Tone circles and chance*

James Myers
National Chung Cheng University
Lngmyers@ccu.edu.tw

DRAFT
June 21, 2006

As first noticed by Bodman (1955), Southern Min tone sandhi gives rise to the system of alternations among the lexical long tones shown in (1) (using the phonetic tone values found in Chiayi county in southern Taiwan, expressed in the 5-point IPA scale). Given its shape, this system is often called the Southern Min Tone Circle (though the Southern Min Tone Lollipop would be a more accurate name, given the "stick" where two tones neutralize into one).

$$
\begin{array}{llllll}
(1) & [55] & \rightarrow & [33] & \leftarrow & [24] \\
 & \uparrow & & \downarrow & & \\
 & [51] & \leftarrow & [21] & &
\end{array}
$$

A great deal of effort has been expended on accounting for this system in an elegant grammar. Wang (1967) used a set of interlocking Greek-variable rules; Yip (1980) used rules where the input [33] and output [33] had different phonological representations; Tsay (1994) proposed an analysis that reversed the traditionally assumed direction. Recently, Barrie (2006) has presented an Optimality-Theoretic analysis in which the Tone Circle emerges via constraints that preserve contrast in the system as a whole.

However, Tsay and Myers (1996) and Moreton (2004) have argued that no phonological analysis is necessary at all if the pattern in (1) involves morphological allotone selection: speakers simply memorize arbitrary tone pairs ([55]~[33], [33]~[21], etc). For Moreton (2004) this conclusion is particularly crucial, since he proves that Optimality Theory is mathematically incapable of handling circles like (1) if they are real.

To the best of my knowledge, however, nobody has ever completed the argument by showing that the pattern in (1) is indeed likely to be a coincidence, as the allotone hypothesis predicts.[1] Of course, without any premises at all, the pattern in (1) is extremely unlikely to arise by chance alone (e.g. why isn't it [13] ↔ [42]?). However, the assumptions needed to account for the Tone Circle are not very complex; the two key assumptions are listed in (2).

(2) a. Southern Min has the five lexical long tones [55], [33], [24], [51], [21].
    b. Tone sandhi consists of pairs of these lexical tones (structure-preserving).

When we now come to consider what we should count as a "successful" outcome of the analysis, it is important to note that it is *not* (1) itself. What makes (1) notable is the circular shape, not the specific value and order of the tones. Thus any five-tone system with a loop

---

* An earlier version of this appeared as part of Myers (2006).

[1] Note that the issue here is whether the *pattern* is a coincidence, not its instantiation across the lexicon nor its use by speakers. Clearly the fact that virtually all Southern Min morphemes conform to this pattern is no coincidence. Thus the proper concept is "allotone selection" and not "allomorph selection."

anywhere in it would attract our attention. Indeed, in the universe of loops formed of these five tones, the pattern in (1) is not the most noteworthy imaginable: more amazing would be a system consisting solely of a five-tone loop, without any "lollipop stick."

Suppose we take the position that in order to count as theoretically interesting, the loop must be at least as big as the one that is actually observed: either four tones, as in the actual tone circle in (1), or five. The probability of getting such a "big" loop by chance is the number of such looped systems divided by the total number of tone systems defined as in (2). To calculate this, we start by encoding a "big" loop in a 5-tone system as an ordered series (t_1, t_2, t_3, t_4, t_5, t_6) where either $t_6 = t_1$ (5-tone loop) or $t_6 = t_2$ (4-tone loop). In the first case, there are 5! (= 1×2×3×4×5) possible ordered series (t_1, t_2, t_3, t_4, t_5), but since it doesn't matter which tone we actually start the loop with (e.g. (t_1, t_2, t_3, t_4, t_5, t_1) = (t_2, t_3, t_4, t_5, t_1, t_2)), the number of 5-tone-loop systems is actually 5!/5 = 4! = 24. The logic works the same way in the second case, except now we must treat the "lollipop stick" as special (e.g. $(t_1, t_2, t_3, t_4, t_5, t_2) \neq (t_2, t_3, t_4, t_5, t_1, t_3)$), so the number of 4-tone-loop systems is 5! = 120. Put together, then, the total number of "big" loops is 144 (= 24 + 120).

As for the number of logically possible tone systems, we start with the observation that for each of the five tones, there are five logically possible outputs (including vacuous rules where input and output are identical). A tone system will then be a set of five tone rules, each starting with a different input and each ending with one of five outputs. If each rule is independent of the others (permitting neutralization, since the real system permits it), the total number of possible tone systems, assuming an ordered set of lexical tones, is (5 possible rules for t_1) × ... × (5 possible rules for t_5) = 5^5 = 3125. This means that the probability of getting a "big" loop by chance in a system conforming to (2) is p = 144/3125 = 0.04508, just marginally significant by the usual conventions ($p < .05$).[2]

Even this rather high p value is artificially low, however. First, consider the effect of adding the third premise shown in (3). This has the flavor of an anti-faithfulness constraint, but it is also motivated on functional grounds, given the potential usefulness of tone sandhi to listeners as a marker of syntactic constituent boundaries (see Tsay, Myers, and Chen 2000). With this addition, the number of logically possible tone systems drops to 4^5 = 1024, so p = 0.140625 (= 144/1024), which is not statistically significant ($p > .05$).

(3) Members of a tone sandhi pair must be distinct (all tones alternate).

Moreover, if we follow Moreton (2004) and assume that a loop of any size is problematic for OT, and thus theoretically interesting, we must actually consider the set of *all* looped systems, including systems containing multiple small loops. There is no need to calculate it exactly, since it must be a proper superset of the set of 5-tone-loop systems, increasing the size of the numerator and thus the p value.

Thus the Southern Min Tone Circle is at best only a marginally significant pattern, and at worst, not significant at all. No phonological analysis is necessary, though the premises in (2ab) and (3) may require an explanation.

[2] Actually, by a convention adopted by many statisticians, p values for "exact" (probability-based) tests like this are often multiplied by 2 to make them "two-tailed," since the alternative hypothesis would be satisfied if the observed outcome was either lower *or* higher than chance. Here this would give us p = 0.09216, which is not significant.

References

Barrie, Michael (2006) Tone circles and contrast preservation. *Linguistic Inquiry* 37 (1):131-141.

Bodman, N. C. (1955) *Spoken Amoy Hokkien*. Revised and published 1987, Ithaca, NY: Spoken Language Services.

Moreton, E. (2004). Non-computable functions in Optimality Theory. In J. J. McCarthy (Ed.), *Optimality Theory in phonology: A reader* (pp. 141-163). Oxford, UK: Blackwell.

Myers, J. (2006). Linguistics as cognitive psychology. In *Pre-Conference proceedings of the 14th Annual Conference of the International Association of Chinese Linguistics & 10th International Symposium on Chinese Languages and Linguistics Joint Meeting* (pp. 150-174). Taipei, Taiwan, May.

Tsay, J. (1994). *Phonological pitch*. Doctoral dissertation, University of Arizona, Tucson.

Tsay, J. & Myers, J. (1996). Taiwanese tone sandhi as allomorph selection. *Proceedings of the Berkeley Linguistic Society, 22*, 394-405.

Tsay, J., Myers, J., & Chen X-J. (2000). Tone sandhi as evidence for segmentation in Taiwanese. *Child Language Research Forum, 30*, 211-218.

Wang, W. S-Y. (1967). Phonological features of tone. *International Journal of American Linguistics, 33*, 93-105.

Yip, M. (1980). *The tonal phonology of Chinese*. Doctoral dissertation, MIT.