

## Phonological Knowledge beyond the Lexicon in Taiwanese Double Reduplication

Key words: tone sandhi, productivity, stochastic modeling

Recent research on phonological productivity has shown that speakers often have knowledge that cannot be gleaned from the lexicon. They know more than the lexicon: for instance, Davidson (2005) and Zuraw (2007) demonstrated that speakers' production and grammatical judgments of novel words do not correlate with lexical frequencies, but must be informed by either formal or phonetically-based a priori principles. They also know less than the lexicon, particularly in cases of phonological opacity: Sanders (2001) showed that Polish speakers do not extend the counterbleeding interaction between o-Raising and Final Devoicing in the lexicon to novel words; the opaque Taiwanese "tone circle," which applies to non-XP-final syllables as in (1), has also been repeatedly shown to be largely unproductive in wug tests (Berko 1958) despite its exceptionlessness in the language (Hsieh 1970, Wang 1993, Zhang et al. to appear). Based on the results of a wug test, we argue in this paper that the tone pattern in Taiwanese double reduplication is a case in which the speakers' knowledge is a combination of more than, less than, and exactly as what their lexicon informs them, and we provide a stochastic OT grammar based on the dual listing/generation model of Zuraw (2000) to account for our wug test results.

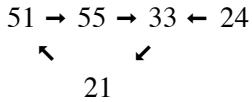
Monosyllabic adjectives in Taiwanese can undergo double reduplication to intensify its meaning. The tone pattern of double reduplication is summarized in (2): its medial syllable follows the general tone sandhi pattern in (1); the initial syllable, however, shows two different behaviors — while base tones 21 and 51 follow the general sandhi pattern, base 55, 33, and 24 have the tone 35 on the initial syllable. The tone pattern of the initial syllable can be transparently captured by the docking of a floating High tone to the left edge of the sandhi tone (vacuously) if the sandhi tone starts High (e.g., 21 → 51-51-21), and to the right edge if the sandhi tone starts Mid (55 → 35-33-55) (cf. Yip 1990).

Using a wug-test paradigm, we tested 14 Taiwanese speakers on how they did double reduplication for three types of monosyllables: Actual-Occurring words with existing reduplicated forms (AO), Actual-Occurring words without reduplicated forms (\*AO), and Accidental Gaps in the Taiwanese syllabary (AG). We calculated (a) the rate with which the expected sandhi tone was used on the medial syllable for each base tone, and (b) the rate with which the floating High correctly docked onto the initial syllable, *using the tone on the second syllable in the subjects' response as the sandhi tone*. The results, given in (3), indicate that the speakers' knowledge of the tone pattern is a combination of proper-, under- and overlearning from the lexicon. First, the speakers productively applied transparent generalizations on floating High docking and that the 24 tone cannot appear on the medial syllable to novel words, indicating that such generalizations are properly learned from the lexicon. Second, the opaque tone circle, though exceptionless in the lexicon, is not fully productive in novel words, indicating that it is underlearned. Third, among the opaque sandhis, there are instances of both overlearning and proper learning. The overlearning is manifested in the lower productivity of 51 → 55, which turns a tone with a shorter intrinsic duration into one with a longer duration (Lin 1988). It is likely that the speakers formed an analytical bias against this sandhi, given that the sandhi occurs in nonfinal positions, which are intrinsically shorter than the final position where the based tone appears. The proper learning is seen in the low productivity of 33 → 21 and high productivity of 55 → 33, which cannot be due to duration, but could be related to the low lexical frequency of the base tone 33 and the high lexical frequency of 55.

We modeled the simultaneously overlearned, underlearned, and properly learned tone pattern from the lexicon with the dual listing/generation model of Zuraw (2000) under the stochastic OT of Boersma and Hayes (2001). The model assumes that existing forms are lexically listed and are protected by highly ranked faithfulness constraint. But lower and stochastically ranked constraints can encode both lexical statistic patterns and phonetically-based generalizations. Transparent patterns in the lexicon can be derived through highly ranked markedness constraints even without lexical listing, and these markedness constraints will also ensure that the patterns are productive in novel words. Opaque patterns, however, cannot be derived by high ranking markedness constraints: for real words, they are derived via high ranking faithfulness constraints that protect lexical listings; but for novel words, such faithfulness constraints are unavailable, causing the patterns to be generally unproductive. Speakers do make certain generalizations about the opaque input-output mappings and encode them as constraints, and the ranking of these constraints is affected by the lexical frequencies of the mappings. These constraints are ranked relatively low in the grammar, preventing the patterns to be completely productive in novel words. The predictions of the our grammar, as given in (4), are relatively well matched with the wug results in (3).

**Data:**

(1) Taiwanese tone sandhi in non-XP-final positions:

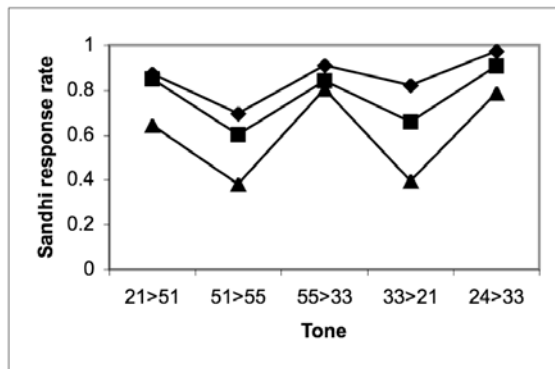


(2) Tone sandhi in double reduplication in Taiwanese:

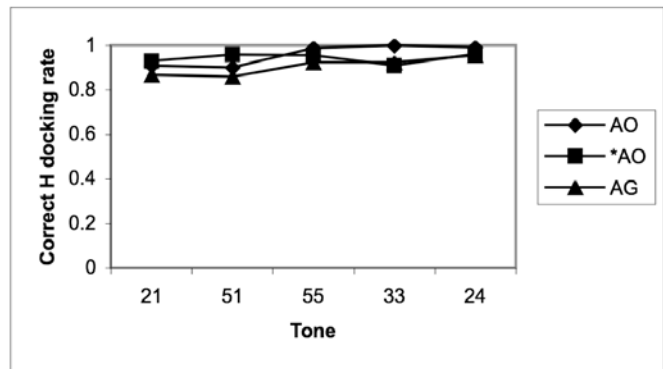
Monosyllabic adjective	Double reduplication
21	<b>51</b> -51-21
51	<b>55</b> -55-51
55	<b>35</b> -33-55
33	<b>35</b> -21-33
24	<b>35</b> -33-24

(3) Wug test results:

a. Sandhi response rate on  $\sigma_2$ :

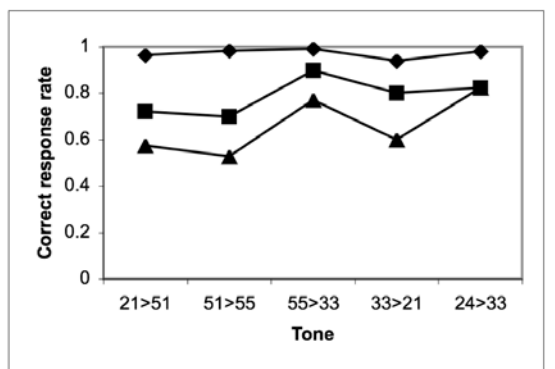


b. Correct floating H docking rate on  $\sigma_1$ :

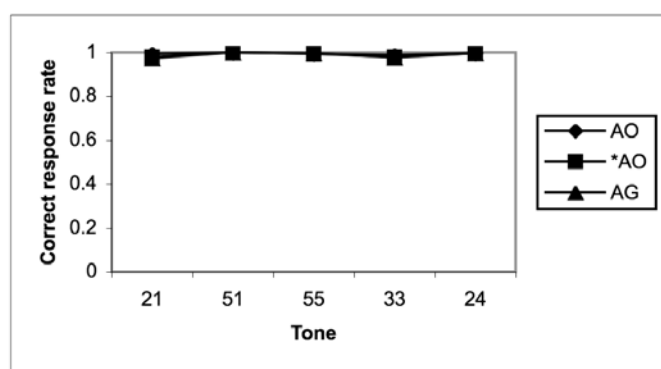


(4) Predictions of the grammar:

a. Sandhi response rate on  $\sigma_2$ :



b. Correct floating H docking rate on  $\sigma_1$ :



**Select References:**

Hsieh, Hsin-I (1970). The psychological reality of tone sandhi rules in Taiwanese. *CLS* 6, 489-503.  
 Zhang, Jie, Yuwen Lai, and Craig Sailor (to appear). Effects of phonetics and frequency on the productivity of Taiwanese tone sandhi. To appear in *CLS* 43.  
 Zuraw, Kie (2000). *Patterned exceptions in phonology*. Ph.D. dissertation, UCLA.