The Functional Load of Tone in Mandarin is as High as that of Vowels

Dinoj Surendran & Gina-Anne Levow

Computer Science Department,
University of Chicago, Chicago, IL, USA
{dinoj; levow}@cs.uchicago.edu

Abstract

Tonal languages, such as Mandarin, convey information using both phonemes and tones. Using a recently proposed framework for measuring the functional load of a phonological contrast (i.e. how much use a language makes of the contrast), we carry out several computations to estimate how much use Mandarin makes of tones. The most interesting result is that identifying the tone of a syllable is at least as important as identifying the vowels in the syllable. Another computation suggests that the contrast between low and neutral tone carries relatively little information.

1. Introduction

While frequency counts represent a measure of how much use a language makes of a linguistic unit, such as a phoneme or a word, it is often more important to consider not the units themselves but the contrasts between them. **Functional load**, a concept rooted in the fields of linguistics and speech recognition [8, 3], refers to a measure of how much use a language makes of a contrast.

The importance of a phonological contrast has potential implications for language learning, future language change, and for speech recognition and phonological modeling. One could argue that it is more important to effectively teach or accurately model a contrast that carries more information. It has been conjectured that a contrast with lower functional load might, in fact, be neutralized completely as a language changes over time [10]. However, this has by no means been confirmed [9].

In the current paper, we focus on assessing the importance of lexical tone in Mandarin Chinese. In Section 2, we briefly introduce the framework recently offered by Surendran and Niyogi [12] of a family of measures for computing the functional load (FL) of several phonological contrasts. These contrasts include not only phonemic oppositions, but also tonal oppositions, as well as several others, and thus provide a suitable framework for our current comparisons.

Previous quantitative definitions of functional load rarely dealt with suprasegmentals such as tone. Those that did, such as [6], had other flaws, such as not taking into account word frequency and/or structure. Therefore, references in the literature to functional load tend to be qualititative, e.g. "in lexical tone languages, tones have a high functional load" [2], "Geman Deng [is a Chinese dialect/language with] 4 tones which are reported to have a low functional load" [7], and "The pitch class languages in Europe such as Norwegian, Swedish, Bosnian, Croatian, Serbian, Latvian, and Lithuanian, are in no way comparable to true tonal languages such as Chinese. In the latter every syllable has an unpredictable tone whereas in pitch class languages most syllables have a predictable tonal contour and the tones are of low functional yield." This imprecision is not limited to suprasegmentals, as witnessed by statements such as "The functional load of voicing in Japanese may be lower than normally thought" [11]. The framework of [12], as exemplified in this paper, permits such statements to be quantified with the appropriate corpora, and more detailed questions to be asked.

We consider the functional load carried by the presence of contrastive lexical tones in general and the information loss that would be incurred by neutralization of the tonal contrast, in Section 4, and compare it to a clearly essential set of contrasts between the vowels of Mandarin and the functional load of stress in several other languages. We find support for the phonological importance of Mandarin lexical tone contrasts. We further evaluate the relative importance of the pairwise contrasts between the tones (Section 4). Section 7 presents conclusions and suggests future work in this area.

2. Information-Theoretical Measure of Functional Load

The analysis presented below employs the family of measures developed by Surendran and Niyogi [12] for computing the functional load (FL) of several phonological contrasts. We describe their information-theoretic framework in as much detail as is required for this paper. The definition is similar to Carter's Percentage of Information Extracted measure [3] and to Hockett's 1955 definition of the FL of binary phonemic oppositions [8]; the definition of FL in [12] includes both as special cases.

First, we assume that a language L is a sequence of discrete units, such as phonemes, morphemes, syllables or words. Second, we assume that L is generated by a stationary ergodic process, so that one can speak of its entropy $H(L)$, where entropy is typically defined as,

$$H(L) = -\sum_{x} p(x) \log_2 p(x) \qquad (1)$$

The units of L can be quite complicated. We consider words to be sequences of syllables, and syllables to be sequences of phonemes. To capture tone or stress contrasts, we can augment the syllable representation with a tone or stress value.

Supposed we refer to a language L, considered as a sequence of units of type $\mathtt{U}$, as $L_{\mathtt{U}}$. To find the FL of obstruents and sonorants, we create a language $L_{\mathtt{U}}^{-son.obs}$ where one can tell sonorants from obstruents, but not distinguish between sonorants or between obstruents. In other words, we convert $L_{\mathtt{U}}$ to $L_{\mathtt{U}}^{-son.obs}$ by a function that takes a syllable in $L_{\mathtt{U}}$ with phonemic component $p_1 \ldots p_n$ to a syllable in $L_{\mathtt{U}}^{-son.obs}$ with phonemic component $p'_1 \dots p'_n$, where

$$p'_i = \begin{cases} S & \text{if} \quad p_i \text{ is a sonorant} \\ O & \text{if} \quad p_i \text{ is an obstruent} \end{cases} \tag{2}$$

The functional load in the language with different sonorant and obstruent contrasts removed is defined to be

$$FL_{\mathtt{U}}(son.obs) = \frac{H(L_{\mathtt{U}}) - H(L_{\mathtt{U}}^{-son.obs})}{H(L_{\mathtt{U}})} \tag{3}$$

FL values should be interpreted as relative, rather than absolute, values. In other words, FL values can only be interpreted by comparing them to other FL values. This is because there is no definitive method to find the entropy of a language, that takes into account, for example, semantic and syntactic structure. Once we accept that any values we get are relative, it turns out that we can use even very crude methods of entropy estimation.

We estimate the entropy of a language by the entropy of the stationary distribution that would result if $M_{\mathtt{U}}$ was generated by a 0-order Markov model. This is not oversimplistic, since one of the key points made in [12] is that the order of the Markov model, and the corpus used, does not matter much since FL values are relative. Having already said that FL values should be interpreted by comparing them to other FL values, we add that they should be compared to other FL values computed the same way they were. For example, values obtained with a unigram syllable model should not be compared with values obtained with a unigram word model.

Of course, there will still be some error in the final values obtained. Fortunately, our results are clear enough to survive such errors.

3. Functional Load Definitions for Mandarin Contrasts

For Mandarin we consider words to be a sequence of syllables, and syllables to be a sequence of phonemes plus a single tone. The tone takes one of five possible values – high level, rising, low, falling and neutral. Suppose we refer to regular Mandarin, considered as a sequence of units of type `U`, as $M_{\mathtt{U}}$, and Mandarin without tones as $M_{\mathtt{U}}^{-tone}$. `U` is either a syllable (`syll`) or word (`word`). The canonical mapping from $M_{\mathtt{U}}$ to $M_{\mathtt{U}}^{-tone}$ takes a syllable in $M_{\mathtt{U}}$ with phonemic component $p_1 \dots p_n$ and tone q to a syllable in $M_{\mathtt{U}}^{-tone}$ with phonemic component $p_1 \dots p_n$.

The functional load of tone in Mandarin is then defined to be

$$FL_{\mathtt{U}}(tone) = \frac{H(M_{\mathtt{U}}) - H(M_{\mathtt{U}}^{-tone})}{H(M_{\mathtt{U}})} \tag{4}$$

For another example, suppose we wanted to find the FL of vowels in Mandarin. This time we convert $M_{\mathtt{U}}$ to a language $M_{\mathtt{U}}^{-vowels}$ where all vowels sound the same. The conversion function takes a syllable in $M_{\mathtt{U}}$ with phonemic component $p_1 \dots p_n$ and tone q to a syllable in $M_{\mathtt{U}}^{-vowels}$ with phonemic component $p'_1 \dots p'_n$ and tone q, where

$$p'_i = \begin{cases} V & \text{if} \quad p_i \text{ is a vowel} \\ p_i & \text{if} \quad \text{otherwise} \end{cases} \tag{5}$$

The functional load of vowels in Mandarin is then defined to be

$$FL_{\mathtt{U}}(vowels) = \frac{H(M_{\mathtt{U}}) - H(M_{\mathtt{U}}^{-vowels})}{H(M_{\mathtt{U}})} \tag{6}$$

Table 1: *The FL of several groups of phonemes and distinctive features in Mandarin, using syllable and word unigram models of language.*

x	$FL_{\mathtt{syll}}(x)$	$FL_{\mathtt{word}}(x)$
Consonants	0.235	0.081
Tones	0.108	0.021
Vowels	0.091	0.022
Stops	0.029	0.006
Fricatives	0.021	0.005
Place	0.065	0.014
Manner	0.034	0.006
Aspiration	0.002	0.0003

Note that the phonological contrast here is not that of being able to tell vowels from non-vowels, but of being able to distinguish between vowels.

4. Computing the Importance of Tone

For every unit x in, say, $M_{\mathtt{U}}$, we found the count $c(x)$ of how many times x appeared in a large corpus. The corpus we used was the Topic Detection and Tracking (TDT) 3 Multilanguage Text (Version 2.0) with [1] transcriptions of Voice of America Mandarin broadcasts, which has about 1.6 million syllables. We converted the character representations to phonemic representations using the on-line resources provided by the Linguistic Data Consortium. Normalizing these counts by $p(x) = \frac{c(x)}{\sum_y c(y)}$, we took $H(M_{\mathtt{U}})$ to be

$$H_0(M_{\mathtt{U}}) = -\sum_x p(x) \log_2 p(x) \tag{7}$$

$FL_{\mathtt{syll}}(tone)$ and $FL_{\mathtt{word}}(tone)$ in Mandarin are 10.8% and 2.1% respectively. Relativeness means that these figures by themselves are meaningless. One cannot say, for example, that tones carry 2% of the information in Mandarin, because speakers know far more than just the vocabulary of the language. To make them meaningful, we compare them to the FL values of something else that we know is important, such as vowels. The corresponding figures for the FL of vowels are 9.1% and 2.2% respectively, so tones are clearly comparable in their importance to vowels. On the other hand, neither are as important as consonants; the FL of consonants are 23.5% and 8.1% respectively. Other results are summarized in Table 1.

One immediate observation from Table 1 is that $FL_{\mathtt{syll}}$ is consistently smaller than $FL_{\mathtt{word}}$. This is because words, being composed of syllables, contain more information than syllables.

5. Comparison across Languages and Phonological Units

Further insight into this result comes from comparing Mandarin with non-tonal languages such as Dutch, English and German. Calculations for these languages were done using the CELEX database [1] and are summarized in Table 3. Syllables for these languages have a stress component that is the closest thing comparable to tone. Syllables in tonal languages can also have a stress component in addition to a tone component; it so happens that the corpus we used here only codes the latter.

[1] Available from http://www.ldc.upenn.edu/

Table 2: *Summary information for corpora used here.* $N_{\mathtt{U}}$ *is the number of* U*-units in the corpus (in millions), while* $H_{\mathtt{U}}$ *is the entropy using a unigram model.*

	$N_{\mathtt{syll}}$	$N_{\mathtt{word}}$	$H_{\mathtt{syll}}$	$H_{\mathtt{word}}$
Mandarin	1.6	0.9	8.3	10.4
Dutch	16.2	9.8	8.2	9.5
English	24.0	16.8	9.2	9.5
German	9.1	5.0	9.3	10.5

Syllables in English can have no, primary or secondary stress; syllables in Dutch and German can be stressed or unstressed. These three languages are closely related and have values that are remarkably similar. They will be referred to as DEG for short.

DEG have relatively high $\mathrm{FL}_{\mathtt{syll}}$(stress) but negligible $\mathrm{FL}_{\mathtt{word}}$(stress), about 3 and 60 times less than the corresponding values for Mandarin. This showing that word, far more than syllable, structure makes stress relatively redundant in DEG.

DEG and Mandarin have similar FL values for Consonants + Vowels + Tone/Stress. The only information available to listeners of the corresponding reduced languages here are CV strings of the syllables/words. But the FL of Consonants and Vowels separately are systematically larger for DEG than Mandarin, with the difference being made up by the amount of FL for Tone/Stress. Roughly speaking, the total fraction of information is split up by DEG amongst consonants and vowels only, while the same information is split up by Mandarin between Consonants, Vowels and Tones.

6. Comparisons among Mandarin Tones

Tables 4 and 5 have FL values for all binary tonal oppositions in Mandarin. To find the FL of tones t_1 and t_2, we convert $M_{\mathtt{U}}$ to a language $M_{\mathtt{U}}^{-t_1t_2}$. Suppose two syllables in $M_{\mathtt{U}}$ have the same phonemic structure but one has tone t_1 and the other tone ft_2. Those two syllables would sound the same in $M_{\mathtt{U}}^{-t_1t_2}$. The conversion function takes a syllable in $M_{\mathtt{U}}$ with phonemic component $p_1 \dots p_n$ and tone q to a syllable in $M_{\mathtt{U}}^{-t_1t_2}$ with phonemic component $p_1 \dots p_n$ and tone q', where

$$q' = \begin{cases} t_{12} & \text{if} \quad q \text{ is } t_1 \text{ or } t_2 \\ q & \text{if} \quad \text{otherwise} \end{cases} \tag{8}$$

t_{12} is a placeholder representing 'tone t_1 or t_2'.

$FL(t_1, t_2)$ is consistently high when neither t_1 or t_2 is the neutral tone. Bearing relativeness in mind, we use the FL of aspiration as a benchmark. Roughly speaking, 'high' here means of at least the same order of magnitude as the FL of aspiration. The neutral tone does (lexically) contrast highly with the two contour tones, but not with the first or third tone.

Note that $FL(t_1, t_2) = 0$ if and only if t_1 and t_2 are in complementary distribution.

We briefly consider the small low FL(neutral,low) and FL(neutral,high) values in more detail. The reason for their size is *not* simply the distribution of tones among syllable tokens. If the distribution of tones in words followed this distribution, then the resulting FL values would be approximately proportional to the values in Table 6. That would require the values of FL(neutral,low) and FL(neutral,high) to be at least 10 times higher than they actually are.

Table 3: *Detailed comparison of the FL of tones in Mandarin with that of stress in Dutch, English and German. The top half of the above table appears in [12].*

	Syllables	Words
Tone		
Mandarin	0.108	0.0213
Stress		
Dutch	0.026	0.0007
English	0.027	0.0001
German	0.034	0.0002
Vowels		
Dutch	0.126	0.052
English	0.133	0.049
German	0.161	0.042
Mandarin	0.091	0.022
Consonants		
Dutch	0.334	0.193
English	0.310	0.176
German	0.336	0.154
Mandarin	0.235	0.081
Vowels + T/S		
Dutch	0.192	0.053
English	0.204	0.068
German	0.234	0.044
Mandarin	0.260	0.082
C + V		
Dutch	0.634	0.512
English	0.562	0.401
German	0.641	0.448
Mandarin	0.453	0.231
C + V + T/S		
Dutch	0.729	0.529
English	0.664	0.463
German	0.742	0.480
Mandarin	0.700	0.485
Obstruents		
Dutch	0.154	0.049
English	0.155	0.070
German	0.188	0.059
Mandarin	0.167	0.048
Sonorants		
Dutch	0.266	0.140
English	0.256	0.119
German	0.218	0.072
Mandarin	0.195	0.057
S + O		
Dutch	0.528	0.361
English	0.518	0.336
German	0.514	0.288
Mandarin	0.488	0.262
S + O + T/S		
Dutch	0.622	0.373
English	0.629	0.401
German	0.613	0.309
Mandarin	0.737	0.530

Table 4: *The FL of all tonal binary oppositions in Mandarin, based on syllable unigrams. All values should be multiplied by 0.01. For comparison, $FL_{\texttt{syll}}$ of aspiration is 0.2 × 0.01.*

	High	Rising	Low	Falling
Rising	1.6			
Low	2.1	1.5		
Falling	2.9	2.2	2.4	
Neutral	0.02	0.3	0.04	0.06

Table 5: *The FL of all tonal binary oppositions in Mandarin, based on word unigrams. All values should be multiplied by 0.01. For comparison, $FL_{\texttt{word}}$ of aspiration is 0.03 × 0.01.*

	High	Rising	Low	Falling
Rising	0.2			
Low	0.4	0.2		
Falling	0.5	0.4	0.4	
Neutral	0.002	0.1	0.001	0.02

The real reason is that for most syllables with neutral tones, in particular the most common ones ('de', 'le', 'men'), there is no third tone variant, at least not in the large dictionary we checked. And while there are first tone variants, there are not enough of them in our corpus to result in a high functional load value.

7. Conclusion

A quantitative information-theoretic measure demonstrates the important role played by tone in Mandarin Chinese. In particular, lexical tone contrast has been shown to have a comparable functional load to that of vowels for Mandarin, and higher FL than stress in English, Dutch, and German. We have also demonstrated that the importance of tonal contrasts varies among the tones of Mandarin.

These findings have possible implications for speech recognition and language development. First, automatic speech recognition systems for tonal languages typically do not make direct use of tonal information. One reason for this is that the underlying tonal sequence is very hard to work out, especially with coarticulation effects [13]. This result suggests that further work on this hard problem will be well worth it. Second, some language reformers have suggested that tones do not need to be represented in a revised alphabet. Our result suggests that such an alphabet would be as hard to use as an alphabet that represented tones but not vowels. And while it is certainly possible to read written texts e.g. Hebrew, that don't have vowels, it is not easy (especially if the text does not separate words).

In future work, we plan to extend these analyses to other tone languages, both to other East Asian languages and to African tone languages.

For more results and papers on functional load, please go to http://dinoj.info/research/fload/index.html .

8. References

[1] Baayen, R. H.; Piepenbrock, R.; and Gulikers, L., 1995. *The Celex Lexical Database* (Release 2) [CD-ROM].

Table 6: *The FL of all tonal binary oppositions in Mandarin, based on the distribution of syllable tones in Table 7.*

	High	Rising	Low	Falling
Rising	0.22			
Low	0.19	0.17		
Falling	0.25	0.23	0.19	
Neutral	0.11	0.11	0.09	0.12

Table 7: *The fraction of syllables with different tones in Mandarin, based on the TDT3 corpus of VOA Mandarin broadcasts.*

Tone	High	Rising	Low	Falling	Neutral
Fraction	0.27	0.22	0.16	0.28	0.07

Philadelphia, PA. Linguistic Data Consortium, University of Pennsylvania (Distributor).

[2] Barry, J.; Blamey, P.; Lee, K.; Cheung, D., 2000. Differentiation in Tone Production in Cantonese-speaking Hearing-impaired Children. *Proceedings of International Conference on Spoken Language Processing.* Beijing, 1:669-672.

[3] Carter, D. M., 1987. An Information-theoretical Analysis of Phonetic Dictionary Access, *Computer Speech and Language,* 2:1-11.

[4] Ching T., 1984. Lexical Tone Pattern Learning in Cantonese Children. *Language, Learning and Communication,* 3:243-414.

[5] Frankel, S., 1999. *Phonology, Verse Metrics, and Music.* Ph.D. dissertation, Department of Music, New York University.

[6] Greenberg, H. H., 1959. A Method of Measuring Functional Yield as applied to Tone in African Languages, *Georgetown University Monograph Series on Language and Linguistics*, 12:7-16.

[7] Grimes, B. F. (ed.), 1996. *Ethnologue*, 13th Edition, SIL International.

[8] Hockett, C. F., 1955. *A Manual of Phonology*, International Journal of American Linguistics, Vol. 21-4, Indiana University Publications.

[9] King, R. D., 1967. Functional Load and Sound Change. *Language*, 43:831-852.

[10] Martinet, A., 1964. *Elements of General Linguistics*, Faber and Faber.

[11] Ohno, K. (to appear) Seidaku: More than a Voicing Difference - for a Better Understanding of the Rendaku Phenomenon. *Voicing in Japanese*, van de Weijer, J.; Nanjo, K.; Nishihara T. (eds.), Berlin: Mouton de Gruyter.

[12] Surendran, D.; Niyogi, P., 2003. Measuring the Usefulness (Functional Load) of Phonological Contrasts, submitted. Preprint available at http://arxiv.org/abs/cs.CL/0311036

[13] Xu, Y, 2003. Contextual Tonal Variations in Mandarin. *Journal of Phonetics*, 25:61-83.