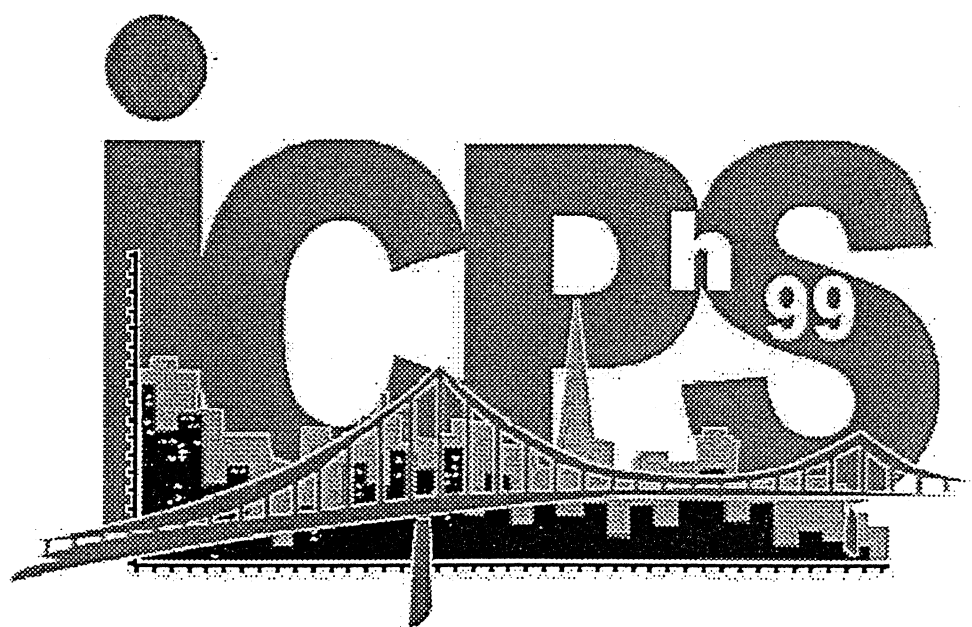

Proceedings of The XIVth International Congress of Phonetic Sciences



San Francisco
1-7 August 1999

Sponsored by: University of California, Berkeley
Co-sponsored by: Acoustical Society of America, Australian Speech Science & Technology
Association, European Speech Communication Association, International Phonetic Association,
Linguistic Society of America

Edited by: John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey

THE SYNTAX-PHONOLOGY INTERFACE IN TAIWANESE: ACOUSTIC EVIDENCE

Jane Tsay*, Jan Charles-Luce⁺, and Yung-Song Guo*

*National Chung Cheng University, Taiwan, ⁺State University of New York at Buffalo, USA

ABSTRACT

Taiwanese tone sandhi (TTS) is a phrase-level tone alternation between the syntactically defined juncture and the context forms of a morpheme. The goals of the acoustic study were to investigate (i) whether the alternation between juncture and context tones is categorical; (ii) whether there are cues for the domain of application of TTS other than the tone alternation itself. The question of categoricity was addressed by comparing the F0's of juncture and context forms that have been described as having the same surface tones, and results showed that F0's of these two tones were not significantly different. Phrase-final lengthening was found in juncture position, showing that there is a prosodic correlate of TTS available in adult speech. We argue that syntactically defined categorical phrase-level phonology like TTS must be mediated by prosodic structure in order to avoid a paradox in language acquisition.

1. INTRODUCTION

1.1. Goals of the study

An acoustic study was conducted to investigate two issues. The first concerned the categoricity of Taiwanese tone sandhi; the other was to see whether there are non-tonal, prosodic correlates for Taiwanese tone sandhi available in adult speech.¹

1.2. Taiwanese Tone Sandhi

Taiwanese tone sandhi (TTS) is a phrase-level pattern where all monosyllabic morphemes have two tonal alternants: one with juncture tone Tj (in syntactically defined pre-juncture position, [1, 2]) and the other with context tone Tc (elsewhere), as schematized below:

(1) Tc Tc Tj # Tc Tj # Tc Tc Tc Tj #

The tone pattern itself can be represented in tone letters, H (High), M (Mid), and L (Low), respectively, as in Table 1.²

	Juncture	Context
Tone 1	H	M
Tone 2	LM	M
Tone 3	HL	H
Tone 4	L	HL
Tone 5	M	L

Table 1. Tone Sandhi in Taiwanese (Chia-Yi dialect)

1.3. Issue 1: Categoricity in TTS

Although TTS has been tacitly assumed in the literature to be categorical, i.e. involving alternations from one tone category to another tone category [3, 4], this has not been demonstrated acoustically. Therefore, the first goal of the acoustic study was to

confirm the categoricity of TTS.

The categoricity claim makes two predictions. The first prediction is that the tone alternation between juncture and context positions in TTS is categorical. For example, it predicts that juncture H (Hj) and context H (Hc) have the same phonetic (F0) properties.

The second prediction is that the apparent neutralization of the two distinct juncture tones, Hj and LMj, into a single context tone Mc is indeed complete.

1.4. Issue 2: Prosodic cues for TTS

If TTS is indeed categorical, another issue arises. Categoricity means that TTS creates potential lexical items (homophones): the juncture form of word A can be identical to the context form of word B. For example, [te¹] can be the juncture form of the word meaning "emperor" or the context form of the word meaning "earth."

This ambiguity causes a paradox for language acquisition. That is, a child learning Taiwanese has to learn words (to know if a syllable is the juncture form of word A or the context form of word B) and syntax (to know if a syllable is in juncture or context position) at the same time. The most plausible way into this vicious circle must be non-tonal (at least non-categorical) and non-syntactic, in particular, gradient, non-neutralizing prosodic cues (for detailed argumentation see [5]). Therefore, the second goal of the acoustic study was to investigate if there are such prosodic correlates for TTS in adult speech (e.g. phrase-final vowel lengthening and/or F0 dropping).

Peng [6] did find lengthening and lowering in phrase-final position, but the results were not conclusive. First, no statistical analyses were done. Second, there was no serious attempt to control segmental and tonal coarticulation of neighboring syllables, an important factor affecting F0 in Mandarin [7, 8], and presumably in Taiwanese as well.

2. METHODOLOGY

2.1. Acoustic Correlates

2.1.1. F0's The primary acoustic correlate of tone is fundamental frequency. Therefore, the first categoricity prediction was addressed by comparing the F0's of juncture and context forms that have been described as having the same surface tones, e.g. Hj and Hc. (However, it has to be kept in mind that if there are intonation effects at the phrase level, Hj might be lowered in F0 due to phrase-final declination. We call this the "Intonation Hypothesis.") The second categoricity prediction was addressed by comparing the F0's of the context Mc that alternates with Hj and the context Mc that alternates with LMj.

2.1.2. **Duration** Since the domain of application of TTS is essentially a phonological phrase [9], it is reasonable, based on what occurs in other languages, to hypothesize that the most salient, non-tonal prosodic cue to a TTS juncture is the duration of the pre-juncture syllable. Therefore, we checked for prosodic correlates by comparing the duration of juncture and context forms, e.g. the duration of syllables with H_j and H_c.

2.2. Experiment 1

2.2.1. **Materials** Several aspects were taken into consideration while designing the materials. First, for easier and more accurate segmentation, the target syllables all started with an obstruent. Second, we hoped to cover all four tone categories, i.e. H, M, L, HL.³ Third, the target syllable must be a legal morpheme. Fourth, the target syllable must be ambiguous in that it can be the juncture form of morpheme A or the context form of morpheme B, for example, [te^L], as illustrated in Table 2.

Juncture	Context	gloss
te ^L	te ^{HL}	emperor
te ^M	te ^L	earth

Table 2. A target syllable in two positions

Following these constraints, fourteen syllables were chosen as targets. A pair of sentences was then made up for each target syllable. To avoid effects from coarticulation, the preceding and following syllables of the target syllable were phonetically matched (especially in tone) as closely as possible in the two sentences. The sentences were written to be as natural as possible.

Fourteen pairs of sentences were prepared and used for comparing F0's and durations of the target syllables. The target syllable always occurred approximately in the middle of the sentence. Examples are given in (2) and (3). (CL = classifier)

- (2) hit e hoŋ^M - te^L # kiu^{LM} tsin tse
that CL loyal-emperor ball very many
"That emperor has many balls."
- (3) hit liap aŋ^M te^L - kiu^{LM} # tsin tat tsī
that CL red earth-ball very worth money
"That red globe is very valuable."

Another three pairs of sentences were created to check whether H_j and H_c were really neutralized to Mc in context position. The sentences in each pair were identical except that in one sentence the target M, si^M in the following examples, alternates with H, while in the other sentence it alternates with LM, as shown in Table 3.

Juncture	Context	gloss
si ^H	si ^M	poetry
si ^{LM}	si ^M	time

Table 3. Neutralizing syllables

A pair of examples is given in (4) and (5).

- (4) gun^H tao^H tiŋ^L tioŋ^M-kək^H si^M-po^L
my home subscribe China poetry-report
"At home, we subscribe to China Poetry Journal."
- (5) gun^H tao^H tiŋ^L tioŋ^M-kək^H si^M-po^L
my home subscribe China times-report
"At home, we subscribe to China Times."

Therefore, there were seventeen pairs of target sentences. The sentences were presented in Chinese characters, which disambiguate homophones.

2.2.2. **Subjects** Seventeen paid subjects who were native speakers of Taiwanese participated in this experiment. They also spoke Mandarin fluently. The age range was from late 20's to early 30's. They were graduate students at State University of New York at Buffalo when this experiment was conducted. However, ten of them were dropped out before the recordings were analyzed due either to dialect variation or incomplete recording.⁴ Analyses were done based on recordings of the remaining three female and four male subjects.

2.2.3. **Procedure** The experiment was conducted in a sound-treated room. Sentences described above with the same number of filler sentences were written in Chinese characters on cards and shown to each subject in random order one sentence at a time. Subjects were asked to read out the sentences and the reading was tape-recorded. The entire set of sentences was repeated five times, but the first time was not analyzed for better control of naturalness in speech. Among the fourteen pairs of sentences that were designed to test categoricity, nine were discarded because some subjects thought those sentences were not natural. The remaining five pairs of sentences covered H, M, and L tones in the targets and none had a HL tone. (The behavior of HL will be discussed in Experiment 2.)

2.2.4. **Measurements** Subjects were tape-recorded one at a time, using a Shure SM2 microphone and a Marantz PMD430 portable cassette deck. Subjects' speech was then digitized at a sampling rate of 10 kHz per second, via a 16 bit analog-to-digital converter (Gradient Technology, Inc. DeskLab TM216 SCSI Audio). Measurements were made on a VAX Workstation (4000 Model 60) from a visual waveform display using a digital waveform editor. The F0's of the beginning, middle, and end points of the target syllables were measured. Duration of the target syllables was also measured.

2.2.5. **Results and Discussion** Regarding neutralization, an ANOVA showed that the F0's of Mc's were not significantly different. ($p > 0.1$)

	Beginning	Middle	End
Mc (H _j)	164	156	151
Mc (LM _j)	160	152	147
difference	4	4	4

Table 4. F0 differences in neutralizing syllables

We can conclude that they were completely neutralized. As to the comparison of F0's of T_j and T_c, an ANOVA

showed no significant difference at the beginning and end points of the syllable, but there was a significant difference ($F(1,6)=12.5, p=0.012$) at the mid point of the syllable.

	Beginning	Middle	End
Tj	180	170	170
Tc	180	173	169
difference	0	-3*	1

(* $p < 0.05$)

Table 5. F0 difference of all five pairs

However, for the mid point, only one pair was significantly different, and in this pair F0 was lower for Tj than for Tc, which is consistent with the Intonation Hypothesis mentioned above and the findings in Peng [6]. Therefore, the categoricity issue is confirmed.

The durations of Tj and Tc were significantly different as expected ($F(1,6)=17.9, p<0.01$).

Juncture	194
Context	147
difference	+47**

(** $p < 0.01$)

Table 6. Duration differences (in msec.)

Therefore, pre-juncture (phrase-final) lengthening is available in adult's speech and might be used by children as a prosodic cue for domains of TTS.

2.3. Experiment 2

2.3.1. Motivation Although the results in Experiment 1 were consistent with our predictions, there remained other issues to address. In particular, differences in F0 between juncture and context forms may not be due to the Intonation Hypothesis, but instead it may be that the F0's of a context tone are influenced by its associated juncture tone (which has been assumed by many phonologists to be the "underlying" tone, e.g. [3, 4]). We call this the "Underlying Tone Hypothesis."

Therefore, Experiment 2 was conducted to verify whether there was influence from the so-called "underlying" tone. Moreover, since Experiment 2 was conducted in Taiwan, we also had better control of subjects. Finally, we took the opportunity of this replication to improve our materials.

2.3.2. Materials While the Intonation Hypothesis predicts that Tj should always be lower than Tc because of phrase-final lowering, the Underlying Tone Hypothesis predicts that the F0's of Tc might be influenced by its underlying tone (i.e. its associated Tj), as shown in Table 7.

Tj "underlying"	Tc "derived"	Influence on Tc from Tj	Predictions
H	M	raised	$M_j < M_c$
M	L	raised	$L_j < L_c$
L	HL	lowered at the beginning	$HL_j > HL_c$ at the beginning
HL	H	lowered at the end	$H_j > H_c$ at the end

Table 7. Predictions of the Underlying Tone Hypothesis

Eight pairs of sentences were prepared using the same method in Experiment 1. All four context tones (M, L, HL, H) were included in the targets.

2.3.3. Subjects Twelve native speakers of Taiwanese, aged between 30 and 50, participated in Experiment 2. All of them were male and were administration staff at National Chung Cheng University. Although they also spoke Mandarin, Taiwanese was the primary language used in their daily life. Also, they all spoke the Chia-Yi dialect of Taiwanese.

2.3.4. Procedure and Measurements The procedure in Experiment 2 was the same as in Experiment 1, except that the materials were displayed on a computer screen and that the entire set of sentences was repeated six times. The recording was done in a sound-treated room using a SONY portable cassette recorder. The recordings were digitized at a sampling rate of 10kHz per second by a 16 bit analog-to-digital converter and the measurements were based on the visual waveform by a digital waveform editor (Kay Elemetrics CSL 4300B). The durations as well as the F0's of the beginning and end points of the target syllables were measured.

2.3.5. Results and Discussion Each tone was analyzed separately. For all of the four tones, Tj was always lower than Tc, consistent with the Intonation Hypothesis. The average F0's are given in Table 8 and Table 9.

Tone	Beginning		
	Juncture	Context	difference
M	129	130	-1
L	116	118	-2
HL	159	164	-5*
H	144	148	-4*

(* $p < 0.05$)

Table 8. F0's at the beginning point

Tone	End		
	Juncture	Context	difference
M	123	125	-2
L	102	109	-7**
HL	125	130	-5
H	142	146	-4

(** $p < 0.01$)

Table 9. F0's at the end point

The differences were very small (from 1 to 7 Hz). In most of the cases, the differences were not significant, although some did show significance (i.e. the end point of L, the beginning point of HL and H).

Thus the Underlying Tone Hypothesis was not supported, since HLj was lower than HLc at the beginning and Hj was lower than Hc at the end. In other words, the differences in F0 in phrase-final position were not caused by the so-called underlying tone, but rather by phrase-final lowering.

As in Experiment 1, the duration of Tj was always longer than the duration of Tc across twelve subjects, and the differences were significant. The average durations are given in Table 10.

Juncture	162
Context	120
difference	+42***

(*** p < 0.001)

Table 10. Duration differences (in msec.)

Again, this prosodic correlate in the pre-juncture position of TTS was found in adult speech.

3. GENERAL DISCUSSION

3.1. Categoricity

Results showed that FO's of juncture tones and context tones were only occasionally significantly different, and where there was a difference, it could be ascribable to intonation. In other words, juncture H, M, L, and HL are the same tone categories as context H, M, L, and HL, respectively. That the neutralization of H_j and LM_j into Mc, which naturally occurs without intonation effects, is complete, further confirms the categoricity of TTS.

3.2. Vowel Duration

We checked for prosodic correlates by comparing the duration of juncture and context forms and found significant differences. Gradient FO differences ascribing to intonation were also occasionally found. The results therefore suggest that cues such as these are available for children learning phrase-level phonology like TTS. The vicious circle in language acquisition, described in Section 1.4, can be broken by the prosodic cues.

3.3. Syntax-Phonology Interface

The case of TTS thus provides a good example of prosodic bootstrapping [10]. Moreover, we claim that phonology which refers to syntax (e.g. [11]) must be mediated by prosodic structure in order to avoid a paradox in language acquisition. In particular, the acoustic properties of TTS and the logic of the acquisition problem lead us to conclude that no categorical phonological patterns at the phrasal level can refer directly to syntax (contra [12, 13]), but must instead be mediated by prosodic structure (e.g. [14, 15]).

ACKNOWLEDGMENTS

Part of this study (Experiment 1) was funded by an NIH NIDCD grant (DC-00957) to Jan Charles-Luce. We thank Paul Luce and James Myers for helpful discussions.

NOTES

1. Taiwanese belongs to the Southern Min family of Chinese.
2. Taiwanese has seven lexical tones. Two of them only appear with checked syllables ending with an unreleased stop. The two checked tones are typically shorter than the other five tones. Since duration was one of the interests of this study, we only examined the five non-checked tones.
3. LM does not appear in context position.
4. Dialects of Taiwanese vary in vowels as well as in tone sandhi.

REFERENCES

- [1] Chen, M. Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4.109-149.
- [2] Lin, Jo-Wang. 1994. Lexical Government and Tone Group Formation in Xiamen Chinese. *Phonology* 11.237-275.
- [3] Cheng, R. L. 1968. Tone Sandhi in Taiwanese. *Linguistics* 41.19-42.
- [4] Yip, Moria. 1980. *The Tonal Phonology of Chinese*. Doctoral

dissertation, MIT. Reprinted in 1990 by Garland Publishers under the title *Tonal Phonology in Chinese*.

- [5] Tsay, Jane. In press. Bootstrapping into Taiwanese Tone Sandhi. *Chinese Languages and Linguistics V: Interactions in Language*. Symposium Series of the Institute of History and Philology No. 2, Academia Sinica, Taipei.
- [6] Peng, S. 1993. Effects of Prosodic Position and Tonal Context on Taiwanese Tones. *Ohio State University Working Papers in Linguistics* 44, 169-190.
- [7] Shen, X. S. 1990. Tonal coarticulation in Mandarin. *Journal of Phonetics* 18, 281-295.
- [8] Xu, Y. 1994. Production and perception of coarticulated tones. *Journal of Acoustical Society of America*, 95(4), 2240-2253.
- [9] Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3.371-405.
- [10] Morgan, J. and K. Demuth. 1996. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- [11] Hayes, B. 1990. Precompiled Phrasal Phonology. In S. Inkelas and D. Zec (eds.) *The Phonology-Syntax Connection*. Cambridge University Press. 85-108.
- [12] Kaisse, E. 1985. *Connected Speech: The Interaction of Syntax and Phonology*. London: Academic Press.
- [13] Kaisse, E. 1990. Toward a typology of postlexical rules. In S. Inkelas and D. Zec (eds.) *The Phonology-Syntax Connection*. Cambridge University Press. 127-143.
- [14] Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press.
- [15] Odden, D. 1990. Syntax, lexical rules and postlexical rules in Kimatuumbi. In S. Inkelas and D. Zec (eds.) *The Phonology-Syntax Connection*. Cambridge University Press. 259-278.