Comparison of Phonetic Naturalness between Rising-Falling and Falling-Rising Tonal Patterns in Taiwan Mandarin

Chia-Hsin Yeh

Department of Linguistics and Germanic, Slavic, Asian, and African Languages
Michigan State University
chiahsin@msu.edu

Abstract

The new tonal pattern, a falling-rising pitch sequence, emerges from nominal reduplications produced by young children, teenagers, and motherese in Taiwan Mandarin. The pattern implies a child-like speech style and an intimate relationship between speakers and addressees, so application of the pattern usually denotes innocence, ingeniousness, and coyness. The morphophonological pattern sounds like a Tone3 (T3) - Tone2 (T2) sequence, a falling-rising pitch contour. The new pattern has an opposite contour of the T2-T3 sequence, a rising-falling contour, which is a phonological output derived from two consecutive T3s based on T3 Sandhi in Mandarin. The study investigated phonetic naturalness of the two tonal patterns in terms of the pitch changes within a monosyllabic domain and the ratio between pitch changes and duration in T2 and T3. Eight Taiwan Mandarin speakers, four males and four females, were recruited to pronounce the two tonal patterns. The current production results showed that the pitch changes of both T2 and T3 syllables in the T3-T2 sequence were significantly fewer than in the T2-T3 sequence, so the new pattern causes fewer articulatory efforts. As a result, the new T3-T2 pattern, a falling-rising sequence, seems more phonetically natural than the T2-T3 pattern, a rising-falling sequence. The results can explain why young Mandarin speakers acquire the T3-T2 pattern earlier than the T2-T3 sequence of the T3 Sandhi.

1. Introduction

Mandarin Chinese has four tonal types, which are a high level tone (Tone 1), a middle rising tone (Tone 2), a low falling tone (Tone 3) and a high falling tone (Tone 4), and two tone sandhi rules, which are T3 Sandhi and Neutral Tone Sandhi [3]. In the T3 Sandhi, the first T3 in two consecutive T3s becomes T2 for various phonetic, morphological, and syntactic factors [9]. In the Neutral Tone Sandhi, pitch contour of the neutral tone, usually in the final syllable, is neutralized to the preceding full tone, as illustrated in Table 1, primarily for phonetic reasons [9]. The pitch contour of the neutral syllables was previously considered as an extension of preceding tonal types or a tonal interpolation between preceding and following tonal types. However, the neutral tones were believed to be specified with a middle pitch value [4]. The tonal patterns of nominal reduplications in Mandarin are usually derived from the Neutral Tone Sandhi.

Table 1: *Pitch Contour of the neutral tone [3]*

Preceding Tones	Examples	Neutral Tones
T1- high level	*ma1ma0* 'mother'	*ma0*- half low
T2 - mid rising	*shu2shu0* 'uncle'	*shu0*- middle
T3 - low falling	*jie3jie0* 'sister'	*jie0*- half high
T4 - high falling	*di4di0* 'brother'	*di0*- low

In recent years, the Neutral Tone Sandhi of nominal reduplications has been gradually replaced by a T3-T2 sequence in Taiwan Mandarin. The T3-T2 pattern is therefore believed to evolve from a variation of the Neutral Tone Sandhi among young speakers. Then the pattern spreads through all speech communities in Taiwan, especially among teenagers. The pattern is extensively applied in an intimate relationship, such as lovers, good friends, and siblings, so application of the pattern usually denotes some additional meanings, such as innocence, ingeniousness, and coyness. The application of the pattern is conditioned by both the tonal types and the lexical types, as illustrated in Table 2. Generally, the new pattern is more likely to be applied in kinship terms than other nouns. In addition, it is more likely to be derived from a T2 or T3 base of nominal reduplications than a T4 base. The T1 base is the least likely to be replaced by the T3-T2 pattern.

Table 2: *Examples of the T3-T2 pattern*

Base		Gloss	Normal	New Pattern
T1	媽 ma1	'mother'	ma1ma0	ma3ma2
	花 hua1	'a flower'	?hua1hua0	**hua3hua2*
	丁 ding1	'name'	ding1ding0	**ding3ding2*
T2	叔 shu2	'uncle'	shu2shu0	shu3shu2
	娃 wa2	'a doll'	wa2wa0	wa3wa2
	妮 ni2	'name'	ni2ni0	ni3ni2
T3	姊 jie3	'sister'	jie3jie0	jie3jie2
	寶 ba3	'a baby'	bao3bao0	bao3bao2
	偉 wei3	'name'	wei3wei0	wei3wei2
T4	弟 di4	'brother'	di4di0	di3di2
	蛋 dan4	'an egg'	?dan4dan0	?dan3dan2
	莉 li4	'name'	li4li0	**li3li2*

The T3-T2 pattern produces a falling-rising pitch sequence which is an opposite pitch contour of the phonological output, the T2-T3 pattern, derived from two consecutive T3s in the T3 Sandhi. The T3 Sandhi has not been fully acquired among three year olds [8], but the T3-T2 patterns are extensively found among two year olds. Developmental differences of the two tonal patterns motivate the current study. This study focuses on phonetic characteristics of the two patterns to investigate whether the phonetic factors influence the first language acquisition of the two patterns.

2. Definitions of Phonetic Naturalness

Both T2 and T3 of the T3-T2 pattern are derived tones based on the assumption that the new pattern was evolved from the Neutral Tone Sandhi. The pitch values of the derived tones are usually found to be significantly different from those of the lexical counterparts. For example, the derived T2 of T3 Sandhi was found to be significantly lower than the lexical T2 in Mandarin [6]. The morphological Tone 35, a middle rising tone, was found to be significantly higher than the lexical Tone 35 in Cantonese [12]. The pitch differences between the derived tones and the lexical counterparts can be correlated to different degrees of phonetic naturalness. Hence, the phonetic changes of the derived T2 and T3 are first investigated. If there is a significant difference, then the phonetic naturalness will be further argued.

In this study, the phonetic naturalness is defined in terms of the pitch changes and the ratio between pitch changes and duration [9, 10] in T2 and T3. The two criteria of the phonetic naturalness are examined. First of all, with everything being equal, the fewer pitch changes the tonal pattern fluctuates from a pitch peak to a valley, the more phonetically natural the tonal pattern will be. For example,

cross-linguistically, level tones are more common than contour tones [10]. The pitch changes actually take more articulatory efforts than expected because there is a speed limit of pitch changes for laryngeal movements to coordinate with supralaryngeal gestures [9]. The fewer articulatory efforts the tonal pattern takes, the more phonetically natural the pattern will be.

Secondly, T2, a rising tone, with longer duration is more natural than with shorter duration, and T3, a falling tone, with shorter duration is more natural than with longer duration at a normal speech range. The speed limit of pitch changes is correlated with tonal types [7, 9]. Generally speaking, T2, a rising tone, usually takes longer duration [9], and T3, a falling tone, usually takes short duration [7]. In other words, as to the same amount of pitch changes, it takes longer duration to rise, but it takes shorter to fall [9]. As a result, longer duration is more natural for T2, but shorter duration is more natural for T3.

In order to evaluate phonetic naturalness in terms of the two criteria, two phonetic characteristics, pitch height and duration, of T2 and T3 are both measured and analyzed. Three loci of pitch changes and duration are targeted, as illustrated in Figure 1. As to the pitch changes, the three loci are pitch changes of T2 syllables, pitch changes of T3 syllables, and boundary pitch changes between T2 and T3. As to the duration, the three loci are T2 duration, T3 duration and bi-syllabic duration of the two sequences.

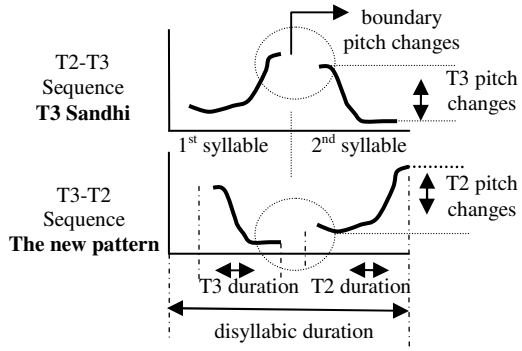

Figure *1: Three loci of pitch changes and duration in the T2-T3 and T3-T2 tonal patterns.*

3. Methods

A reading format is adopted in the production experiment. In addition to the two tonal conditions, the T3-T2 pattern and the T3 Sandhi, two other tonal conditions, a lexical T3-T2 sequence and the Neutral Tone Sandhi, are also included as control variables in the experiment. Hence, there are four tonal conditions total in the experiment.

3.1 Participants

Eight Taiwan Mandarin native speakers, four males and four females, from 22 years old to 25 years old, were recruited at Michigan State University. Six of the participants live in northern Taiwan and the other two participants live in middle Taiwan. All of them have the minimum dialectal background, such as Taiwanese and Hakka.

3.2 Stimuli

The six kinds of syllables, *ma, shu, jie, di, ge,* and *bo* are chosen as a reduplication base from the six kinship terms, 媽媽 *ma1mao0* 'mother,' 叔叔 *shu2shu0* 'uncle,' 姊姊 *jie3jie0* 'elder sister,' 弟弟 *didi0* 'younger brother,' 哥哥 *ge1ge0* 'elder brother,' and 伯伯 *bo2bo0* 'uncle.' The six syllables are reduplicated to create four different disyllabic stimuli in terms of the four tonal conditions: the T3 Sandhi, the Neutral Tone Sandhi, the T3-T2 pattern, and a lexical T3-T2 sequence, which constitute 24 stimuli total. Taking the syllable *jie* for example, the four disyllabic stimuli are: 姊姊 jie3jie3 (underlying) → *jie3jie0* (surface) 'elder sister,' an instance of the Neutral Tone Sandhi, 解解 jie3jie3 → *jie2jie3* 'try to resolve,' an instance of the T3 Sandhi, 姊姊 jie3jie3→ *jie3jie2* 'elder sister with an intimate relationship,' an instance of the new T3-T2 patter, and 解結 jie3jie2 → *jie3jie2* 'to untie a knot,' an instance of a lexical T3-T2 sequence.

3.3 Procedure

Participants were provided with a list of 24 stimuli in a context and were asked to practice reading the 24 stimuli in the context. Then the 24 stimuli showed up in a sentential carrier: 我說__這個字 *wo3 shuo1 __ zhe4 ge0 zi4* 'it is __ that I say' on a laptop screen at random. They were asked to repeat the stimuli in the carrier for five times and their pronunciations were recorded by Praat Version 5.1 [Boersma 2009] with the dynamic microphone, SHURE SM48. When the stimulus of the new T3-T2 pattern shows up on the screen, participants would be reminded to read the sentence in a childlike manner. Participants would not be explicitly told to pronounce the new pattern. If participants cannot come up with the new pattern, more contexts would be provided. In fact, all participants had no problem to apply the new pattern.

4. Results

All recordings were spectrographically analyzed by Praat 5.1 [2] in terms of the 10 phonetic variables: the highest pitch height, the lowest pitch height, pitch changes and duration in both T2 and T3 syllables, boundary pitch changes and disyllabic duration. Two extreme values, a maximum and a minimum value, out of five repetitions were excluded from the results. Hence, there were 72 (6 syllables x 4 tonal sequences x 3 repetitions) recordings total for each participant. For each tonal condition, there were 144 (8 participants x 6 syllables x 3 repetitions) stimuli total. A T-test analysis was conducted to investigate the two dimensions, pitch differences and phonetic naturalness, as shown in Table 3. As to the analysis of pitch differences, results of the highest pitch height, the lowest pitch height and duration of both T2 and T3 syllables were analyzed to investigate pitch differences of derived tones, which is called the lexical contrast to compare derived T2 and T3 of the T3-T2 pattern and lexical T2 and T3 of the lexical T3-T2 sequence. As to the analysis of phonetic naturalness, results of pitch changes of T2 and T3, boundary pitch changes, and disyllabic duration were analyzed to compare phonetic naturalness between the T3-T2 sequence derived from the new pattern and the T2-T3 sequence derived from the T3 Sandhi, which is called the sandhi contrast.

Table 3: *Two Dimensions of T-Test Analysis*

Pitch differences of derived tones	lexical contrast	highest, lowest pitch and duration
Phonetic naturalness	sandhi contrast	pitch changes and disyllabic duration

4.1 T-Test of Pitch differences

In the lexical contrast, the highest pitch and the lowest pitch of T2 and T3 syllables in the new T3-T2 pattern was

slightly higher than in the lexical T3-T2 sequence, but there was no significant difference: the highest pitch of T2, $t(286)= 0.0113$, $p= 0.4955$, the lowest pitch of T2, $t(286)= 0.8669$, $p= 0.1934$, the highest pitch of T3, $t(286)= 0.8449$, $p= 0.1994$, and the lowest pitch of T3, $t(286)= 1.3161$, $p= 0.0945$, as shown in Figure 2. As to the duration, T2 of the lexical T3-T2 sequence was longer than that of the new T3-T2 pattern, which was close to a significant level, $t(286)= 1.5654$, $p= 0.0593$. T3 of the lexical T3-T2 sequence was significantly longer than that of the new T3-T2 pattern, $t(286)= 7.6459$, $p= 0.000< 0.001$***.

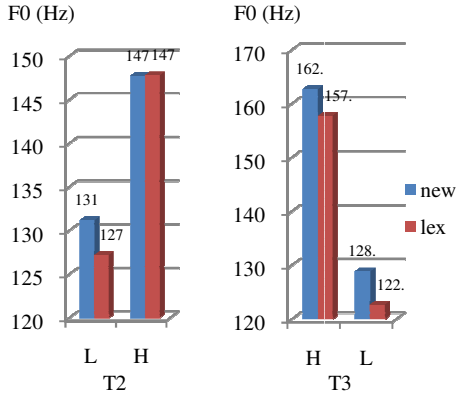

Figure *2: Highest Pitch (H) and Lowest Pitch (L) in T2 (left) and in T3 (right) in the lexical contrast between the new pattern (new) and the lexical T3-T2 sequence (lex).*

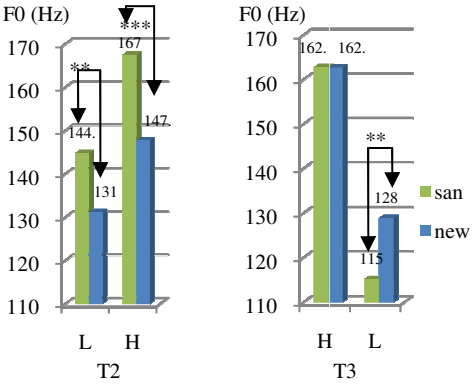

Figure *3: Highest Pitch (H) and Lowest Pitch (L) in T2 (left) and in T3 (right) in the sandhi contrast between the new T3-T2 pattern (new) and the T3 Sandhi (san).*

4.2 T-Test of Phonetic Naturalness

In the sandhi contrast, the highest pitch and the lowest pitch of T2 syllables in the T3 Sandhi was significantly higher than in the new T3-T2 pattern: the highest pitch of T2, $t(286)= 3.6514$, $p= 0.0002< 0.001$***, and the lowest pitch of T2, $t(286)= 2.7362$, $p= 0.0033< 0.01$**. The highest pitch of T3 syllables in the T3 Sandhi was not significantly different from that in the new T3-T2 pattern, $t(286)= 0.0139$, $p= 0.4945$, but the lowest pitch of T3 syllables in the T3 Sandhi was significantly lower than in the new T3-T2 pattern, $t(286)= 2.8571$, $p= 0.0023< 0.01$**. These results are shown in Figure 3. As to the duration, T2 of the T3 Sandhi was significantly shorter than that of the new T3-T2 pattern, $t(286)= 6.5372$, $p= 0.000< 0.001$***. T3 of the T3 Sandhi was significantly longer than that of the new T3-T2 pattern, $t(286)= 9.9811$, $p= 0.000< 0.001$***.

The pitch changes of T2 and T3 syllables in the T3 sandhi were significantly larger than in the new T3-T2 pattern: the pitch changes of T2 syllables, $t(286)= 4.8959$, $p= 0.000< 0.001$***, and the pitch changes of T3 syllables, $t(286)= 5.6519$, $p= 0.000< 0.001$***. The boundary pitch changes between T2 and T3 syllables of the T3 Sandhi were also significantly larger than those of the new T3-T2 pattern, $t(286)= 2.924$, $p= 0.0019< 0.01$**. The results are illustrated in Figure 4. As to the disyllabic duration, the new T3-T2 pattern was slightly longer than the T3 Sandhi, but there was no significant difference, $t(286)= 0.9345$, $p= 0.1754$, as illustrated in Figure 5.

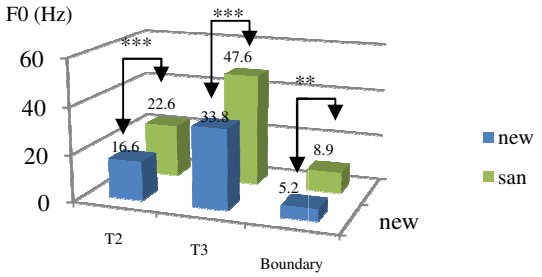

Figure *4: Pitch changes of T2 syllables, T3 syllables and tonal boundary in the new T3-T2 pattern (new) and the T3 Sandhi (san).*

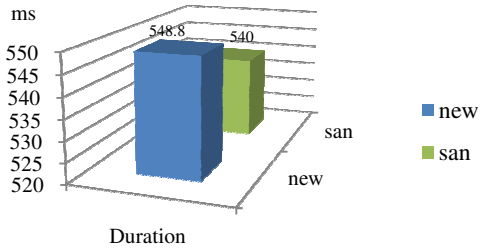

Figure *5: Disyllabic duration of the new T3-T2 pattern (new) and the T3 Sandhi (san).*

5. Discussion

The pitch values of derived T2s and T3s in the new T3-T2 patterns were not significantly different from those of lexical counterparts in the lexical T3-T2 sequence, but the monosyllabic duration of T2s and T3s in the new T3-T2 patterns was significantly shorter than in the lexical T3-T2 sequence. The results indicated that the new pattern is not only judged as a T3-T2 sequence perceptually, but also produced as a T3-T2 sequence acoustically. However, the derived T2 and T3 in the new pattern have not been fully neutralized to their lexical counterparts. The shorter duration of the new pattern suggests a shorter processing time in speech production. The shorter processing time may result from lexical familiarity of kinship terms or the bound unit of the new T3-T2 pattern.

When it comes to the sandhi contrast between the new pattern and the T3 Sandhi, the highest pitch and the lowest pitch of T2 syllables in the T3 Sandhi were both higher than in the new pattern, but the duration of T2 syllables in the T3 sandhi was shorter than in the new pattern. The highest pitch of T3 syllables in the T3 Sandhi was not significantly different from that in the new pattern. However, the lowest pitch of the T3 syllables in the T3 Sandhi was significantly lower than in the new pattern, and the duration of T3 syllables in the T3 Sandhi was

significantly longer than in the new pattern. These phonetic differences of the T2 and T3 syllables between the T3 Sandhi and the new pattern are indeed correlated to different degrees of phonetic naturalness.

As argued above, the phonetic naturalness is defined in terms of the two criteria: fewer pitch changes, and longer duration of T2 syllables and shorter duration of T3 syllables. The pitch changes of T2 and T3 syllables in the T3 Sandhi were significantly larger than in the new T3-T2 pattern. The boundary pitch changes between the two syllables in the T3 Sandhi were also significantly larger than in the new T3-T2 pattern. The results suggested that the T3 Sandhi should be less phonetically natural than the new T3-T2 pattern. As to the ratio between pitch changes and duration in T2 and T3, the disyllabic duration of the T3 Sandhi was not significant different from that of the new pattern. However, the pitch changes of both T2 and T3 syllables and the boundary pitch changes in the T3 Sandhi were significantly larger than in the new pattern. The results indicated that laryngeal movements in the T3-T2 sequence can coordinate with supralaryngeal gestures more smoothly than in the T3 Sandhi, so the new pattern takes fewer articulatory efforts to achieve pitch targets. In fact, more creaky voices and glottalization features were found in T3 syllables of the T3 Sandhi.

In addition, the duration of T2 syllables in the T3 Sandhi was shorter than in the new pattern. The shorter duration and the larger pitch changes of derived T2 in the T3 Sandhi takes more articulatory efforts than the longer duration and the smaller pitch changes of derived T2 in the new T3-T2 pattern. The shorter duration of rising tones tends to cause undershooting [9], so it takes more laryngeal manipulations to produce a distinct rising contour with shorter duration. As to the T3, its highest pitch in the T3 Sandhi was not significantly different from in the new pattern. However, the lowest pitch was significantly lower. The lower offset pitch in a falling contour tends to cause creaky voices glottalization features [1], which cause more articulatory efforts. Moreover, falling tones are relatively shorter [7], so the longer duration of T3 in the T3 Sandhi may exacerbate the degrees of glottalization features and creaky voices. These phonetic differences, such as larger pitch changes, longer duration of T3 syllables and shorter duration of T2 syllables, in the T3 Sandhi so not conform to the criteria of phonetic naturalness. Therefore, the T3-T2 pattern, a falling-rising pitch contour, is more phonetically natural than the T2-T3 sequence, a rising-falling pitch contour, derived from the T3 Sandhi.

6. Conclusion

The T3-T2 pattern, a falling-rising pitch contour, was found to be more phonetically natural than the T2-T3 sequence, a rising-falling pitch contour, derived from the T3 Sandhi based on the production results. The finding may explain a cross-linguistic tendency that a concave pitch contour, a falling-rising pitch sequence, is more common than a convex pitch contour, a rising-falling pitch sequence, roughly 352:80, among all Chinese languages [10]. The phonetic naturalness indeed plays a role. The finding can also explain why young Mandarin learners acquire the T3-T2 pattern earlier than the T3 Sandhi. Since the falling-rising pitch contour takes fewer articulatory efforts, children with an inexperienced articulatory control may have fewer difficulties in producing the T3-T2 pattern.

In fact, evolution of the new T3-T2 pattern can be determined by the immature laryngeal production of the Neutral Tone Sandhi among young speakers. Mandarin young speakers tend to pronounce the first tone at a middle pitch range, which is believed as a neutral position of laryngeal control [4], because of inexperienced laryngeal control. The middle onset pitch is characteristic of T2 and T3, so the first tone may be misperceived as T2 or T3, which may explain why the first tone of nominal reduplications has evolved into T3. A neutral tone following T2 and T3 is also produced at a middle pitch range [3]. The acoustic account explained why T2 and T3 bases of nominal reduplications are more susceptible to the new pattern than T4 and T1 bases. The nearly middle to middle pitch sequence is easy for young speakers to pronounce. However, final syllable lengthening was cross-linguistically found during child language acquisition [5], which raises a difficulty of producing the characteristic short duration of neutral syllables. Longer duration is found to be associated with a rising tone with everything being equal [11], so it can be a reason for young speakers' production of neutral tones to be misperceived as T2. These explanations suggested that the new T3-T2 pattern be evolved from a variation of Neutral Tone Sandhi produced by young speakers. The inference needs further investigation. To sum up, the T3-T2 pattern, a falling-rising pitch contour, is more phonetically natural than the T2-T3 sequence, a rising-falling pitch contour, derived from the T3 Sandhi.

7. References

[1] Bent, T., 2005. *Perception and production of non-native prosodic categories.* Ph.d. Dissertation. Northwestern University at Evanston, 138-140.

[2] Boersma, P.; Weenink, D., 2009. Praat: doing phonetics by computer (Version 5.1.02) [Computer Program]. <http://www.praat.org/>

[3] Chao, Y. R., 1968. *A Grammar of Spoken Chinese*. University of California Press: Berkeley.

[4] Chen, Y.; Xu, Y., 2006. Production of Weak Elements in Speech Evidence from F0 Patterns of Neutral Tone in Standard Chinese. *Phonetica* 63, 47-75.

[5] Nathani, S.; Oller, D. K.; Cobo-Lewis, A. B., 2003. Final Syllable Lengthening (FSL) in infant vocalizations. *Journal of Child Language* 30, 3-25.

[6] Peng, S., 1996. *Phonetic Implementation and Perception of Place Coarticulation and Tone Sandhi*. Ph.d. Dissertation. Ohio State University at Columbus, 17-24.

[7] Peng, S., 1997. Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics* 25, 371-400.

[8] Wang, C. Y., 2008. *The Acquisition of T3 Sandhi in Mandarin-speaking Children.* Unpublished manuscript. Michigan State University.

[9] Xu, Y., 2004. Understanding tone from the perspective of production and perception. *Language and Linguistics* 5, 757-797.

[10] Yip, M., 2002. *Tone*. Cambridge University Press, 29-30.

[11] Yu, A.. In press. Tonal effects on perceived vowel duration. To appear in *Laboratory Phonology* 10.

[12] Yu, A., 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24 (1), 187-214.