

Wug-Testing the “Tone Circle” in Taiwanese

Jie Zhang, Yuwen Lai, and Craig Turnbull-Sailor

The University of Kansas

1. Introduction

Tone sandhi refers to tonal alternations conditioned by adjacent tones or the prosodic and/or morphosyntactic position in which the tone occurs (Chen 2000, among others). For example, in Mandarin Chinese, an underlying complex contour tone 213 is only realized as such in word-final position. In nonfinal positions, it simplifies to either a high rising tone 35 or a low falling tone 21 depending on the tone on the final syllable, as shown in (1).¹

(1) Mandarin Chinese tone sandhi:

213 → 35 / 213

213 → 21 / {55, 35, 51}

Tone sandhi patterns are often extremely complex, sometimes so much so that they do not have workable synchronic analyses (see Chen 2004 for an example). This may have been caused by a number of internal and external factors. Internally, since a tone is primarily implemented with one articulator — the vocal folds, it is easily influenced by adjacent tones due to the difficulty in fast transitions from one vocal fold state to the next; tone is also easily influenced by segmental features due to properties of laryngeal anatomy. These factors determine that tones are inherently amenable to change. Externally, the close contact of different tone languages, especially in China, also causes mutual influence of the already complicated tone patterns.

Taiwanese tone sandhi is a case in point. The tonal inventory of Taiwanese is given in (2) (from Peng 1997: p.374). In this paper, we will only be concerned with tones on non-checked syllables (syllables that do not end in an obstruent coda; in other words, open syllables or syllables that end in a sonorant coda), which are not underlined in (2). Notice that among the five tones on non-checked syllables, the two falling tones 21 and 51 have the shortest intrinsic duration. This will be significant when we discuss the tone sandhi pattern.

Typical of a southern Min dialect, Taiwanese tone sandhi for non-checked syllables is characterized by a “tone circle” in non-XP-final positions (Chen 1987, Lin 1994), as shown in (3). The tone circle has two distinct properties. First, it represents a series of counterfeeding opacity in the sense of Kiparsky (1973); and given that the series of opacity is circular, there is no rule-ordering analysis for the pattern. Second, some of the sandhis in the circle are phonetically arbitrary in nature (Moreton 2004). Given that the sandhi occurs on nonfinal syllables, which are shorter than the final syllable due to lack of final lengthening (Oller 1973, Klatt 1975, Wightman et al. 1992, among others), we may consider the sandhi that turns a longer tone into a shorter tone to have a phonetic motivation based on duration. But only one of the sandhis in the circle has such durational basis — 33 → 21; the 21 → 51 and 55 → 33 sandhis do not have this durational motivation, and the 51 → 55 sandhi, which turns a shorter tone into a longer tone, is durationally aberrant. The sandhi outside the circle — 24 → 33 —

* This research is inspired by Hsieh (1970, 1975, 1976) and Wang (1993), and we owe a special thank to them. We also thank Jane Tsay and James Myers for providing us with a Taiwanese frequency corpus, Paul Boersma and Mietta Lennes for helping us with Praat scripts, and James Myers and Charles Chien-Jer Lin for their helpful comments. This research is supported by the University of Kansas General Research Funds #2301522 and #2301760.

¹ Tones here are transcribed in Chao letters (Chao 1968), whereby the speaker’s pitch range is represented on a five-point scale — “1” represents the lowest pitch, and “5” represents the highest pitch. A two-number sequence indicates that the tone is a long level tone or a contour tone; for example, 51 is a falling tone.

(4) Hypotheses:

On productivity:

- (a) Tone sandhis in the tone circle are unproductive.
- (b) The phonotactically driven 24 → 33 is productive.

On phonetic effects:

- (c) Within the tone circle, the productivity is the highest when the sandhi changes a longer tone into a shorter tone (33 → 21); the productivity is the lowest when the sandhi changes a shorter tone into a longer tone (51 → 55).
- (d) The two shorter tones 21 and 51 are the preferred sandhi tones.

2. Experimental Methods

2.1 Stimuli Construction

Following Hsieh (1970)'s experimental design, we constructed five types of disyllabic words in Taiwanese. The first type includes real words, denoted by AO-AO (AO = actual occurring morpheme). These words served as the control for the experiment. The other four types are wug words: *AO-AO, where both syllables are actual occurring morphemes, but the disyllable is non-occurring; AO-AG (AG = accidental gap), where the first syllable is actual occurring, but the second syllable is an accidental gap in Taiwanese syllabary; AG-AO, where the first syllable is an accidental gap and the second syllable is actual occurring; and AG-AG, where both syllables are accidental gaps. The AGs were hand-picked by the second author, who is a native speaker of Taiwanese. In each AG, both the segmental composition and the tone of the syllable are legal in Taiwanese, but the combination happens to be missing.

For each word type, we used five different tonal combinations: the tone on the first syllable was one of the five tones in the non-checked inventory; the tone on the second syllable was kept constant to 33, as Taiwanese tone sandhi is positionally conditioned, not tonally conditioned, and 33 is the closest to the middle of the pitch range. These five tonal combinations represent five sandhi types in non-XP-final positions. Eight words for each word type × sandhi type combination were used, making a total of 200 test words (8×5×5). We also used 160 filler words, which had tones other than 33 in final position.

Given that the main interest of the experiment lies in the comparison of different sandhi types under wug tests, we executed the following frequency controls across the sandhi types using a Taiwanese spoken corpus compiled by Tsay and Myers (2005). For real words AO-AO, we matched the frequency of the disyllables and the frequency of the first syllable using the frequency count of the syllable occurring in nonfinal positions. For the other two types of words with a real first syllable, namely *AO-AO and AO-AG, we also matched the frequency of the first syllable across the sandhi types. For AG-AO and AG-AG, although the first syllable is an accidental gap, its correct sandhi form can be either an existing syllable or an accidental gap; therefore, for the eight words within each sandhi type, we ensured that half of them have existing sandhi syllables, while the other half have non-existing sandhi syllables. Finally, to control for neighborhood effects to some extent, for all the wug words, we ensured that the disyllable is not a real word with *any* tonal combinations, not just the tonal combination we used. We specifically controlled for tonal neighbors as research on homophony judgment (Taft and Chen 1992, Cutler and Chen 1997), similarity rating (Vitevitch et al., in progress), phoneme (toneme) monitoring (Ye and Connine 1999), and legal-phonotactic judgment (Myers 2002) has all shown that phonemic tonal differences are perceptually less salient than segmental differences, which entails that tonal neighbors, in a sense, make closer neighbors.

2.2 Experimental set-up

The experiment was conducted with SuperLab (Cedrus) in the Phonetics and Psycholinguistics Laboratory at The University of Kansas. There were 360 stimuli in total (200 test items + 160 fillers). Each stimulus consisted of two monosyllabic utterances read by the second author, separated by an 800ms interval. The stimuli were played through a headphone worn by the subjects. For each stimulus, the subjects were asked to put the two syllables together and pronounce them as a disyllabic word in Taiwanese. Their response was collected by a Sony PCM-M1 DAT recorder through a 33-3018 Optimus dynamic microphone placed on the desk in front of them. The sampling rate for the DAT

recorder was 44.1kHz. The digital recording was then down-sampled to 22kHz onto a PC hard-drive using Praat. There was a 2000ms interval between stimuli. If the subject did not respond within 2000ms after the second syllable played, the next stimulus would begin. The stimuli were divided into two same-sized blocks (A and B) with matched stimulus types, and there was a five-minute break between the blocks. Subjects rotated in whether they took block A or block B first. Within each block, the stimuli were automatically randomized by SuperLab. Before the experiment began, there was a short introduction in Taiwanese that the subjects heard through the headphone, which explained their task both in prose and through examples. There was then a practice session that consisted of 14 words from the experiment (two of each from AO-AO, *AO-AO, AO-AG, AG-AO, and AG-AG; two real-word fillers, two wug fillers). The experiment began after a verbal confirmation from the subjects that they were ready. The entire experiment took around 45 minutes.

2.3 Subjects

Nine native speakers of Taiwanese (6 female, 3 male) participated in the experiment. They were recruited through flyers on KU campus and word of mouth. Each subject was paid \$10 for participating in the study.

2.4 Data Analyses

Due to the structure-preserving nature of the Taiwanese tone sandhi, the sandhi tones on the first syllable of the test words were transcribed by the three authors — a native speaker of Taiwanese (Lai), a native speaker of Beijing Mandarin (Zhang), and a native speaker of American English (Turnbull-Sailor), all of the whom are phonetically trained. Our agreement based on first-time auditory impression was about 95%. With help from pitch tracks in Praat, we agreed on virtually all the tokens. For the handful of tokens on which we did not reach an agreement, we took the native Taiwanese speaker Yuwen Lai's judgment.

We calculated both the correct-response and non-application rates for each word type \times sandhi type combination, and carried out two-way Repeated-Measures ANOVAs with word type and sandhi type as independent variables. We expected to see differences among different sandhi types in both the correct-response and non-application rates, especially in wug words. In particular, subjects should perform the 33 \rightarrow 21 sandhi with the greatest success and the 51 \rightarrow 55 sandhi with the least success. We also tallied the number of responses for each output tone in the tone circle. Due to the circular nature of the sandhis, if there is no preference for any tone to be the sandhi tone, there should be no difference in the number of occurrence among the tones; any differences would indicate that the tones with more frequent occurrences are the preferred sandhi tones. According to our hypotheses, we expected to see more frequent occurrences for the shorter tones 21 and 51. Finally, we tallied the number of responses for each output tone for the sandhi outside the circle — 24 to 33, expecting to see a higher number of 33 responses.

3. Results and Discussion

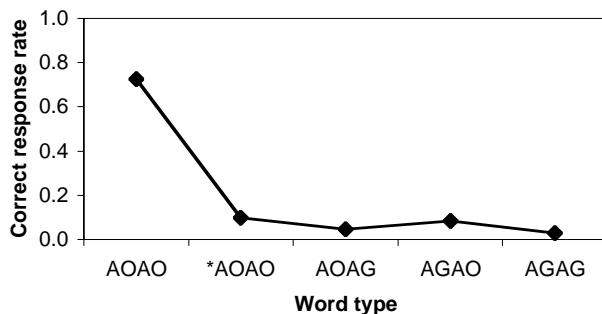
3.1 Correct-Response Rates

The results for correct-response rates according to word types and sandhi types are given in (5) and (6), respectively. A two-way Repeated-Measures Huynh-Feldt ANOVA, which adjusts the degrees of freedom to correct for sphericity violations, shows that the effect of word type is significant ($F(3.797, 29.656)=118.882, p<0.001$), so is the effect of sandhi type ($F(3.077, 24.619)=6.278, p=0.002$). Within the effect of word type, the real words had a significantly higher correct response rate than all the wug words, unsurprisingly. We also found that *AO-AO had a significantly higher rate than AO-AG and AG-AG, and AG-AO had a significantly higher rate than AG-AG. Within the effect of sandhi type, three sets of significant comparisons were found: 24 \rightarrow 33 — the phonotactically driven sandhi outside the circle — had a higher rate than three of the sandhis in the circle — 55 \rightarrow 33, 21 \rightarrow 51, and 51 \rightarrow 55; 33 \rightarrow 21, which is the sandhi that turns a longer tone into a shorter tone and thus has a durational basis, also had a higher rate than two other sandhis in the circle — 21 \rightarrow 51 and 51 \rightarrow 55; and 51 \rightarrow 55, which is the sandhi that turns a shorter tone into a longer tone and is thus

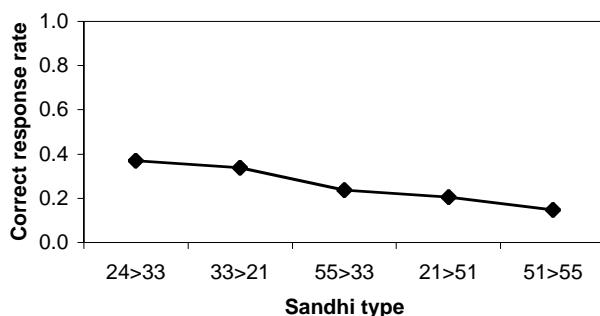
durationally aberrant, had a lower rate than both $24 \rightarrow 33$ and two other sandhis in the circle — $33 \rightarrow 21$ and $55 \rightarrow 33$.

The interaction between the two main effects, as shown in (7), is also significant ($F(8.869, 70.954)=2.456, p=0.017$), primarily due to the different behavior of the real words. There is no significant interaction between the two main effects among the wug words.

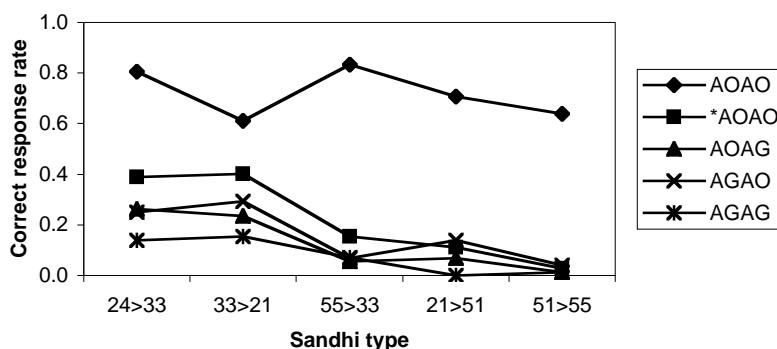
(5) Correct-response rates according to word types:



(6) Correct-response rates according to sandhi types:



(7) Correct-response rates — word type \times sandhi type:



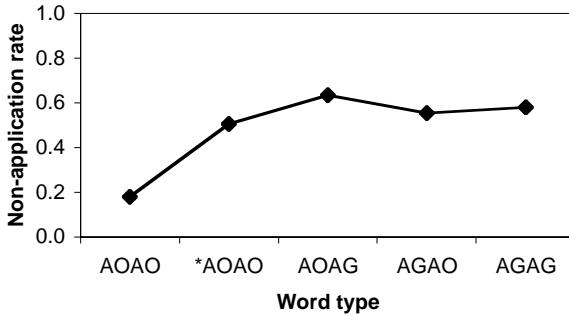
3.2 Non-Application Rates

The results for non-application rates according to word types and sandhi types are given in (8) and (9), respectively. A two-way Repeated-Measures Huynh-Feldt ANOVA shows that the effect of word type is significant ($F(3.999, 31.992)=65.116, p<0.001$), so is the effect of sandhi type ($F(3.314, 26.514)=33.256, p<0.001$). Within the effect of word type, the real words unsurprisingly had a significantly lower non-application rate than all the wug words; *AO-AO had a significantly lower rate

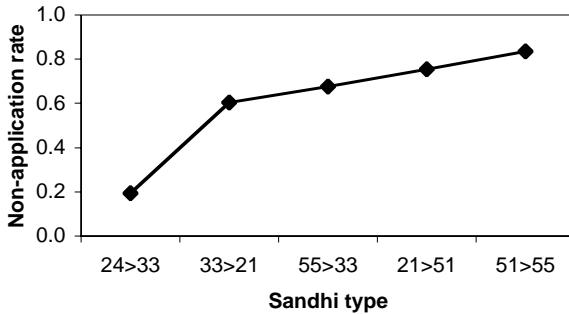
than AO-AG and AG-AG; and AG-AO had a significantly lower rate than AG-AG. Within the effect of sandhi type, three sets of significant comparisons were found: the phonotactically driven sandhi 24 → 33 had the lowest rate; the durationally driven 33 → 21 had a lower rate than two other sandhis in the circle — 21 → 51 and 51 → 55; and the durationally aberrant 51 → 55 had a higher rate than all other sandhis.. Expectedly, the patterns seen in non-application rates are generally the opposite to the correct-response rate patterns.

Due to the different behavior of the real words, the interaction between the two main effects, as shown in (10), is also significant ($F(9.110, 72.879)=4.795, p<0.001$). There is again no significant interaction between the two main effects among the wug words.

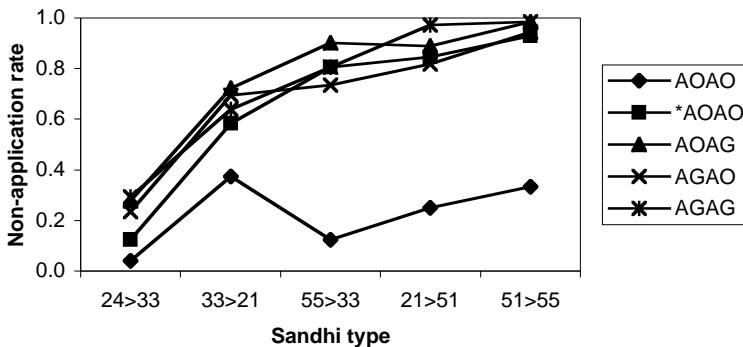
(8) Non-application rates according to word types:



(9) Non-application rates according to sandhi types:



(10) Non-application rates — word type × sandhi type:



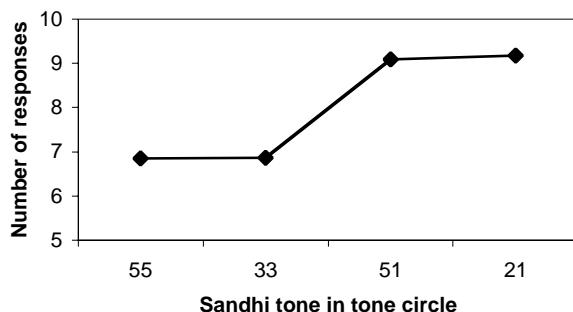
3.3 Sandhi Tone Preferences

The results for the number of responses for each output tone in the tone circle are given in (11). A Repeated-Measures ANOVA shows that there is a significant main effect for tone ($F(2.701,$

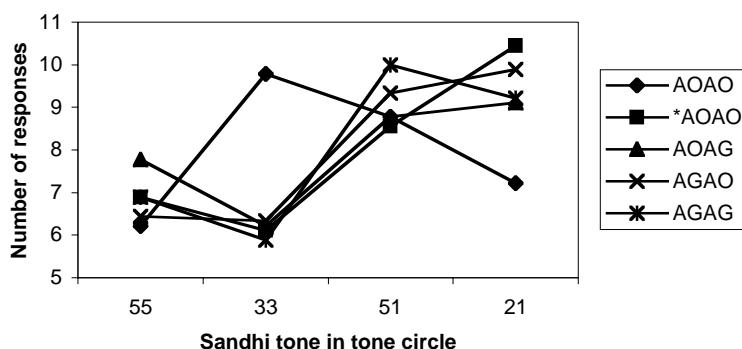
21.609)=5.414, $p=0.008$). Numerically, the two shorter tones 21 and 51 had more responses than the two longer tones 55 and 33. But only the 21 vs. 55 difference is significant.

The interaction between word type and tone, as shown in (12), is also significant ($F(6.695, 53.564)=3.355$, $p=0.005$). This is again caused by the different behavior of the real words. There is no significant interaction between word type and tone for the wug words.

(11) Output responses for sandhis in the tone circle:



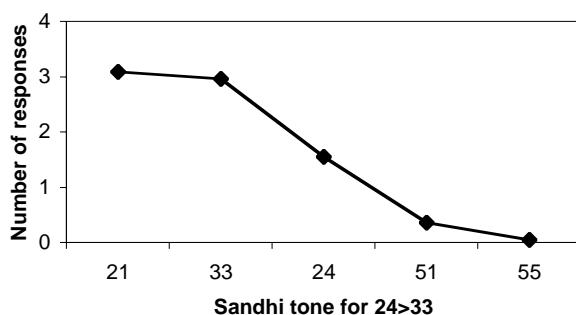
(12) Output responses for sandhis in the tone circle — word type \times tone:



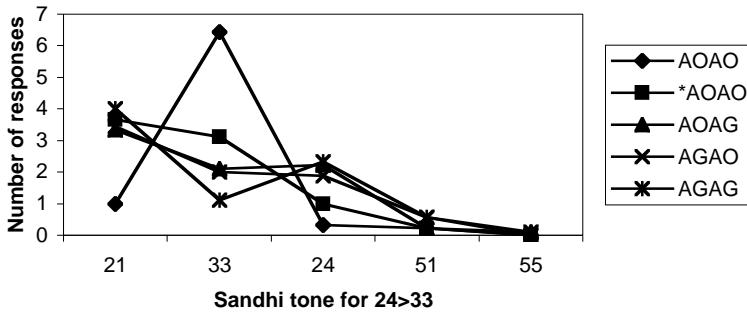
The results for the number of responses for each output tone for the sandhi $24 \rightarrow 33$ are given in (13). A Repeated-Measures ANOVA again shows that there is a significant main effect for tone ($F(1.704, 13.634)=8.784$, $p=0.005$). Interestingly, the tone that had the most responses was not the correct sandhi tone 33 or the non-application 24, but the low falling tone 21, presumably due to its short duration. Statistically, 21 and 33 had significantly more responses than 51 and 55.

Intuitively, the interaction between word type and tone, as shown in (14), is significant, as the real words mostly had the correct response 33, while the wug words did not.

(13) Output responses for $24 \rightarrow 33$:



(14) Output responses for 24 → 33 — word type × tone:



3.4 Summary

Our results on the wug tests can be summarized as follows. In terms of productivity, our hypothesis that the tone circle is unproductive is generally supported — only 11.5% of wug words in the tone circle had the correct sandhi; 82.9% had non-application. The hypothesis that 24 → 33 is productive, however, is not supported, as only 26.0% of the wug words with 24 on the first syllable had the correct sandhi. But we did find that the correct response rate was higher for 24 → 33 than for the sandhis in the tone circle, and the majority of the mistakes was not the non-application 24, but the short falling tone 21, indicating that the phonotactic generalization that 24 cannot occur in nonfinal positions is still quite effective. In terms of the effects of phonetics on the productivity of sandhis in the tone circle, all our hypotheses are corroborated: the highest productivity was found in the sandhi that turns a longer tone into a shorter tone (33 → 21); the lowest productivity was found in the sandhi that turns a shorter tone into a longer tone (51 → 55); and the shorter tones 21 and 51 were the preferred sandhi tones — the effect is significant for 21, and is in the right direction for 51. The durational effect can also be seen in the 24 → 33 sandhi through the preference for the short 21 as the sandhi tone.³

Crucially, the productivity differences cannot be accounted for by word frequency, which was either controlled for or did not exist as the word was a wug word. Furthermore, the productivity differences cannot be accounted for by the frequency of occurrence of the tones, either: according to our corpus, the Taiwanese tone frequencies, going from the most common to the least common, are 55, 51, 24, 33, and 21. Therefore, the preference for 21 as a nonfinal sandhi tone cannot be due to either *its* frequency or the frequency of its underlying tone 33.

4. Conclusion

In sum, our experimental results are consistent with previous research in that we have also shown that the opaque tone circle is largely unproductive when wug-tested. This casts further doubt on the productivity of opaque phonological patterns, in the same vein as works by Sanders (2001) and Sumner (2003). But we were able to more firmly establish that different sandhi patterns have different degrees of productivity, and certain tones are preferred as sandhi tones, most likely due to the more comprehensive frequency controls and statistical analyses. In particular, we have shown that the phonetic duration of tones is relevant to the synchronic grammar, as reflected in the higher productivity of the 33 → 21 sandhi, the lower productivity of the 51 → 55 sandhi, and the overall preference for 21 and dispreference for 55 as the nonfinal sandhi tone.

³ The preference for 21 as the sandhi tone for 24 → 33 may have been caused by a dialectal difference in Taiwanese. Thanks to James Myers for pointing this out to us.

References

- Chao, Yuen Ren (1968). *A grammar of spoken Chinese*. University of California Press, Berkeley, CA.
- Chen, Matthew Y. (1987). The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4: 109-149.
- Chen, Matthew Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge University Press, Cambridge, UK.
- Chen, Matthew Y. (2004). Changting Hakka tone sandhi: Analytical challenges. *Language and Linguistics* 5.4: 799-820.
- Cutler, Anne and Hsuan-Chih Chen (1997). Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics* 59.2: 165-179.
- Du, Tsai-Chwun (1988). *Tone and stress in Taiwanese*. Ph.D. dissertation, University of Illinois, Urbana Champaign.
- Hsieh, Hsin-I (1970). The psychological reality of tone sandhi rules in Taiwanese. In *Papers from the 6th Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago. 489-503.
- Hsieh, Hsin-I (1975). How generative is phonology. In E. F. Koerner (ed.), *The transformational-generative paradigm and modern linguistic theory*. John Benjamins, Amsterdam, The Netherlands. 109-144.
- Hsieh, Hsin-I (1976). On the unreality of some phonological rules. *Lingua* 38: 1-19.
- Kiparsky, Paul (1973). Abstractness, opacity, and global rules. In Osamu Fujimura (ed.), *Three dimensions of linguistic theory*. TEC Company, Ltd., Tokyo, Japan. 57-86.
- Kirchner, Robert (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27: 341-350.
- Klatt, Dennis H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics* 3: 129-140.
- Lin, Hwei-Bing (1988). *Contextual stability of Taiwanese tones*. Ph.D. dissertation, University of Connecticut.
- Lin, Jowan (1994). Lexical government and tone group formation in Xiamen Chinese. *Phonology* 11: 237-275.
- Moreton, Elliot (2004). Non-computable functions in Optimality Theory. In: John J. McCarthy (ed.), *Optimality Theory in phonology*. Blackwell Publishing, Malden, MA. 141-164.
- Myers, James (2002). An analogical approach to the Mandarin syllabary. *Chinese Phonology* 11, *Special Issue – Glide, Syllable and Tone*: 163-190.
- Myers, James and Jane Tsay (2001). Testing a production model of Taiwanese tone sandhi. *Proceedings of the Symposium on Selected NSC Projects in General Linguistics from 1998-2000*. National Science Council, Taipei, Taiwan. 257-279.
- Oller, D. Kimbrough (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54: 1235-1247.
- Peng, Shu-hui (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics* 25: 371-400.
- Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado, Boulder.
- Sanders, Nathan (2001). Preserving synchronic parallelism: Diachrony and opacity in Polish. In *Papers from the 37th Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago. 501-515.
- Sumner, Megan (2003). *Testing the abstractness of phonological representations in Modern Hebrew weak verbs*. Ph.D. dissertation, State University of New York at Stony Brook.
- Taft, Marcus and Hsuan-Chih Chen (1992). Judging homophony in Chinese: The influence of tones. In Hsuan-Chih Chen and Ovid J. L. Tzeng (eds.), *Language processing in Chinese*. Elsevier Science Publishers B. V., North-Holland. 151-172.
- Tsay, Jane, Jan Charles-Luce, and Yung-Song Guo (1999). The syntax-phonology interface in Taiwanese: Acoustic evidence. *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California, Berkeley. 2407-2410.
- Tsay, Jane and James Myers (2005). *Taiwanese spoken corpus*. National Chung Cheng University, Taiwan.
- Vitevitch, Michael, Jie Zhang, and Shinying Chu (in progress). *Phonological neighbors in Mandarin Chinese: What role does lexical tone play?* Ms. in progress, The University of Kansas.
- Wightman, Collin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707-1717.
- Wilson, Colin (2003). Experimental investigation of phonological naturalness. In Gina Garding and Mimu Tsujimura (eds.), *WCCFL 22 Proceedings*. Cascadilla Press, Somerville, MA. 533-546.
- Wilson, Colin (2005). *Learning phonology with substantive bias: an experimental and computational study of velar palatalization*. Ms., UCLA.
- Ye, Yun and Cynthia M. Connie (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes* 14: 609-630.
- Zhang, Jie (2002). *The effects of duration and sonority on contour tone distribution*. Routledge, New York.
- Zhang, Jie (2006). Testing the role of phonetic naturalness in Mandarin tone sandhi. Ms., The University of Kansas.

Proceedings of the 25th West Coast Conference on Formal Linguistics

edited by Donald Baumer,
David Montero, and Michael Scanlon

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

Proceedings of the 25th West Coast Conference on Formal Linguistics
© 2006 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 1-57473-415-6 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Zhang, Jie, Yuwen Lai, and Craig Turnbull-Sailor. 2006. Wug-Testing the “Tone Circle” in Taiwanese. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. Donald Baumer, David Montero, and Michael Scanlon, 453-461. Somerville, MA: Cascadilla Proceedings Project.

or:

Zhang, Jie, Yuwen Lai, and Craig Turnbull-Sailor. 2006. Wug-Testing the “Tone Circle” in Taiwanese. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. Donald Baumer, David Montero, and Michael Scanlon, 453-461. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1479.