The Frequency of Tones

Author(s): Ian Maddieson

Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society (1978), pp. 360-369

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

The Frequency of Tones.

Ian Maddieson
University of California, Los Angeles.

Introduction

In order to establish reliable estimates of the frequency of occurrence of tones of certain types, a sample of over 200 tone languages has been studied.

The language sample

There are a number of well-known difficulties in the way of constructing high-quality language samples, even if only currently-spoken languages are considered. A proper random sample cannot be taken since the number of languages spoken is not known - both because some languages have escaped attention altogether and because there are no agreed criteria as to what count as separate languages as opposed to dialects of the same language. Furthermore, the quality and quantity of information available on known languages varies enormously, as does the accessibility of the data. These last factors can lead to a language sample which is biased through inclusion only of the most familiar languages, or those whose descriptions are most accessible.

The language sample in this study was selected according to a principle designed to avoid as many of these problems as possible. Each small group of closely-related languages, according to the best available genetic classifications, was considered. One tone language was selected to represent each group if one or more of the languages was tonal. Factors such as the clarity and accessibility of data only played a role in deciding which language to include from within a group. Languages with no (known) close relatives were treated as the equivalent of a group and were included. This approach ensures a language sample which includes the maximum of genetic diversity while representing language families in a manner which is roughly proportional to the number of tone languages they include. The problem of language/dialect discrimination is avoided, since cases which raise this problem obviously concern members of the same language group and only one representative can be included from each group. And the principle directs an outward search for the data to satisfy the requirements of the quota structure, thus reducing the influence of 'bibliographical convenience' (Bell 1978) on the resulting sample. Naturally, many problems concerning inadequacy and insufficiency of data still cannot be overcome, and time has placed limits on the extent to which the ideal sample has yet been approached. If all gaps could be filled, over 400 languages would have been included in the sample.

Interpretation and standardization of data

To make comparisons between languages meaningful, and derive valid generalizations, the languages must be described in a similar fashion. Otherwise, differences in descriptive paradigm are confounded with differences between languages at some common descriptive level. In this study the objective is to standardize the descriptive level at some central point along what is roughly an 'abstract-concrete' parameter ranking phonological descriptions according to the depth of the analysis in a given language and the degree of abstraction permitted by the theoretical model used. The reasons for this choice are partly pragmatic: many available descriptions do not supply enough information about higher-level processes to enable a more highly 'abstract' account to be developed; likewise, many descriptions give inadequate detail on phonetic realizations, etc, for a more 'concrete' account to be developed. However, most of the descriptions which would fall toward either extreme of an 'abstract-concrete' parameter do provide the information required to move them toward the center.

The procedures involved in standardizing the data can be thought of as involving essentially three aspects:-

- i) determining the number of basic contrastive tone units,
- ii) determining the shapes of the tones,
- iii) determining the number of contrastive pitch heights used in the tones and deriving a transcription.

The steps taken in determining the number of tones include attempting to correct over- or under-differentiation. Over-differentiation, resulting in a claim that a language has more tones than it is assigned here, can occur when predictable variant shapes of the same tone are considered distinct; when nontonal features such as contrasting syllabic structures or phonation types are held to distinguish 'tones' when the pitch patterns are in themselves similar; or when tone patterns are considered to be unitary despite evidence that a sequence of more basic tonal units is concerned.

Under-differentiation, resulting in a claim that a language has fewer tones, can occur when tones are considered the same because of a complementarity of distribution despite a lack of similarity in the tones themselves; when abstract or diacritic features are used to predict tonal differences; or when more complex tone shapes are unjustifiably regarded as composed of a sequence of simpler tones simply on reductionist principles rather than because evidence exists to show their sequential nature.

Each language was examined to determine how many tones it would have in the light of such considerations as the above. An example may help to illustrate how this was done. Standard Thai has a type of closed syllable with a short vowel followed by a stop consonant. These syllables can be either high-toned or low-toned. However, as the open or sonorant-final syllables may also have high or low tones at similar pitch levels, the case for recognizing additional distinct 'short tones' is not accepted, as this would be to over-differentiate. On the other hand, Gandour

(1975) has proposed that the (long) rising and falling tones should be regarded as sequences of low and high, and high and low tones respectively. This argument is based only on distributional criteria, rather than on showing that there are some ways that falling and rising tones behave like other sequences of high and low tones in Thai, or result from a coalescence of such units. Consequently, the elimination of rising and falling tones from the inventory is not accepted, since this would result in under-differentiation. The result is that Standard Thai is considered here as a language with 5 tones (it also has a mid tone) rather than a language with either 7 or 3 tones.

Beyond the kind of considerations outlined above, tones which are considered 'marginal' to a system have been eliminated from the basic inventory. These cases include tones which are of very rare occurrence in a language. Such marginal items are eliminated for two principal reasons. First, many available descriptions of languages reflect a level of familiarity too imperfect to be sure that such marginal items would have been reported. Hence comparability of descriptions is best served by matching such de facto elimination of marginal tones by eliminating them from the more detailed descriptions. Secondly, there is a sense in which the structure of an inventory of tones is best represented by an account which reflects the relative frequency of occurrence of tones within the language. The elimination of marginal tones represents a first crude step towards such an account.

A basic shape was determined for each tone in a system in terms of the relative pitch height of each tone and the direction and extent of any movement. The major effort was directed to correctly identifying basically level tones, requiring only one 'pitch target' for their characterization, and essentially contour tones, requiring two (or more) pitch targets for their characterization. In speech, F_0 is rarely held in a steady state. Pitch movement occurs because of co-articulation between tones in sequence, because of differences of phonation type and other segmental influences, because of stress placement and the superimposition of intonational contours, and because of mechanical factors inherent in the production of F_0 . Whatever pitch movement can be reasonably attributed to such factors is discounted in determining basic tone shapes. Thus many phonetically-occurring contours are regarded as resulting from the effects of such natural phonetic factors on tones which can be described with a single pitch target, i.e. they are level tones.

Likewise, factors such as co-articulation of tones and imposition of intonational patterns can result in an enlargement or a reduction or levelling of the pitch movement associated with an underlying contour tone. A basic shape for each contour tone when it is unaffected by the influence of these factors has been sought.

Some of the influences on tone shapes which are discounted can be summarized by such 'natural phonetic rules' as the following:

a) unmarked intonational effects tend to cause level tones to fall, especially finally (this effect seems most apparent with the lowest

level tone), and tend to compress the range of rising tones while extending the range of falling tones;

b) tones tend to begin at a pitch level which is more central in the pitch range than their (first) target level.

The way in which such rules are used to interpret tone shapes may be illustrated with reference to Standard Thai again. One set of shapes for the 5 Standard Thai tones is given by Abramson (1962). These, shown below on Chart 1, represent 'citation shapes' of the tones and show quite marked pitch movement for all five tones. They may be compared with the shapes for tones in connected speech found by Gandour and Maddieson (1976), shown on Chart 2.

An effect of the citation context can be seen in that the final falling sections of Abramson's tones 2 and 3 can be seen to be absent in the corresponding tones on Chart 2. The overall downward speech melody still makes tones 3 and 4 descend somewhat on Chart 2, but in both charts tones 2, 3 and 4 occupy separate portions of the pitch range. Their shapes can be accounted for if only one pitch target is assumed for each of them, plus the effects of rules (a) and (b) above. They are considered to be high, mid and low level tones. Note that the high and low tones both begin at points closer to the mid tone than their end-points are in Chart 2. A centralized beginning to the tones can be seen also in the two remaining tones, 1 and 5. These two tones cross the pitch ranges of the level tones and thus must be characterized by more than one target. Because the centralized beginning is general it is not assumed to be part of the basic tone shape and hence these tones can be described with only two targets each. Tone 1 is regarded as a falling tone (not a rising-falling) and tone 5 as a rising tone (not a falling-rising). The greater extent of the pitch movement for the falling tone relative to the rising tone, which is more noticeable on Chart 2, can also be understood as an adaptation to an unmarked downward intonation contour. Imposition of falling intonation is facilitated by extending the downward pitch movement of a tone and truncating the upward pitch movement of a tone. Because of these considerations, the underlying pitch range covered by these two contour tones is assumed to be essentially the same, both traversing a distance equivalent to that between the high and low level tones.

Following determination of the tone shapes, tones were coded by means of a standard transcription employing a 5-point pitch

scale. Each tone inventory was examined to determine how many distinct beginning- or end-points ('pitch heights') were required to characterize the tones in the inventory, and each pitch height was assigned to a number on the 5-point scale. Thus, for Standard Thai, 3 pitch heights were recognized, with a level tone at each height, a tone falling from the highest to the lowest of the three, and a tone rising from the lowest to the highest of the three. This transcription enables comparisons to be made concerning the proportion of the pitch-range covered by contour tones in those languages which use 3 or more pitch heights. In these cases the contour tones can be classed as using the whole of the pitch-range used for tones in the language, or as using only a part of it. It is not possible to make direct cross-language comparisons of the pitch heights themselves, as the transcription is derived independently for each language. This is necessary because the data to establish cross-language comparability of pitch phenomena is generally lacking, except in situations such as bilingualism or when borrowings occur between tone languages, see Maddieson 1977.

Results

At the present time, data from 207 languages has been examined. Some of the results are given below.

i) Number of tones in the inventory. The totals have been computed for the number of languages having phonological inventories containing 2 tones, 3 tones, 4 tones, etc. The totals are given in Table I.

Table I

<u>no of tones</u> :	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8(+)</u>
languages :	106	62	14	12	11	2	0

A striking result is that 2-tone systems are overwhelmingly the most frequent. Of the 207 languages, roughly 51% have 2-tone systems. This very striking predominance of 2-tone systems does not seem to have been noted before in the literature on general tonology. Systems with 3 tones are also common, being found in about 30% of the languages sampled. There are only small numbers of languages found with more than 3 tones, with systems of 4, 5, and 6 tones being about equally common (or perhaps more accurately, equally scarce). There are only about 6 or 7% of the languages surveyed that fall in each of these categories. It will be shown later that the step from systems of 3 tones to systems of 4 or more tones also corresponds with the point at which contour tones begin to play a significant part in the make-up of the tone inventories. A further sharp reduction occurs in the frequency of systems with more than 6 tones. Only two 7-tone systems are included in the sample, and there are none with 8 or more tones. Of course, this does not mean that such systems do not occur. The purpose of this study is not to compile a list of all the possible tone sys-

tems that do occur, but rather to derive estimates of how frequently systems with given characteristics occur. We may conclude only that probably less than 1 in 200 tone languages contrasts 8 or more tones.

ii) Frequency of tone types. The total numbers of tones of various types contained in the inventories of the languages sampled were also computed. The frequency of tone types is given in Table II.

Table II

level tones	:	465			84%	
contour tones	:	87	{	falling	42	8%
				rising	38	8%
				fall-rise	4	1%
				rise-fall	3	
total	:	<u>552</u>				

The most striking fact is that, overall, 84% of the tones in the languages sampled are level tones: the number of level tones vastly outweighs the number of contour tones. It will be shown below that this is largely a result of the structure of the common 2- and 3-tone systems.

Among the contours, falling and rising tones are effectively equally frequent in occurrence. The complex bidirectional tones occur very rarely, but both types (falling-rising and rising-falling) occur about as often as each other. These results conflict with earlier claims that falling tones outnumber rising tones in the languages of the world (cf Cheng 1973, Ohala 1973, Hombert 1978). The likelihood is that this claim reflects patterns observed in surface pitch contours, rather than reflecting inherent shapes of tones. It will be seen later that an implicational universal stating that an inventory must have a falling tone before it can have a rising tone would also be false. It seems that any productive or perceptual mechanism that may favor falling pitch in speech does not lead to more falling than rising underlying shapes for tones.

iii) Frequency of tone systems. The frequency of tone systems of various structures was also examined. A summary of the distribution by languages of contour tone types is given in Table III.

Table III

languages with only level tones	:	163	(79%)
languages with level(s) & simple contour(s)	:	37	(18%)
languages including complex contours	:	7	(3%)

Languages whose tone systems contain only level tones comprise the vast majority of tone languages. There are only 37 languages in the sample with systems consisting of a level tone or tones and a simple contour tone or tones, and only 7 including complex contours as well.

The makeup of each individual system in terms of level (L), falling (F), rising (R), falling-rising (FR) and rising-falling

(RF) tones was also analyzed and the results are given in Table IV. This shows the number of languages in the sample with systems containing each different combination of tones found. Systems with 2 tones, 3 tones, 4 tones, etc are listed separately. So, for example, under 3-tone systems an entry reading 2L + F : 7 means that there were 7 languages whose systems consisted of 2 level tones and one falling tone.

Table IV

<u>2-tone systems</u>	2L : 106		
<u>3-tone systems</u>	3L : 49	2L + F : 7	2L + R : 3
	L + F + R : 2		
<u>4-tone systems</u>	4L : 7	3L + F : 2	3L + R : 2
	2L + F + R : 1	L + F + R + FR : 2	
<u>5-tone systems</u>	5L : 1	3L + 2F : 1	3L + 2R : 1
	3L + F + R : 7	2L + F + R + FR : 1	
	L + F + 2R + FR : 1		
<u>6-tone systems</u>	4L + 2R : 1	3L + F + 2R : 3	4L + F + R : 1
	3L + 2F + R : 1	2L + 2F + 2R : 1	2L + F + 3R : 1
	2L + 2F + R + RF : 1	2L + F + 2R + RF : 1	
<u>7-tone systems</u>	3L + 2F + 2R : 1	3L + F + 2R + RF : 1	

In all of the languages with inventories containing only 2 tones, both of the tones are level. Although other types of 2-tone systems may occur, they must be regarded as decidedly rare - probably occurring in less than 1 in 200 tone languages. 3-tone systems are also predominantly found to contain only level tones. This is the case for 80% of the 3-tone languages, the remaining 20% having systems including 1 falling or rising tone, or, in a few cases, both. Although the most frequent single type of 4-tone system has 4 level tones, there are as many 4-tone systems which include a contour or contours as those which consist of only level tones. A language with 5 tones which are all level is a rarity. There is only one in the sample. The most common type of 5-tone system is one containing 3 levels, one fall and one rise. This is represented by 7 of the 12 languages concerned, i.e. by about 58% of them. Clearly, when the system contains as many as 5 tones contours are normally found to be included. All the 6- and 7-tone systems contain contours; no systems with 6 or 7 level tones are included in the sample and, in fact, no such systems have been reported in the literature.

Thus, although Wang's (1967) claim that systems with 4 tones must include at least one contour tone is clearly false, contour tones are certainly not unexpected in systems with 4 tones, they are highly probable in systems with 5 tones, and are required in

systems with 6 or more tones. As noted earlier, contour tones are rare in 3-tone systems, but become quite common in 4-tone systems. This dividing line which falls between 3 and 4 tone systems is also marked by the restriction of complex contours to those languages which have 4 or more tones.

An odd result is that falling-rising tones only occur in 4- and 5-tone systems, and rising-falling tones only occur with 6- and 7-tone systems in the languages sampled. In view of the small numbers involved, this is likely to be only a sampling error with no significance.

iv) The extent of contours. When 3 or more pitch heights are contrasted in a language, the portion of the pitch range which contour tones cover can be compared. For example, in a system which includes 3 level tones and 1 falling tone, the fall may extend over the whole pitch range from high to low, or may extend only over a part of the range, i.e. from high to mid or from mid to low. The data in Table V show how many of the simple contours in languages with 3 or more contrasting pitch heights cover the whole range or a part of the pitch range. Data are given separately for languages whose inventories include only 1 falling tone (F) or 2 falling tones (2F), and only 1 rising tone (R) or 2 rising tones (2R), as well as totals for all contour tones, all falling tones and all rising tones.

Table V

	falling		rising		totals		all
	F	2F	R	2R	falls	rises	
whole pitch range	<u>10</u>	5	4	<u>5</u>	15	9	24
partial pitch range	<u>3</u>	7	5	<u>11</u>	10	16	26

An interesting pattern emerges in Table V. The extent of contours is dependent on the direction of movement of the contour and the number of tones in the system that move in a given direction. Two asymmetries stand out in the table (the crucial numbers are underlined). Although full-range and partial-range contours are about equally frequent overall, when there is only one falling tone in the system, it most often extends over the whole tone range. On the other hand, when there are 2 rising tones in the system, the most commonly found type covers only a part of the tone range. It is clearly not the case that simple contours usually cover only a part of the tone range, as suggested by Wang (1967).

The observed asymmetries require explanation. As a first hypothesis, it may be suggested that there is systemic pressure to distinguish 2 tones which move in the same direction by separating them into 2 distinct registers. Both 2F and 2R systems contain more partial-range contours. But there may also be phonetic factors, such as the tendency of falling tones to be extended and rising tones to be truncated, which result in a greater effect of this pressure on the underlying shapes of falling than rising tones.

Where distinguishing between 2 tones which move in the same direction is not a factor, a falling tone is generally found to cover the whole pitch range.

Tone universals

This study also provides an opportunity to test the reliability of generalizations concerning tones, including proposed 'tone universals'. In this connection, it has already been noted that falling and rising tones are equally common and that systems with 4 level tones and no contours occur. Wang's claim that no 4-tone systems occur without at least 1 contour could have been tested by a different strategy involving a direct search for counterexamples; but generalizations such as the claim that falling tones outnumber rising tones can only be tested by constructing a language sample in which the frequencies of tone types are represented in their true ratios. In research on language universals, the direct construction of hypotheses and search for counterexamples is complemented by studies such as the present one using sampling techniques.

For example, an implicational hierarchy has been proposed concerning tone types (Maddieson 1978). This is summarized by the two statements below:-

- a) If a language has contour tones it also has level tones.
 - b) If a language has complex contours it also has simple contours.
- These say that no language would have a contour tone without at least 1 level tone in its inventory, and no language would have a complex tone without at least 1 simple contour in its inventory. None of the languages in the sample violates this hierarchy. Because of the way the sample is constructed it can be said that at the very least the normal pattern of distribution of tone types is captured by this hierarchy.

It is important to note that this hierarchy imposes more highly-structured constraints on the distribution of simple and complex contours in tone inventories than are imposed by the simple fact of their rarity. For example, if contour tones were randomly distributed in languages in the same overall ratio to level tones represented in the sample, several 2 or 3-tone systems containing only contours would have been expected to occur.

In fact, the distributional limits on tone types seem even more highly structured than this hierarchy indicates. This can be seen from the data in Table VI, which compares the overall ratio of 84% level tones to 16% contours in the whole sample with the separate percentages of level and contour tones in systems containing different numbers of tones.

Table VI

	<u>number of tones in the system</u>						<u>overall</u>
	2	3	4	5	6	7	
% level	100	93	78	58	47	43	84
% contour	0	7	22	42	53	57	16

The overall ratio is most closely matched by the ratio in 4-tone systems. However, this overall ratio is not at all typical of the systems with other than 4 tones. Instead the level to contour ratio is highly dependent on the number of tones in the system; the percentage of levels decreases and the percentage of contours increases as the number of tones increases. In systems with 6 or more tones a majority of the tones involved are contours.

Concluding thoughts

This paper has reported on the frequency of tones and tone systems. It has served to highlight such facts as that the most 'normal' tone system is one with only two level tones. It has provided new, reliable data concerning the relative frequency of level tones and contours of different types. It has provided an opportunity to test certain ideas concerning the constraints on the composition of tone inventories, and revealed that there is more to be done before these are understood better. It provides a source of data from which to derive expectations of the likely historical changes as tone systems evolve.

Note: This work was supported by a grant from the NSF.

References

- Abramson, A. 1962. "The vowels and tones of Standard Thai: acoustical measurements and experiments" *International Journal of American Linguistics*, Publication 20.
- Bell, A. 1978. "Language samples" in *Universals of Human Language*, edd. J.H. Greenberg et al., Stanford University Press.
- Cheng, C-C. 1973. "A quantitative study of Chinese tones" *Journal of Chinese Linguistics* 1: 93-110.
- Gandour, J. & I. Maddieson. 1976. "Measuring larynx movement in Standard Thai using the cricothyrometer" *Phonetica* 33: 241-267.
- Hombert, J-M. 1978. "A model of tone systems" in *Tone and Accent*, ed. D.J. Napoli, Georgetown University Press, Washington D.C.
- Maddieson, I. 1977. "Tone loans: a question concerning tone spacing and a method of answering it" *Studies on Tone (UCLA Working Papers on Phonetics 36)*: 49-83.
- Maddieson, I. 1978. "Universals of tone" in *Universals of Human Language*, edd. J.H. Greenberg et al., Stanford University Press.
- Ohala, J.J. 1973. "The physiology of tone" in *Consonant Types and Tone*, ed. L.M. Hyman, University of Southern California.
- Wang, W. S-Y. 1967. "The phonological features of tone" *International Journal of American Linguistics* 33: 93-105.