawing on vast evidence from phonological processes involving tones and consonant types in several African languages, showed that "consonants affect tone, but tone does not affect consonants" (Hyman and Schuh 1972; Hyman 1973). Even E. Pike, in a survey, cites this claim with approval (Pike 1974).

Yet the belief that initial voicing in Tibetan depends on the tone is wide-spread in the literature: R. A. Miller (1955) does not state it explicitly, but he makes no distinction between the consonant "phonemes" for the "class of high tones" and the "class of low tones"; also Badaraev (1967), Simon (1970), Goldstein and Nornang (1970), Yukawa (1971), and Kitamura (1974) adopt the same view in their analyses, which are essentially very similar to P. M. Miller's.

To give a specific and representative example of this kind of analysis that of Yukawa (1971) will be presented. Yukawa mainly used the same informant as the present author.

Monosyllabic words consisting of a short open syllable have either of two tones, viz. high level, denoted by $\bar{\quad}$, e. g. $\bar{\text{so}}$ (tooth); or low rising, denoted by \prime , e. g. $\prime\text{ča}$ (tea). Long-syllable monosyllabics (i. e. those having a long vowel or being closed by a consonant) are (1) high level, denoted by $\bar{\quad}$, e. g. $\bar{\text{šaa}}$ (east); (2) low rising, denoted by \prime , e. g. $\prime\text{čaŋ}$ (north); (3) high-low [falling], denoted by \backslash , e. g. $\backslash\text{mii}$ (eye); or (4) low rising-falling, denoted by \wedge , e. g. $\wedge\text{kaŋ}$ (snow). The latter two in isolated pronunciation frequently have a glottal stop instead of, or within, the last mora.

In disyllabic words, the first syllable, whether short or long, is always high level or low level; the second syllable has either of the (1), (2), or (3) types, syllables corresponding to (4) being neutralized with (3). [Kitamura 1974, in cooperation with Yukawa and using the same informant, further claims that (2) is neutralized with (1), so that the second syllable of a word obligatorily begins on the high pitch. (Present author's note)]

The tonal types are summarized in the following table:

Monosyllabics		Disyllabics (pitch contour of each syllable shown within brackets [])			
short	long	short } short	long } short	short } long	long }
(a) $\bar{\text{O}}$	(c) $\bar{\text{O}}$	(a) $\bar{\text{O}}\bar{\text{O}}$	[$\bar{\text{O}}\bar{\text{O}}$]	(c) $\bar{\text{O}}\bar{\text{O}}$	[$\bar{\text{O}}\bar{\text{O}}$]
(b) $\prime\text{O}$	(d) $\prime\text{O}$	(b) $\prime\text{O}\text{O}$	[$\text{O}\bar{\text{O}}$]	(d) $\prime\text{O}\text{O}$	[$\text{O}\prime\text{O}$] (Kitamura: [$\text{O}\bar{\text{O}}$])
	(e) $\backslash\text{O}$			(e) $\backslash\text{O}\text{O}$	[$\text{O}\backslash\text{O}$]
	(f) $\wedge\text{O}$			(f) $\wedge\text{O}\text{O}$	[$\text{O}\wedge\text{O}$]

[Adapted translation from Yukawa (1971) p. 156 f.]

Based on this description, two alternative treatments in the generative framework were given in Kjellin (1974), with tones as suprasegmentals and without reference to consonant types.

The way the consonants are generally treated in the literature will be exemplified with Yukawa's description of the labial stops and the laryngeals:

" (1) /p/ in initial position is a voiceless aspirated labial stop; in tones beginning with high pitch it is voiceless throughout the aspiration, but in tones beginning with low pitch the last half of the aspiration is voiced. Word medially between vowels the closure is extremely weak and, further, most of the aspiration part is a voiced sound; occasionally it may appear as a sound near [β]. Medially after consonants the closure is not weak but the aspiration part, again, is quite voiced. (These differences can be sufficiently well explained as differences between environment features affecting the same phoneme.)

" (2) /b/ in initial position is an unaspirated labial stop accompanied by a not so strong laryngealization; in tones beginning with high pitch it is almost completely voiceless, but in tones beginning with low pitch only the beginning is voiceless. In syllable-initial position word medially it is almost a voiced sound, and the closure does not weaken." [Trans. Yukawa(1971)p.142]

After a lengthy discussion, Yukawa concludes that the syllable-final labial stop belongs to the phoneme /b/ rather than /p/.

The laryngeals:

" (1) ... Syllables having /h/ always begin with a high pitch. This is an extremely low-frequency phoneme.

" (2) When a Tibetan syllable seems to begin with a vowel, does one have to think that there is some consonant phoneme there? One may say roughly that when such a vowel stands in word-initial position, in tones beginning with high pitch it begins with a glottal stop, but in tones beginning with low pitch, it begins with a soft vocal Ansatz...." [Translated from Yukawa (1971) p. 145]

Then again after a lengthy discussion he decides not to accept a phoneme /ʔ/, although, of course, he does not dispute the existence of the phonetic "glottal stop" and "gradual beginning".

This kind of solution, however, evidently leads to difficulties, whether one believes in "phonemes" or not. In languages like English or Japanese, where there are only two types of vowel onset, namely the 'ordinary' and the voiceless (that is, h), some researchers posit some consonantal entity, such as "glottal stop", for the 'ordinary' type, others do not even mention it; but nobody forgets the h. In Tibetan, on the other hand, there are obviously three types of vowel onset, namely unaspirated (glottal stop), aspirated (h), and voiced ("gradual beginning"). The first two occur with high pitch, the last one with low pitch. In a discussion aiming at finding "phonemes", it must be entirely arbitrary which two to pair as "allophones", e. g. h/h, and blame the voicing of one on the low tone; or ?/h and blame the 'hardness' of the one on the high tone; or do as Yukawa (and many others actually do), recognize only one, 'accidentally' occurring with only one of the

two pitch types. Why could not ʔ have been selected as well, instead of h, and then the h/f difference attributed to the pitch differences? Actually, Yukawa nearly does so in the case of a final offglide which he observed after word-final open syllables, saying that phonetically after a vowel ending on the high pitch level there is an h, but a vowel ending on the low pitch level is often closed by a glottal stop or by nothing at all (a 'gradual end'?). Of course, he makes no "phonemes" out of these, though as we have seen, one is arbitrarily selected as an initial "phoneme".

Turning next to those descriptions in the literature which do mention the connection between consonant type and pitch type, one of the most ambitious is that of Sedláček (1959):

"... the distinction between high and low tones depends predominantly on the kind of the initial with which a syllable begins. The same regularity may be observed e. g. in Ancient Chinese where the voiceless initials... always carried with them high tones, while the [voiced] initials... and the smooth vocalic ingress without explosion... always had low tones. In Lhasa Spoken Tibetan, the voiced initials have changed substantially into voiceless ones, so that, e. g., the distinction between the syllables ša and ža lies exclusively in their different tones, i. e., the former is pronounced in the high falling tone... and the latter in the low rising tone ..."

Although this sounds a promising outset, he goes on to say that:

"The object of this study is to examine the tones and their changes in all possible syllables beginning with three categories of letters of the Tibetan alphabet, namely prescribed, superscribed, and postscripted letters..."

and this will also be the objection to his study. Sedláček used magnetic tape recordings of reading-pronunciation forms of isolated syllables of the Tibetan alphabet and also a number of compounded forms, listened to them and transcribed them. As a result he obtained a set of rather unbelievable tones. Direct consultation with one of his two informants revealed that most of Sedláček's reported tones are more or less incorrect. So unfortunately his ambitious paper cannot be used as a reliable source of information.

Unfortunately again, the same must be said about the next most ambitious treatment, that of Richter (1964). Richter used tape recordings ("in ihrer Qualität ausgezeichnet gelungene Tonbandaufnahmen") of a language course, not even recorded by himself but by his professor, in Peking 1955. In general, Richter is less incorrect than Sedláček (they both used the same informant, whom the present author has also had occasion to work with), but although the title of Richter's work implies a phonetic description, he pays more attention to Written Tibetan and the pronunciation of letters, even in the section entitled "Phonetische Varianten".

It is from this vantage point he observes that "For the formation of the pitch contour of a syllable, the consonantal environment of the vowel...

is of crucial importance. Depending on whether the initial consonant is voiced or voiceless, the pitch contour of the syllable varies correspondingly."

Richter further claims that each syllable has "its inherent pitch contour", which "results from the (Written Tibetan [Richter's parentheses]) initial and final of the syllable". Thus Richter reports four tones, (a) mid-high flat (1 2), (b) falling (1 23, or 1 34), (c) (mid) rising (1 43), and (d) mid-low flat (1 4), each being "required" by certain letters. (Note that Richter's tone indexes are inverted, using the lowest number for the highest tone). In the rest of the work, however, one finds several additional, unexplained tone indications.

Also Badaraev (1967), in his extensive and excellent review, probably is referring to the writing, when, on page 128, he says that,

"... all initials are divided into the initials that form low-pitch tones... and the initials that form high-pitch tones... The fundamental aspects of the tones - flat, rising, falling - evidently depend on and derive from the finals and other factors".

He denies his own words in the discussion of linguistic (phonological) change on page 179, where he believes that the tone "must not be referred to the consonants nor to the vowels".

Rerix (Roerich), the influential Russian Tibetologist, in his posthumous monograph (1961), passes over the subject of tones with just a few lines, saying that "just as in Old Chinese, syllables with a voiceless initial is pronounced with high tone, and syllables with a voiced initial with low tone." But he claims, further, that syllables with clusters (of course he is referring to the written-language correlates of the spoken language, which does not possess initial clusters on the surface) are pronounced with high tone; this is not always true. Moreover, Rerix claims that syllables ending in -g have the high rising tone, and this is never true. Besides, he only reports three tones, viz. "rising, flat, and falling", and this might indicate that he observed another dialect than what he believed he did (that of the Central Tibetan provinces U and Tsang, with Lhasa and Shigatse as the main cities).

Free from orthographic influence, the most reliable source is probably Sprigg (1955), where the numerous and well-transcribed examples compensate for the obscurity of the peculiar Firthian phonological treatment. In this paper on the tones, the first mention of them is in footnote 3 on the twelfth page, namely, I. High flat, II. Low flat, III. Falling, IV. Rising-falling, V. Rising. However, most of the article deals with criteria for the delimiting of words, and it is with that aim in mind that Sprigg sets up a "tonal system" stated "not in relation to the syllable but to the word". Simultaneously, "it is ... possible to go a long way towards delimiting the word by utilizing for the purpose certain of the phonetic features that characterize the initial and the final consonants of syllables (C-, -C) in inter-syllabic junction". So, as criteria for the tonal system he posits pitch, duration of vowel, word-initial features, and voice quality (clear voice for

high tone and breathy voice for low tone). Had not Sprigg locked himself up in Firthian theory, he would have been able to show twenty years ago what the present author is trying to show now, namely that the "tones" in Tibetan are not lexical but derive from features of the consonants. Instead, he goes the opposite way, and so in Sprigg (1972) he even claims that "initial-consonant clusters in Tibetan, or rather, Tibetan syllable-initial features generally, are better treated not segmentally but unitarily, as applicable en bloc to monosyllabic lexical items, providing an, as it were, supra-segmental scheme of classification with the same status as tonal classification."

Apart from Sprigg's own conclusions, his treatment comes the closest to showing that Lhasa Tibetan is not a tone language.

There are several other authors and articles that would deserve mention, but this review shall close with a quotation from Gleason (1961). Gleason contrasts tone with intonation (so that, in his definition, a non-tonal language is an "intonational language"). Classifying tonal languages into four types depending on intonational "overriding", he includes Lhasa Tibetan into the least tonal type, with the following statements:

"Words in citation form have clear lexical tone and many minimal pairs can be found. Indeed Tibetan is classically listed as a tone language. In utterances, however, tonal distinctions are very nearly leveled out, and a sentence intonation system takes its [sic] place... Only in very restricted places can any of the tonal contrasts be found. On a sentence level, Lhasa Tibetan is in no sense a typical tone language." (Gleason, 1961, p. 300.)

In the present author's opinion, the "minimal pairs" mentioned by Gleason are only spurious, surface-phonetic pairs. In the lexicon there are no minimal pairs where tone is the distinctive feature, because "tone" need not be posited at all in the lexicon. Therefore, on the lexical level, Lhasa Tibetan is in no sense at all a tone language. This standpoint will be outlined in the following section.

2. The Present Author's Treatment

2.1. The Informant

The informant was an adult male native of Lhasa. Between seven and twenty five years of age he lived and studied at a temple (Ngor Gongba) near Shigatse in the Tsang province to the West of Lhasa. Then he left Tibet and came to Japan, where he stayed until 1975 as a researcher on the Tibetan language, culture, and Lamaism. He considers his own idiolect representative of the Lhasa dialect, although it must be suspected that it might contain elements of the Tsang dialect. The two dialects, however, are very similar.

2.2. The Language

2.2.1. Syllables

Tibetan is a syllabic language, where each syllable has its own meaning. Syllables can be compounded two and two to obtain derived meanings. Spurious polysyllabic words can be shown rather to be phrases. Each syllable has stress (in the wide, generally used but poorly defined sense). Only certain function words (clitics) are unstressed.

No syllable can begin with a vowel; thus the minimal syllable is CV. The maximal surface syllable is CVC (cf. non-surface syllables, below). Often a CVV syllable occurs in more or less free variation with CVC; "more or less free" depending on the nature of the final consonant. The final vowel in such a case is an elongation of the preceding vowel, but it retains certain important features of the replaced consonant, notably the pitch specification. In maximal compounds, i. e., CVCCVC, the main variants are CVCCVV and CVVCVC. That is, the syllable-final consonant replaced by a vowel is either that of the first syllable or that of the last syllable. Occasionally both are replaced: CVVCVV.

Syllable-initial and syllable-final consonants seem to belong to different sets, and they will consequently be treated separately.

2.2.2. Surface Initial Consonants

The word-initial sounds which have been observed in surface-phonetic form are summarized in Table 1. This table also introduces a proposal for the transcription of spoken Tibetan. It is proposed that capital letters be used in the representations of sounds that are associated with the higher of the two basic fundamental-frequency (F_0) levels (referred to here as the HIGHF target), and that lower-case letters be used in the representations of sounds that are associated with the lower F_0 level (referred to here as the lowf target; the F_0 targets and related F_0 contours are treated separately below).

Letter combinations in the table represent unit sounds, and in many cases the choice of symbols reflects the author's opinion as to the phonological derivations of the particular sounds. However, the combination of n and g for the velar nasal is merely traditional and arbitrary, as is also the use of h to indicate aspiration and e. g. ts for affrication. The rows and columns of the table have the following contents:

Columns:

- (A) labial stops and sonorants (i. e., nasals and glides);
- (B) dentals, including laterals
 - 1. stops and sonorants,
 - 2. fricatives,
 - 3. affricates;
- (C) retroflex sounds;
- (D) palatalized sounds
 - 1. affricates and certain sonorants similarly derived from dentals
 - 2. fricatives,
 - 3. glides,

- 4. stops derived from velars;
- (E) velar stops and sonorants;
- (F) laryngeals.

Rows:

1. HIGHF voiceless unaspirated...
 - a. nonsonorants (i. e., obstruents and voiceless glides),
 - b. nasals,
 - c. glides;
2. HIGHF voiceless aspirated...
 - a. nonsonorants,
 - b. nasals,
 - c. glides;
3. lowf voiced...
 - a. nonsonorants,
 - b. nasals,
 - c. glides;
4. lowf voiceless obstruents.

	(A)	(B1	2	3)	(C)	(D1	2	3	4)	(E)	(F)
1a.	P	T		TS	TR	TY			KY	K	?
b.	M	N				NY				NG	
c.	W	L			R			Y			
2a.	PH	TH	S	TSH	THR	THY	SY		KHY	KH	H
b.	MH					NHY				NGH	
c.		LH			RH						
3a.	b	d		dz	dr	dy			gy	g	
b.	m	n				ny				ng	
c.	w	l			r			y			f
4.	ph	th	s	tsh	thr	thy	sy		khy	kh	

Table 1. Surface-phonetic word-initial consonants.
(For legend, see above text.)

2.2.3. Surface Medial Consonants

Medial consonants, i. e., the initials of the second syllable of compounds, show no distinctions as concerns HIGHF or lowf even in surface-phonetic forms; they are all perceived as being associated with high F_0 . Nor do they usually show the triple distinctions of voiceless-aspirated, aspirated or voiced. Actually, they are more or less voiced, or at least unaspirated, unless the particular word is pronounced very slowly and 'carefully'. This even concerns the stops and affricates of rows 1 and 2 in Table 1; those of row 4 are generally voiced even in slow speech, and those of row 3 are, of course, always voiced. For these reasons the phonological exposition below will only deal with non-medial consonants, supposing the medial ones can be derived through some simple rules reducing the number of distinctive features in them.

2.2.4. Surface Final Consonants

The consonants which may occur in final position are very few compared with the total inventory, viz.

(1) b g m ng M NG R

with the same notation as in Table 1. These respective finals are phonetically almost the same whether absolute word-final or followed by the second syllable of a compound. There is some disagreement in the literature whether the final, generally unexploded stops b and g are voiced or voiceless. Oscillograms unequivocally show them to be voiced, even when the following word begins with s. The phonological distinction in finals, therefore, is not that of voiced-voiceless, but of sonorant-nonsonorant. Even so, as was mentioned briefly above, the consonantal articulation of the finals may be lost leaving an elongation of the preceding vowel in its place. This elongated part of the resulting long vowel will be called an "alpha-vowel", because it retains several features of the consonant it replaces, notably the HIGHF or lowf specification. So, for instance, if the preceding "nucleus" vowel is HIGHF and the lost final consonant is lowf, this is heard as a "falling tone" on that syllable, e. g. MAa (from MAg 'down').

This alpha-vowel conversion seems to be more or less optional: it affects velars more frequently than labials, and R almost always. Thus e. g. maR or maA 'butter'.

If the word-final C is an obstruent it is sometimes replaced by a glottal stop-like sound, especially in phrase-final position.

2.2.5. Surface and Lexical Vowels

Eight oral vowels and their nasalized variants have been observed. They are listed in Table 2.

	<u>Oral</u>		<u>Nasal</u>		
i	ü	u	ĩ	ũ	(ũ)
e	ö	o	ẽ	õ	(õ)
	ä	a	ǣ		(ǣ)

Table 2. Surface-phonetic vowels. Diaereses are used as in German to indicate fronted vowels. (For legend, see text.)

There seems to be a phenomenon that resembles vowel harmony, which, however, will not be treated here.

The nasal vowels enclosed in parentheses occur in free variation with the combination a o u + labial or velar nasal. Nasal vowels are always long. Long vowels in general were treated in 2.2.4. together with final consonants, as being derived from a syllable-final -VC sequence. Therefore, the nasal vowels will be treated as phonological sequences of an oral vowel+nasal consonant.

The fronted, or unlauded, vowels also are always long and never occur with final consonants. This suggests that in fact they are the reflexes of some lost final consonants. Since the set of conceivable dental finals *d n N l is missing from the series of final consonants, the unlauded vowels can be suspected to derive from sequences of a back vowel plus dental consonant, and will be treated as such here. Hence the lexical forms of Tibetan syllables will only contain the five vowels,

(2) i e a o u.

Consequently, the list of final consonants (1) has to be revised for the lexical representation, but this question will be postponed until the lexical initial consonants have been determined.

In principle, the vowels have the higher of the two vocal pitches, and in one word there is always at least one vowel or part of a vowel having this HIGHF target, even if all the rest of the word is lowf. Therefore, a vowel having such a HIGHF target will be transcribed in the upper case; e. g., mA, maA, etc. (see section 2.2.9.).

2.2.6. Lexical Initial Consonants

The phonological treatment of segments as well as of tones is greatly simplified if certain generalizations are made from Table 1., and a much smaller set of lexical, or underlying consonants posited. The proposed set is given in Table 3., where it should be noted that there are no special indications at all concerning tones or pitches. That is, "tone" does not exist in this lexicon. The derivations of tone-like pitch contours will be explained below.

	(A)	(B 1	2	3)	(C)	(D)	(E)	(F)
1.	p	t		ts			k	ʔ
2.	ph	th	s	tsh			kh	h
3a.	b	d	z	dz			g	
3b.	m	n					ng	
3c.	w	l			r	y		fi

Table 3. The lexical syllable-initial consonants. Columns as in Table 1. Rows: 1, voiceless unaspirated nonsonorants; 2, voiceless aspirated nonsonorants; 3, voiced (a) obstruents, (b) nasals, (c) glides. Note that the upper-case/lower-case graphemic distinction is unnecessary here.

In Table 3, compared to Table 1, column D is reduced to the glide *y*. All the other surface palatal forms will be regarded as phonologically /C+y/. In the same way the retroflex consonants are regarded as /C+r/ (column C). This solution is supported by the orthography also; but there seems to be no evidence for splitting the affricates into /C+s/, the aspirated stops into /C+h/, or the like.

A conspicuous difference between Table 1 and Table 3 is the small number of rows in the latter. It is proposed that row 4 of Table 1 be derived from row 3a of Table 3; that is, the surface lowf voiceless aspirated consonants are the result of a process which devoices the lexically voiced obstruents in certain environments (see below). Thus /b/ → ph, /d/ → th, the newly posited /z/ → s, etc. This is suggested by the fact that these ph th s, etc. have the lower vocal pitch associated with them instead of the higher. In word medial position they are very seldom devoiced, and at least not aspirated, as was said above about medial consonants. Further, positing the lowf voiceless group of consonants as an independent type in the lexicon would predict the existence of a lowf voiceless *h, i. e. a lowf cognate of H. Such a sound, however, has not been observed.

The MH listed in Table 1, a voiceless [m] followed by high vocal pitch, is only found in certain assimilated (?) occurrences of the negative prefix /ma-/: If the verb following /ma-/ has as its initial in the lexical form a consonant of type 1 (Table 3), this /ma-/ will also be pronounced in the higher vocal pitch, i. e. MA- (surface type 1b; Table 1. But if the verb in the lexical form begins with a consonant of type 2 (Table 3), this /ma-/ will also be pronounced as HIGHF aspirated in the surface form, in effect devoiced, i. e. MHA-. Even so, the initial consonant of the verb root itself, being word medial in the actual pronunciation, will of course usually be unaspirated or even voiced.

However, if the verb following /ma-/ begins with a consonant which, if there was no negative prefix, would belong to surface type 3 or one which would belong to surface type 4, then the negative prefix will be pronounced in the lower vocal pitch and always voiced, whereas, of course, the verb

itself, being the second syllable of the word, will be in the higher vocal pitch, as was said above about the second syllable. Examples:

- (3) KOR 'boil' (transitive aspect) MAKOO 'do not boil'
KHOR 'boils' (by itself) MHAKOO 'does not boil'
- nyO 'buy' maNYO 'do not buy';
thyE 'do' maDYE 'do not do'

In this way, MH may be considered to be derivable by some simple rule. The other two aspirated nasals are derivable by some similar, partly morphological rule; peculiarly enough they are found in the following words only:

- (4) NHYEe 'to get tanned' (intransitive aspect) (cf. neg. MHANYEe)
(cf. also NYEe 'to tan (a hide)' (transitive; neg. MANYEe)
NGHÖö 'to get frizzled (in oil)' (intrans.) (cf. neg. MHANGÖö)
(cf. also NGÖö 'to frizzle (in oil)' (trans.); neg. MANGÖö)

As concerns the other surface HIGHF sonorants (types 1b, 1c, and 2c, Table 1), it is proposed that they derive from lexical type 3b and 3c sonorants taking part in consonant clusters. (Independent evidence that syllables may have consonant clusters in the lexical representations are given in 2.2.8.). It is suggested that these sonorants are HIGHF because they are not absolutely word-initial, but preceded by a "prefix" (see section 2.2.8. for this term). When the cluster is simplified - i. e., the "prefix" is lost - the distinction between lexically prefixed and nonprefixed words will remain. This would suit the fact that all other non-initial sonorants are HIGHF, whereas the initial ones are lowf.

In the case the lexical "core" consonant (see 2.2.8. for this term) is b, d, or g (type 3 stops), the presence of a lexical prefix prevents it from being devoiced, while non-cluster lexical /b d g/ are proposed to undergo the devoicing process. Thus, again, the distinction between lexically prefixed and non-prefixed words is retained on the surface, but this time without having to resort to F₀ change.

These processes and their purported effect on the vocal pitch were treated in some detail in Kjellin (1975b), q. v.

2.2.7. Lexical Final Consonants

The following consonants are posited as lexical finals. Some of them, especially /z/, often function as morphological suffixes.

- (5) b d g m n ng r l z

Among these finals, the dental ones never appear as such in surface forms, but cause vowel Umlaut and convert into alpha-vowels (see 2.2.5.).

As has been mentioned above, the phonological distinction in the finals is not one of voiced versus voiceless (they are all voiced), but rather one of sonorant versus obstruent. For the pitch specification, this entails the effect that final sonorants get the higher of the two pitch levels, and obstruents get the lower.

The occurrence of the final z is purely hypothetical, though it finds support in the orthography. It occurs as a second member of a final cluster, where the first member is a sonorant, specifically a nasal. This nasal is assimilated to the lowf of the z, and then the z itself is lost in the general cluster simplification process. Hence the lower-case nasals in the list of surface finals (1), section 2.2.4. The orthography (which uses the letters s and d in complementary distribution and not just a simple z) would suggest that the z also might occur after the final obstruents, but there is no synchronic evidence for this in the observed idiolect. Since no syllable may begin with a vowel, there will never be any chance for z to "transmigrate" across a word-medial syllable boundary and "survive", as the "prefixes" can do.

2.2.8. Consonant Clusters

There is evidence that two consonants can combine and form clusters in the same syllable. In the actual spoken language such clusters will never be heard word-initially, but they are found medially if the first syllable in a compound is open, i. e., if it has an open slot in place of a final. Then, if the second syllable begins with a cluster, the first member of these two consonants will move across the syllable boundary to become final in the first syllable:

- (6) CV&CCV → CVC&CV
 (where & represents syllable boundary.)

In the composition of clusters, the lexical initials as given in Table 3 are regarded as the "core" initials, or root initials, which may be preceded by one of certain consonants from a limited set. This limited set of pre-initials ("prefixes" in the terminology of traditional Tibetology, for often, but not always, they have morphological functions) happen to be almost the same consonants as those which may occur as finals:

- (7) "prefixes" : b d g m ng r l z
 lexical finals: b d g m n ng r l z

What is common about these two sets is voicing, that is, they are all Type 3 consonants.

Positing clusters is supported by the orthography, which, however, has more complex clusters than are necessary for the synchronic phonology. A distributional constraint is that homorganic clusters are not allowed. That is, for instance, g may not precede /g/, /k/, /kh/, or /ng/, but it may precede labials and dentals. (A problem might arise in the case of /ng/. The orthography has /ŋ/, not /ng/, but when it is actually pronounced in the middle of words it functions like an /ng/, often also nasalizing

the preceding vowel without umlauting it. But this orthographic /h̄/ may precede orthographic /ng/, and orthographic /ng/ itself is not allowed as a "prefix". However, no attempt will be made to solve this rather philological problem within the scope of the present paper.)

The following examples illustrate the evidence for positing clusters in the lexical forms. Upper/lower-case letters for the HIGHF/lowf distinction is only used to the right of the arrows (the actually observed pronunciation forms), since this distinction is considered to be absent from the lexicon.

(8) /btyu/	→ TYU	'ten'
/gtyig/	→ TYIḡ	'one'
/btyu & gtyig/	→ TYUGTYIi	'eleven' (with "alpha-vowel")
/bzyi/	→ syI	'four' (note that "at least one vowel or
/btyu & bzyi/	→ TYUBSYI	'fourteen' [part of vowel must be
/bzyi & btyu/	→ syibTYU	'forty' [HIGHF])
/rgu/	→ ḡU	'nine'
/btyu & rgu/	→ TYURḡU	'nineteen'
/rgu & btyu/	→ ḡubTYU	'ninety'
/lŋa/	→ NGA	'five'
/btyu & lŋa/	→ TYÖÖNGA	'fifteen'

The last example in (8) also shows the effect of vowel lowering (some kind of vowel harmony ü→ö) before /a/. The umlaut here, of course, is due to the dental final /l/ received from the second syllable in the compound; this is quite strong evidence for the hypothesis that cluster-making "prefixes" do exist, and that they "transmigrate" and survive when the syllabification is altered. This is similar to the gnostic versus agnostic alteration in English, further exemplified by word pairs such as knōwledge - acknowledge and pterygoid - helicopter. This process may not be synchronic in English, but in Tibetan it is, and it occurs very frequently.

2.2.9. "Tones"

As is clear from the many implications above, no tones are needed in the lexicon. Instead, the pitch levels and movements between the levels are derived (Kjellin 1975b). The crucial factor is whether the initial consonant is voiced or not. If the initial consonant is voiced or derived by devoicing from a voiced obstruent, the pitch target is low for that consonant. If the initial consonant is voiceless, the pitch target at that point is high. The special case of "not absolutely initial" sonorants was explained in section 2.2.6.

For the final consonant, if there is one, the crucial factor is whether it is an obstruent or a sonorant. If it is an obstruent, the pitch target is low; if it is a sonorant, the pitch target is high. A vowel is a sonorant, and its pitch target is high. In a word consisting of one long syllable (CVV)

with the initial consonant having the lowf target, this lowf target is mapped onto the following vowel, which thus also becomes lowf. The assimilation process stops there, always leaving at least one HIGHF target in any given word, e. g., maA. If the final segment also has the lowf target, no mapping can occur; e. g., mAa. There seems to be no "reason" why there should be a HIGHF element in the middle of such a word, or why the mapping process should not carry on to the end of the word, but it may be suggested that this is a requirement for the stress to be pronounced and perceived. The actually occurring target combinations (pitch-contour types) are summarized in (9).

- (9) mA maA mAa MA MAA MAa

The actual F_0 contours of these syllables were given in Kjellin (1975a and 1976).

In disyllabic words, the principles are exactly the same as in monosyllables, with the extension that the lowf target mapping process carries on throughout the first syllable up to the syllable boundary. The following types can occur:

- (10) MAMA MAAMA MAMAA MAMAa MAAMAa
 maMA maaMA maMAA maMAa maaMAa

The rise always seems to take place in the boundary between the syllables. Thus the second syllable is always perceived as beginning on the high pitch level, even though the initial consonant of the second syllable is usually voiced. This is compatible with the finding that "The central values of F_0 in a voiced consonant... immediately before the second vowel... were also considerably higher when the second syllable was stressed" (Lea 1973, p. 58; observations on English). Even all the investigators who think that Tibetan is a tone language have noted that the second syllable is "non-distinctively high". It may be no coincidence that consonant voicing and aspiration also are non-distinctive in medial position.

This concludes the outline of the author's proposed treatment of Lhasa Tibetan.

References

- Badaraev, Bal-Dorzhi (1967), *Ob osnovax transkriptsii i transliteratsii dlja tibetskogo jazyka*. Burjatskij institut obshchestvennyx nauk, Ulan-Ude.
- Chomsky, Noam, and Morris Halle (1968), The Sound Pattern of English. Harper and Row.
- Gleason, H. A. Jr. (1961), Review. *Language* 37:2, 294-308.
- Goldstein, Melvyn C., and Nawang Norngang (1970), Modern Spoken Tibetan: Lhasa Dialect. University of Washington Press, Seattle.
- Hyman, Larry M. (1973), The role of consonant types in natural tonal assimilation. In: Consonant Types and Tone, ed. Larry M. Hyman; Southern California Occasional Papers in Linguistics, No. 1 (Los Angeles: Linguistics Program, University of Southern California).

- Hyman, Larry M., and Russel G. Schuh (1972), Universals of Tone Rules. Evidence from West Africa. Working Papers on Language Universals 10, 1-50. (Revised version in Linguistic Inquiry 5:1, 81-115, 1974).
- Jäschke, H. A. (1881), A Tibetan-English Dictionary with Special Reference to the Prevailing Dialects. Reprinted, London 1958.
- Kitamura, Hajime (1974), The Pronunciation of Modern Tibetan. Tokyo University of Foreign Languages, Institute of the Languages and Cultures of Asia and Africa, Tokyo. (In Japanese).
- Kjellin, Olle (1974), A Preliminary Report on Tibetan Tones. Annual Bulletin, RILP 8, 115-123.
- Kjellin, Olle (1975a), Progress Report on an Acoustical Study of Pitch in Tibetan. Ann. Bull, RILP, 9, 137-150.
- Kjellin, Olle (1975b), How to Explain the "Tones" in Tibetan., Ann. Bull, RILP, 9, 151-166. Also presented to the 8th International Conference on Sino-Tibetan Languages and Linguistics, University of California, Berkeley, 1975.
- Kjellin, Olle (1976), Observations on Consonant Types and "Tone" in Tibetan. Unpublished thesis, University of Tokyo.
- Lea, Wayne A. (1973), Segmental and Suprasegmental Influences on Fundamental Frequency Contours. In: Hyman (ed.), Consonant Types and Tone. Southern California Occasional Papers in Linguistics, No. 1, (Los Angeles: Linguistics Program, University of Southern Calif.)
- Lehiste, Ilse, and Gordon E. Peterson (1961), Some Basic Considerations in the Analysis of Intonation. J. Acoustical Society of America, 33:4, 419-425.
- Miller, P. M. (1951), The Phonemes of Tibetan (U-Tsang Dialect) with a Practical Orthography for Tibetan-speaking Readers. J. Asiatic Society, Letters. 17:3, 191-216.
- Miller, R. A. (1955), Studies in Spoken Tibetan, I: Phonemics. J. American Oriental Society 75, 46-51.
- Pike, Eunice V. (1974), A Multiple Stress System versus a Tone System. International J. of American Linguistics, 40:3, 169-175.
- Rerix, Ju. N. (1961), Tibetskij jazyk. Izd. vostochnoj literatury. Moscow.
- Richter, Eberhardt (1964), Grundlagen der Phonetik des Lhasa-Dialektes. Schriften zur Phonetik, Sprachwissenschaft und Kommunikationsforschung, Nr. 8. Akademie-Verlag, Berlin.
- Sedláček, Kamil (1959), The Tonal System of Tibetan (Lhasa Dialect). T'oung Pao 47:3/5, 181-250.
- Simon, Walter (1970), Some Suggestions Towards a Romanisation of Modern Tibetan (Lhasa Dialect). In: Roman Jakobson and Shigeo Kawamoto (eds.), Studies in General and Oriental Linguistics, pp. 535-539, TEC, Tokyo.
- Sprigg, R. K. (1955), The Tonal System of Tibetan (Lhasa Dialect) and the Nominal Phrase. BSOAS (Bulletin of the School of Oriental and African Studies) 17:1, 133-153.
- Sprigg, R. K. (1972), A Polysystemic Approach, in Proto-Tibetan Reconstruction, to Tone and Syllable-Initial Consonant Clusters. BSOAS 35: 3, 546-587.
- Yukawa, Yasuyoshi (1971), Some Problems in Tibetan Phonology. In: Yukawa, Y., Basic Problems in Linguistics, Ch. 5, pp. 141-160. Taishuukan, Tokyo (In Japanese).