Effects of Phonetics and Frequency on the Productivity of Taiwanese Tone Sandhi*

Jie Zhang, Yuwen Lai, and Craig Sailor
The University of Kansas

1 Introduction

The productivity of a linguistic pattern refers to its ability to apply to new items (Bybee 2001). The understanding of productivity is important to theoretical linguistics as it provides crucial evidence about the generalizations and cognitive abstractions that speakers make (Bybee 2001, Pierrehumbert 2003).

The productivity of a phonological process crucially depends on various properties of the process. Type frequency, which refers to the dictionary frequency of a particular pattern, has been shown by many works to positively influence the productivity of the pattern (Bybee 1985, 2001, Pierrehumbert 2003, 2006, among others). For example, Pierrehumbert (2006) showed that speakers' ability to extend velar softening as in *electric~electricity* to pseudo Latinate stems can be attributed to the regularity of this pattern in a relatively high number of word pairs. Token frequency, which refers to the frequency of occurrence of a unit, however, has been shown to distract away from the productivity of the pattern that the unit exhibits, as the high token frequency of a form weakens its associations with other forms (Bybee 1985, 2001, Moder 1992, among others). Moder (1992), for instance, found that in English irregular past tense formation, the *sweep~swept* class of words has a considerably higher token frequency than the *string~strung* class of words even though they have comparable type frequency, and the *sweep~swept* class is also less productive. The phonetic grounding of a pattern may also have an effect on productivity (Wilson 2003, 2006, Zhang and Lai 2006). For instance, Wilson (2003) showed that learners extend a nasal assimilation pattern (l → n / [+nasal]V__) more easily to novel words than a random alternation pattern (l → n / [+dorsal]V__). Finally, phonological opacity also has a strong effect on productivity, in that opaque patterns are often unproductive (Sanders 2001, Zhang *et al.* 2006). Sanders (2001), for example, has shown that Polish speakers do not extend the counterbleeding interaction between o-Raising and Final Devoicing to novel

---

* This research was 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, Mickey Waxman for statistical advice, and James Myers, Charles Chien-Jer Lin, Alan Yu, and Harald Baayen for their comments at various stages of this research. All remaining errors, however, are strictly our own. This research is supported by the University of Kansas General Research Funds #2301522 and #2301760 and a research grant from the Chiang Ching-Kuo Foundation for International Scholarly Exchange in 2006-2007.

words in a wug test (Berko 1958) experiment. The goal of the current research is to investigate the interaction of the effects of frequency, phonetics, and opacity on the productivity of Taiwanese tone sandhi, in which all three effects are operative.

2 Taiwanese tone sandhi

2.1 Basic patterns

Tone sandhi refers to tonal alternations conditioned by adjacent tones or the prosodic position in which the tone occurs. Taiwanese tone sandhi is typical of Southern Min dialects of Chinese in being positionally conditioned and characterized by circular opacity: tones in nonfinal positions of syntactic phrases undergo sandhi, and four out of five tones in the tonal inventory on non-checked syllables are involved in a "tone circle," as in (1) (Chen 1987). The sandhi 24 → 33 is phonotactically transparent as the ban on rising tone in non-XP-final positions is a true generalization in Taiwanese.

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

```
51 → 55 → 33 ← 24
  ↖     ↙
    21
```

The different sandhi patterns in Taiwanese also have different phonetic bases. Phonetic studies by Lin (1988) and Peng (1997) showed that the two falling tones 51 and 21 have considerably shorter intrinsic durations than 55, 33, and 24. Given that the sandhi occurs on nonfinal syllables, which are known to be shorter than final syllables due to the lack of final lengthening (Oller 1973, Klatt 1975, Wightman *et al.* 1992, among others), the 33 → 21 sandhi has a durational basis, as it turns a longer tone into a shorter tone; the 51 → 55 sandhi is an anti-duration change, as it turns a shorter tone into a longer tone; the other two sandhis are durationally neutral.

Finally, the different base tones in Taiwanese have different type and token frequencies, as estimated from a spoken corpus by Tsay and Myers (2005). In (2), the syllable type frequency of a tone refers to the number of different syllables in the Taiwanese syllabary represented in the corpus that can carry the tone; the morpheme type frequency of a tone refers to the number of different monosyllabic morphemes represented in the corpus that can carry the tone; and the token frequency of a tone is the number of occurrence of the tone in the entire corpus. Interestingly, these frequency counts do not put the tones in the same order. This provides us with an opportunity to study which frequency count has the greatest effect on productivity. Furthermore, the effects of frequency may also be manifested differently for speakers in different environments; for example, speakers who speak the language daily may exhibit a stronger frequency effect than those who only speak it occasionally.

(2) Tone frequency counts in Taiwanese:

Syllable type frequency: 55 > 51 > 24 >21 >33
Morpheme type frequency: 55 > 24 > 51 > 33 > 21
Token frequency: 55 > 24 > 33 > 51 > 21

2.2 Previous research on the productivity of Taiwanese tone sandhi

Previous research has shown that opacity, phonetics, and frequency all have an effect on the productivity of the sandhi pattern. Hsieh (1970) showed that the application rate of the entire Taiwanese tone sandhi pattern was as low as 10-30% in novel disyllabic words. Hsieh (1975), Wang (1993), and Zhang *et al*. (2006) all showed that the phonotactically transparent sandhi 24 → 33 had a higher application rate to novel words than opaque sandhis in the tone circle. Hsieh (1975) indicated that there might be a frequency effect, as the 55 → 33 sandhi, which has the highest counts for both type and token frequencies, had a higher application rate in novel words than other opaque sandhis in the tone circle. Finally, Zhang *et al*. (2006) reported that Taiwanese speakers tested at the University of Kansas showed a higher productivity for the duration reducing sandhi 33 → 21 than for the duration increasing sandhi 51 → 55.

2.3 The current research

Our current experiment aims to better quantify the effects of opacity, phonetic duration, and frequency on the productivity of Taiwanese tone sandhi by using a greater number of tokens and speakers and conducting more comprehensive statistical analyses. We also explicitly address the question of how these effects may differ for different speaker populations by comparing two groups of speakers — one from Chiayi, Taiwan, who uses the language in their daily functions; one from Lawrence, Kansas, who only uses the language occasionally.

Our hypotheses are the following. Regarding opacity, we expect the opaque tone circle to be generally unproductive, but the phonotactically transparent 24 → 33 to be generally productive. Regarding phonetics, we expect a significant correlation between productivity and duration reduction within the opaque tone circle; in particular, the duration reducing sandhi 33 → 21 should be the most productive, and the duration increasing sandhi 51 → 55 should be the least productive. Regarding frequency, on the other hand, we expect the most frequent sandhi across the board — 55 → 33 — to be the most productive, and the least frequent sandhi according to both morpheme and token frequencies — 21 → 51 — to be the least productive. We also hypothesize that the occasional speakers of the language from Lawrence may behave differently from the everyday speakers from Chiayi in that the former may have a stronger phonetic effect, while the latter may have a stronger frequency effect.

3 Experimental methods

3.1 Stimuli construction and the nature of subjects

The basic method of our experiment was to present the subjects with two monosyllables and ask them to pronounce the syllables together as a true disyllabic word in Taiwanese. Our analyses focused on the tone on the first syllable of the subjects' responses.

There are two within-subjects factors in the experiment. The first factor is Word Type. Following Hsieh (1970)'s experimental design, we constructed five types of disyllabic words in Taiwanese. The first type is 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. The second within-subject factor is Sandhi Type, which is determined by the tone on the first syllable of the disyllables. There are five different sandhi types, represented by the five tones in the tonal inventory on non-checked syllables — 24, 55, 33, 21, and 51. The tone on the second syllable was kept to a constant 33, as Taiwanese tone sandhi is positionally conditioned, not tonally conditioned, and 33 is the closest to the middle of the pitch range. 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 on the second syllable.

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 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), 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.

The experiment also has a between-subject factor, which is Speaker Group. We recruited 11 speakers from Chiayi (5 males, 6 females), Taiwan, who had an average age of 44 and all used Taiwanese in their daily functions, and 16 speakers on KU campus in Lawrence (6 males, 10 females), who had an average age of 31 and had been in the US for an average of 3.8 years — these speakers only spoke Taiwanese for an average of 45 minutes a week, usually in phone calls to their families in Taiwan. The Lawrence speakers are also considerably younger than the Chiayi speakers, and sociolinguistic studies have shown that younger speakers use Taiwanese less often due to the increasing influence of Mandarin (Sandel 2003, Scott and Tiun 2007, among others).

3.2 Experimental set-up

The experiment was conducted with SuperLab (Cedrus) in the Phonetics and Psycholinguistics Laboratory at KU and the Phonetics Laboratory in the Institute of Linguistics at National Chung Cheng University in Chiayi. There were 360 stimuli in total, and 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 instructed to put the two syllables together and pronounce them as a true disyllabic word in Taiwanese. Their response was collected by either a Sony PCM-M1 DAT recorder (in Lawrence) or a Marantz solid state recorder PMD 671 (in Chiayi) through a 33-3018 Optimus dynamic microphone placed on the desk in front of them. The sampling rate for both the DAT and solid state recorders was 44.1kHz. The DAT recording was down-sampled to 22kHz onto a PC hard-drive using Praat (Boersma and Weenink 2005). 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, the subjects heard a short introduction in Taiwanese 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.

3.3 Data analyses and hypotheses

Due to the structure-preserving nature of the sandhi (Du 1988, Lin 1988, Peng 1997, Tsay *et al*. 1999, Myers and Tsay 2001), 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 (Sailor), all phonetically trained — using a 1-5 scale with the help of pitch tracks in Praat. There was clear agreement for nearly all cases among the authors. In cases of disagreement, the judgment of the native Taiwanese author was taken. Based on the transcriptions, the correct response rates for the tone sandhis in each Word Type × Sandhi Type combination for each speaker were then calculated.[1]

Mixed-design ANOVAs with Word Type and Sandhi Type as within-subject factors and Speaker Group as a between-subject factor were conducted for the correct response rates. Regression analyses with the durational property of the sandhis and the type and token frequencies of base tones as predictors for the correct application rates of the opaque sandhis in wug words (*AO-AO, AO-AG, AG-AO, AG-AG) were also conduced for the two groups of speakers. All statistics were carried out in SPSS.

As stated in §2.3, we expect opacity, phonetics, and frequency to all have an effect on the productivity of the sandhi pattern. In particular, we expect a global effect of opacity: for both Lawrence and Chiayi speakers, the ANOVA and subsequent post-hoc analyses should show a significant effect for Word Type, with a significant difference between real words and wug words, and a significant effect for Sandhi Type, with a significant difference between the transparent sandhi 24 → 33 and the opaque sandhis in the tone circle. The effects of phonetics and frequency, however, may be Speaker-Group-specific, in that the correct application rates of the sandhis may be more strongly correlated with the durational property of the sandhis for the Lawrence speakers, but with frequency counts for the base tones for the Chiayi speakers, in the regression analyses.

4 Results

The main effects of Speaker Group, Word Type, and Sandhi Type on the correct response rate are plotted in Figure 1. The effect of Speaker Group is not significant: $F(1, 25)=0.014$, $p=0.907$. The effect of Word Type is significant: $F(3.445, 86.126)=370.323$, $p<0.001$. So is the effect of Sandhi Type: $F(2.886, 72.158)=22.403$, $p<0.001$). Post-hoc analyses with Bonferroni adjustments showed that real words AO-AO have a significantly higher correct response rate than all wug groups ($p<0.001$ for all paired comparisons): the average correct response rate is 71.2% for real words, but only 15.9% for wug words. Post-hoc

[1] We recognized that assuming structure-preservation in the sandhi tone transcriptions might have caused us to miss the subtle differences between accurately applied sandhis and slightly inaccurately applied sandhis and between sandhi tones in real words and sandhi tones in wug words. But the differences among the correct choice of the sandhi tone, the incorrect choice of the sandhi tone, and non-application of the sandhi were magnitudes bigger. We therefore opted to ignore the subtle pitch differences that might have existed and transcribe the sandhi tones categorically based on the structure-preservation assumption.

analyses also showed that the phonotactically transparent sandhi 24 → 33 has a significantly higher correct response rate than all opaque sandhis ($p<0.001$ for all paired comparisons except for 24 → 33 vs. 33 → 21, for which $p=0.15$): the average correct response rate is 42.1% for 24 → 33, but only 23.1% for the opaque sandhis.

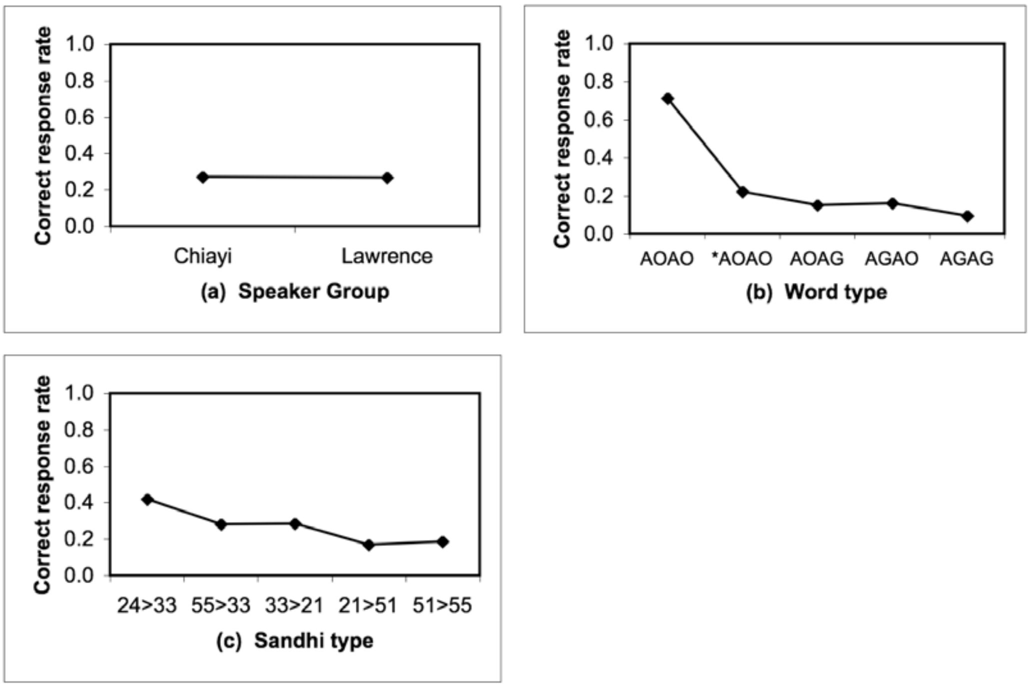

Figure 1: Effects of (a) Speaker Group, (b) Word Type, and (c) Sandhi Type on the correct response rates for tone sandhis in Taiwanese.

The interactions between the within-subject factors and the between-subject factor are plotted in Figure 2. The Word Type × Speaker Group interaction is not significant: $F(3.445, 86.126)=0.916$, $p=0.447$. Neither is the Sandhi Type × Speaker Group interaction: $F(2.886, 72.158)=2.135$, $p=0.106$.

Two separate ANOVAs and subsequent post-hoc analyses on the two subject groups showed that for both Chiayi and Lawrence speakers, real words AO-AO have a significantly higher correct response rate than all wug groups at the $p<0.001$ level; for Chiayi speakers, the phonotactically transparent sandhi 24 → 33 has a significantly higher correct response rate than all the opaque sandhis in the tone circle at the $p<0.05$ level except for 55 → 33, and for Lawrence speakers, 24 → 33 has a significantly higher correct response rate than all opaque sandhis at the $p<0.05$ level except for 33 → 21. We therefore consider it relatively uncontroversial that Taiwanese speakers in general had considerably more trouble in (a) the application of tone sandhis in wug words, and (b) the application of opaque tone sandhis.

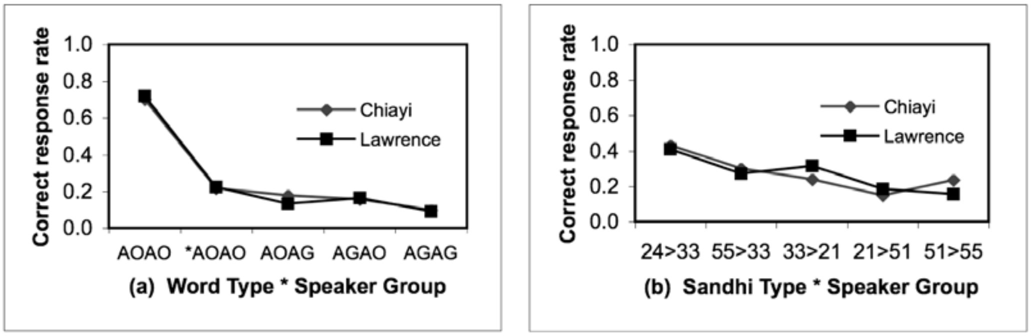

Figure 2: Interactions (a) between Word Type and Speaker Group and (b) between Sandhi Type and Speaker Group on the correct response rates for tone sandhis in Taiwanese.

To investigate the potential differences between the two groups of speakers caused by phonetics and frequency, we conducted a further mixed-design ANOVA with only the correct response rates for the opaque sandhis in wug words. The main effect of Word Type is still significant: $F(3, 75)=7.51$, $p<0.001$. So is the main effect of Sandhi Type: $F(1.984, 49.611)=8.587$, $p<0.005$. The interactions between the within-subject factors and the between-subject factor were regraphed in Figure 3, which highlighted the differences between the two speaker groups. The Word Type × Speaker Group interaction is still not significant: $F(3, 75)=1.447$, $p=0.236$. But the Sandhi Type × Speaker Group interaction is now significant: $F(1.984, 49.611)=3.428$, $p<0.05$).

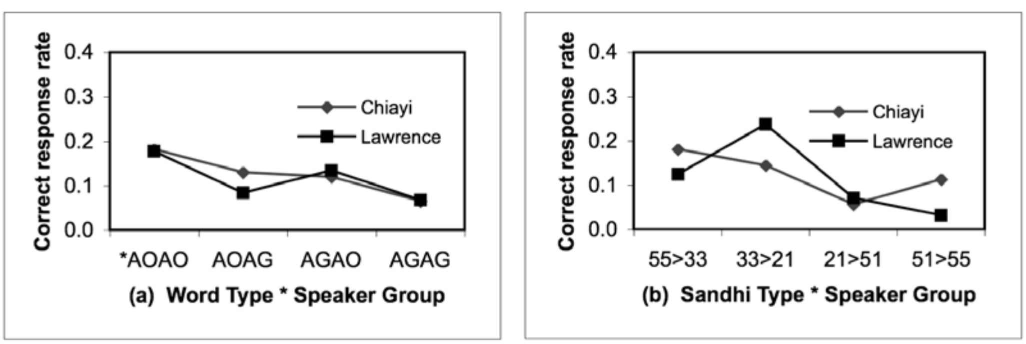

Figure 3: Interactions (a) between Word Type and Speaker Group and (b) between Sandhi Type and Speaker Group on the correct response rates for opaque tone sandhis in Taiwanese wug words.

The significant interaction between Sandhi Type and Speaker Group indicates that the two groups of speakers behaved differently in the application of opaque sandhis to wug words. Figure 3b shows that for Chiayi speakers, the highest productivity is found in the sandhi whose base tone has the highest frequency counts (55 → 33), and the lowest productivity is found in the sandhi whose base tone has the lowest morpheme type and token frequencies (21 → 51); for Lawrence speakers, however, the sandhi with the greatest duration reduction — 33 → 21 — has the highest productivity, while the sandhi with the greatest duration increase — 51 → 55 — has the lowest productivity. These results indicate that Chiayi speakers' behavior may be closely related to frequencies, while Lawrence speakers' behavior may be closely related to phonetic duration.

To further investigate the effects of frequencies and duration on sandhi productivity in the two speaker groups, we conducted a number of regression analyses, with syllable type frequency, morpheme type frequency, and token frequency of the base tone and the amount of duration reduction of the tone sandhi as predictors for the correct response rate for the sandhi. The values for these predictors are given in Table 1. The frequency counts are from Tsay and Myers (2005) and have been *log*-transformed. The duration reduction data were deduced from the durations of the tones read in isolation reported Lin (1988). A positive value indicates a duration decrease from the base tone to the sandhi tone, and a negative value indicates a duration increase.

	55 → 33	33 → 21	21 → 51	51 → 55
Syllable type frequency	2.499	2.391	2.407	2.467
Morpheme type frequency	2.872	2.715	2.687	2.736
Token frequency	5.085	5.058	4.972	5.020
Duration reduction (ms)	1	69	-9	-59

Table 1: Predictors for correct response rates for opaque tones sandhis in Taiwanese wug words used in regression analyses. The frequency counts refer to the *log* frequencies of the base tone in the sandhi. Duration reduction indicates the amount of duration decrease from the base tone to the sandhi tone.

The regression results for Chiayi and Lawrence speakers are given in Table 2 and Table 3, respectively. Table 2 shows that for Chiayi speakers, the only significant predictor for correct response rates is token frequency of the base tone. The effect of duration reduction is not significant. Lawrence speakers, however, behaved quite differently, as shown in Table 3. Duration reduction is a highly significant predictor for correct response rates; but none of the frequency counts has a significant effect.

In summary, the experimental results showed that, in terms of the effects of opacity, our hypothesis that the tone circle is generally unproductive is supported: the ANOVA results indicated that both Chiayi and Lawrence speakers had considerably greater trouble applying the opaque sandhis in wug words. Further tabulations indicated that in wug words, the opaque sandhis only had an average of 12.0% correct response rate, but a 81.7% non-application rate. Contra our hypothesis, the phonotactically transparent sandhi 24 → 33 is also not entirely productive, with average correct response rates of 42.1% for all words and 31.9%

for wug words. But these rates are significantly higher than those for the opaque sandhis. Moreover, a tabulation of the non-application rate for 24 → 33 in wug words returned the figure 31.9% — strikingly low compared to the 81.7% for opaque sandhis, indicating the phonotactic generalization "a rising tone cannot occur in nonfinal positions" is quite effective.

Our hypothesis that the effect of phonetics and frequency may be Speaker-Group-specific, in that Lawrence speakers will be more strongly affected by duration, while Chiayi speakers will be more strongly affected by frequency counts, is also borne out by the regression analyses.

	Sig.	R^2	β
Syllable type frequency	n.s. (p=0.224)	0.104	0.322
Morpheme type frequency	n.s. (p=0.055)	0.238	0.488
Token frequency	*** (p=0.016)**	**0.350**	**0.591**
Duration reduction (ms)	n.s. (p=0.494)	0.034	0.184

Table 2: Regression results for Chiayi speakers. "*Sig.*" indicates the significance value of the predictor, with "n.s." representing "not significant at the p<0.05 level," and "*" representing "significant at the p<0.05 level." R^2 indicates the amount of variation in the correct response rate that can be accounted for by the predictor. β is the standardized coefficient.

	Sig.	R^2	β
Syllable type frequency	n.s. (p=0.166)	0.132	-0.364
Morpheme type frequency	n.s. (p=0.899)	0.001	0.035
Token frequency	n.s. (p=0.079)	0.204	0.452
Duration reduction (ms)	***** (p<0.001)**	**0.616**	**0.785**

Table 3: Regression results for Lawrence speakers. "*Sig.*" indicates the significance value of the predictor, with "n.s." representing "not significant at the p<0.05 level," and "***" representing "significant at the p<0.001 level." R^2 indicates the amount of variation in the correct response rate that can be accounted for by the predictor. β is the standardized coefficient.

5 Discussion

It is somewhat surprising that the only frequency effect that we found for the Chiayi speakers was of that of token frequency. This is inconsistent with earlier research, which established that productivity was usually positively correlated with type frequency, and that a high token frequency could in fact distract away from the pattern that the unit exhibits (Bybee 1985, 2001, Moder 1992, among others). We surmise that this discrepancy may have been caused by the fact that the processes studied by earlier research were generally morphophonological processes with many lexical exceptions and instantiated by relatively small numbers of types; but the sandhi processes here are exceptionless in the

Taiwanese lexicon, instantiated by many syllables, morphemes, and words. Therefore, the type frequency effect may have reached a ceiling. Moreover, the token frequency here is not the frequency of occurrence of a very specific lexical item in the corpus, but the frequency of occurrence of a very general tone type. Therefore, the high token frequency does not distract away from a general pattern, as there is not a more general pattern to be distracted away from, but further instantiates the sandhi pattern that the tone exhibits.

It is also interesting to ponder how the difference in frequency effects arose for the two groups of speakers. We assume that the Chiayi speakers exhibited a strong frequency effect due to their everyday usage of the language as the combined result of both their location and their age; the Lawrence speakers, who were younger and used Taiwanese less often to start with, had stopped using the language regularly after they moved to the US, causing the attrition of the lexical strength of all words, which consequently caused the frequency effects to fall below a certain threshold of detection.

We also surmise that the difference in phonetic effects between the two speaker groups is correlated with their difference in frequency effects. Due to the incompatibility of the two effects — for instance, phonetic effects predict a high productivity of 33 → 21 and a low productivity of 51 → 55; but frequency effects predict a high productivity for 55 → 33 and a low productivity for 33 → 21, for Chiayi speakers, the phonetic effects may have simply been overridden by the strong frequency effects. But when the frequency effects weaken due to the lack of usage, as for our Lawrence speakers, the phonetic effects surface. Therefore, failing to detect the phonetic effects for Chiayi speakers does not necessarily mean that they do not exist.

We recognize that our experimental result on the speaker group effect may have two opposing interpretations. One interpretation is that it demonstrates the usage-based nature of phonology *à la* Bybee (1985, 2001), as it shows that the speakers' behavior is heavily influenced by the context of usage. But conversely, it could also be interpreted as demonstrating that the only true linguistic effects are those of opacity and phonetics, as they are shared by all speakers; the effects of frequency are metalinguistic. Without committing ourselves to either of these extreme positions, we would rather like to propose that a more fruitful alterative is to consider a system that incorporates both processing and phonological factors that can predict the observed behaviors of both Chiayi and Lawrence speakers, and we take this as the ultimate goal of our current research project. This echoes the position espoused in works such as Albright (2002), Albright and Hayes (2003), Hayes and Londe (2006), Pierrehumbert (2003, 2006), Wilson (2006), and Zuraw (2000, 2007).

6 Conclusion

Our research investigates the interaction of the effects of frequency, phonetics, and opacity on the productivity of phonological patterns, with Taiwanese tone sandhi, in which all three effects are operative, as a case study. We also explicitly address how these effects may differ for speaker populations that have different patterns of language usage.

Our results show that opacity trumps both phonetics and frequency in predicting the productivity of tone sandhi patterns in Taiwanese: the phonotactically transparent sandhi 24 → 33 applies to wug words with a significantly greater success rate than opaque sandhis in the tone circle despite the fact that the rising tone 24 does not have the greatest frequency counts, nor does the sandhi enjoy the greatest durational advantage. Moreover, the two groups of speakers behaved similarly in this respect, indicating that the effect is robust against changes in the language environment.

The effects of frequency are only evident for frequent users of the language from Chiayi, Taiwan, for whom the highest and lowest sandhi productivities corresponded to the highest and lowest token frequencies for the base tones, respectively. The effects of phonetics, on the other hand, only surface for infrequent users of the language from Lawrence, KS, for whom the sandhi with the greatest duration reduction was the most productive, while the sandhi with the greatest duration increase was the least productive. The lack of phonetic effects for Chiayi speakers may have been caused by the strong frequency effects, which overrode the phonetic effects due to their incompatibility. But when the frequency effects weaken due to lack of usage, as is the case for the Lawrence speakers, the phonetic effects surface.

The results of the current study indicate that a fruitful direction for research on phonological productivity is in the integration of processing and phonological factors into a unified model of speakers' behavior. They also point to a potentially important methodological issue in phonological data collection: different subject populations may exhibit different phonological behaviors depending on their patterns of language use, and we as researchers need to be aware of this potential effect and incorporate it to the best of our abilities in our analyses and modeling of phonology.

References

Albright, Adam (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language* 78: 684-709.

Albright, Adam & Bruce Hayes (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90: 119-161.

Berko, Jean (1958). The child's learning of English morphology. *Word* 14: 150-177.

Boersma, Paul and David Weenink (2005). *Praat: doing phonetics by computer* (Computer program). Retrieved Nov., 2005, from http://www.praat.org/.

Bybee, Joan L. (1985). *Morphology: A study of the relation between meaning and form*. Benjamins, Philadelphia.
Bybee, Joan L. (2001). *Phonology and language use*. Cambridge University Press, Cambridge, UK.
Chen, Matthew Y. (1987). The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4: 109-150.
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.
Hayes, Bruce and Zsuzsa Czyráky Londe (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23: 59-104.
Hsieh, Hsin-I (1970). The psychological reality of tone sandhi rules in Taiwanese. In *Papers from the 6th Meeting of the Chicago Linguistic Society*, 489-503. Chicago, IL.
Hsieh, Hsin-I (1975). How generative is phonology. In *The transformational-generative paradigm and modern linguistic theory*, E. F. Koerner (ed.), 109-144. John Benjamins, Amsterdam, The Netherlands.
Hsieh, Hsin-I (1976). On the unreality of some phonological rules. *Lingua* 38: 1-19.
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.
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*, 257-279. National Science Council, Taipei, Taiwan.
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.
Pierrehumbert, Janet B. (2003). Probabilistic phonology: Discrimination and robustness. In *Probabilistic linguistics*, Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds.), 177-228. The MIT Press, Cambridge, MA.
Pierrehumbert, Janet B. (2006). The statistical basis of an unnatural alternation. In *Laboratory Phonology VIII, Varieties of Phonological Competence*, L. Goldstein, D. H. Whalen, and C. Best (eds.), 81-107. Mouton de Gruyter, Berlin.
Sandel, Todd L. (2003). Linguistic capital in Taiwan: The KMT's Mandarin language policy and its perceived impact on language practices of bilingual Mandarin and Tai-gi speakers. *Language in Society* 32.4: 523-551.
Sanders, Nathan (2001). Preserving synchronic parallelism: Diachrony and opacity in Polish. In *Papers from the 37th Meeting of the Chicago Linguistic Society*. 501-515. Chicago, IL.
Scott, Mandy and Hak-Khiam Tiun (2007). Mandarin only to Mandarin-plus: Taiwan. *Language Policy* 6: 53-72.
Taft, Marcus and Hsuan-Chih Chen (1992). Judging homophony in Chinese: The influence of tones. In *Language processing in Chinese*, Hsuan-Chih Chen and Ovid J. L. Tzeng (eds.), 151-172. Elsevier Science Publishers B. V., North-Holland.
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*, 2407-2410. University of California, Berkeley.
Tsay, Jane and James Myers (2005). *Taiwanese spoken corpus*. National Chung Cheng University, Taiwan.
Wang, H. Samuel (1993). Taiyu biandiao de xinli texing (On the psychological status of Taiwanese tone sandhi). *Tsinghua Xuebao* (*Tsinghua Journal of Chinese Studies*) 23: 175-192.
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 *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, Gina Garding and Mimu Tsujimura (eds.), 533-546. Cascadilla Press, Somerville, MA.

Wilson, Colin (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30: 945-982.

Ye, Yun and Cynthia M. Connie (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes* 14: 609-630.

Zhang, Jie and Yuwen Lai (2006). Testing the role of phonetic naturalness in Mandarin tone sandhi. *Kansas Working Papers in Linguistics* 28: 65-126. Also under review.

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*, Donald Baumer, David Montero, and Michael Scanlon (eds.), 453-461. Cascadilla Proceedings Project, Somerville, MA.

Zuraw, Kie (2000). *Patterned exceptions in phonology*. Ph.D. dissertation, UCLA.

Zuraw, Kie (2007). Exceptions and lexical processing. Invited talk presented at *the 43rd Meeting of the Chicago Linguistic Society*, Chicago, IL.