Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20). 2008. Volume 1.
Edited by Marjorie K.M. Chan and Hana Kang. Columbus, Ohio: The Ohio State University. Pages 87-107.

Tonetic Sound Change in Taiwan Mandarin: The Case of Tone 2 and Tone 3 Citation Contours[1]

Robert Sanders
University of Auckland

Recent acoustic studies of Taiwan Mandarin citation tones reveal disagreement about whether Tone 2 is dipping or rising and whether Tone 3 is falling or dipping. The present study surveys 33 informants who were subdivided by age, gender and language background. We can generalize that older groups and females are collectively more conservative than their respective counterparts, and likewise that monolinguals and females are collectively more stable than their respective counterparts. We speculate that the primary motivation for this tonetic change is reanalysis of fixed contour patterns by a new generation of speakers, redetermining which portion is meaningful and then making phonetic adjustments to render the newly meaningful part more salient and the remaining background noise less so. When a newly acquired toneme contour comes to resemble the existing contour of another toneme, the second toneme is forced to readjust in order to maintain perceptual distance. This case study offers a plausible model for tonetic sound change in any lexical tone language and helps explain why Chinese exhibits rich diversity in tonetic details across speech communities that are otherwise genetically and geographically very close.

1. Introduction

Over the last two decades acoustic studies examining the citation tones of Taiwan Mandarin (henceforth TM) have started to report pitch contours for single Tone 2 and/or Tone 3 syllables that differ from the prescriptive standard claimed for Putonghua and Beijing Mandarin (henceforth BM). In particular, most of these studies claim that Tone 2 now takes a dipping contour (e.g. Chiung 1999, Fon 1997, Li et al. 2006), though at least one study continues to claim that it is a rising contour (Shi & Deng 2006). Likewise, most studies claim that Tone 3 is now a low, falling tone (e.g. Chiung 1999, Li et al. 2006, Shi & Deng 2006). However, one scholar claims in multiple studies that it remains a dipping contour (Fon 1997, Fon & Chiang 1999, Fon et al. 2004), while a study of two decades

[1] The author would like to Professor Shi Feng of Nankai University, as well as his many graduate students, for their enthusiastic assistance in carrying out the actual acoustic analysis.

ago observed alternation between falling and dipping contours (Shih 1988). These conflicting empirical claims about the citation contours of TM Tone 2 and Tone 3 are summarized in Table 1 below:

Table 1: Conflicting Descriptions of TM Tone 2 & Tone 3 Citation Contours

Tone 2		**Tone 3**	
Study	**Claimed Contour**	**Study**	**Claimed Contour**
Shi & Deng 2006	Rising	Fon 1997, Fon & Chiang 1999, Fon et al. 2004	Dipping
Chiung 1999, Fon 1997, Li et al. 2006	Dipping	Chiung 1999, Li et al. 2006, Shi & Deng 2006	Low falling
		Shih 1988	Dipping/ Low Falling

The picture projected above is rather puzzling, not only because of the striking discrepancies in empirical claims despite the facts that each of these studies had access to the same speaker population, and that they were all carried out using the same acoustic measurements, but also because, with the exception of Shih (1988), not one of them reported even the slightest degree of variation in pronunciation, either within the pronunciation of a single speaker or across different speakers. The present study attempts to solve the mystery of why this is so.

With the exception of Chiung (1999), which makes use of a robust sampling size of 22 informants, one common characteristic of the above studies is that they are all based upon informant sampling sizes of no more than four individuals, and even the one exceptional study by Chiung was limited to sampling bilingual TM and Taiwan Southern Min (henceforth TSM) graduate students of roughly the same age studying at the University of Texas. Given what we know about sociolinguistic variability and given the striking divergence in findings reported above by different researchers, a decision was made to carry out our own survey using a larger number of informants and which took into consideration the language background, gender and age of those informants. This survey, then, was designed to answer the following questions:

1. What are the empirical acoustic facts for TM Tone 2 and Tone 3?
2. Is it possible for all of the claims presented in Table 1 above to hold some degree of truth (a nuanced truth), or must we discard some of these claims in favor of others (a more absolute truth)?

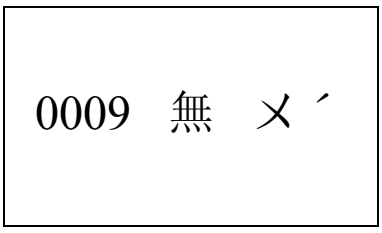
3. If we discover a nuanced truth, can correlations be drawn between various citation tone contour pronunciation and the sociolinguistic variables of age, gender or degree of monolinguism?
4. What might have motivated the observed synchronic state of flux in the first place? Is it more likely to have been external in its origin (i.e. the substratum influence of TSM) or internal?
5. What implications, if any, does the observed pattern of pronunciation of Tone 2 and Tone 3 in TM have for our general understanding of tonetic sound change?

2. Research Design and Procedure[2]

Altogether 33 informants were recruited to take part in this survey. All were born and minimally educated in Taiwan through the beginning of high school before migrating to New Zealand, although most arrived in New Zealand as adults. Of these 33 informants, 11 were monolingual speakers of TM while 22 were bilingual speakers of TM and TSM; 12 were male and 21 were female; 20 were aged between 50 and 64, 5 were aged between 30 and 49 and 8 were aged between 17 and 29.

Multiple examples of commonly-used characters representing each of the four tones of Mandarin Chinese were selected and then put onto MS Powerpoint slides, together with the desired target pronunciation of each character indicated in *Zhuyin Fuhao*. Altogether seven Tone 2 characters (無, 兒, 魚, 拔, 遲, 鵝, 移) and nine Tone 3 characters (五, 耳, 雨, 把, 齒, 惡, 椅, 耦, 此) were selected. An example of such a Powerpoint slide is shown in Figure 1 below:

Figure 1: Sample Character Presentation Slide



These slides were then presented to informants in a random order, one at a time. In order to insure that each character would be read as an individual syllable and not slurred together with a preceding or following syllable, advancement from one slide to the next only proceeded once the informant had finished pronouncing the slide in front of him/her and had paused.

[2] The content of Section 2 is taken verbatim and in total from the corresponding section of Sanders (2008).

Each set of recordings was then acoustically analyzed using Mini Speech Lab, a software program developed at Nankai University, which in addition to being able to display F0 pitch contours for individual syllables, is also able to calculate the average tone letters on a 1-5 pitch scale for any set of syllables sharing the same contour and then display the 'average' contour for the entire set of tokens of that tonal category against the 1-5 scale. Figure 2 below is an example of a set of seven different Tone 2 syllables of a single speaker, with hand-drawn lines tracking each syllable's dipping contour:

Figure 2: Pitch tracks of 7 Tone 2 tokens of subject A

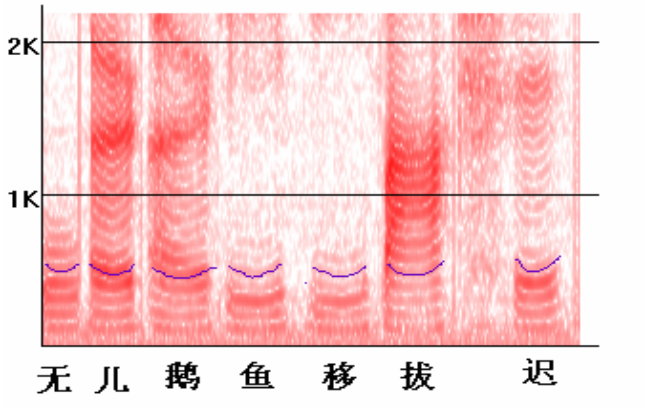
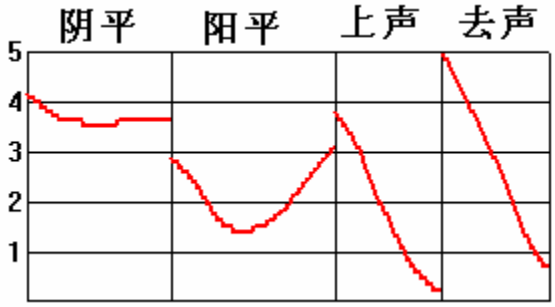


Figure 3, on the other hand, shows the 'average' contour for each tone token set for the speaker shown in the above figure, superimposed onto the 1-5 scale tone grid.

Figure 3: Average pitch contour for all four sets of subject A's tonal tokens superimposed on a 1-5 tone grid



We can see in Figure 3 that all seven tokens of Tone 2 yield an average dipping contour of 213, while the nine tokens of Tone 3 yield a falling contour of 31 or 41 (which contrasts with the average Tone 4 falling contour of 51).[3]

[3] Chiung (1999) claims a 53 contour for TM Tone 4, yielding a distinct low-high register contrast between a falling TM Tone 3 and a falling TM Tone 4. Many of the informants analyzed in this study who favored a falling contour for Tone 3 displayed a similar low-high register contrast with Tone 4.

2.1. Intra-Speaker Variation

It soon became evident that for certain informants, characters belonging to the same tonal category, e.g. Tone 3, were sometimes collectively pronounced with more than one pitch contour by the same speaker, i.e. sometimes the pronunciation had a dipping contour and sometimes a falling contour. An example of contour variation within the pronunciation of a single speaker is shown in Figures 4-5 below:

Figure 4: Pitch contour pronunciations of 7 Tone 2 syllables for subject B

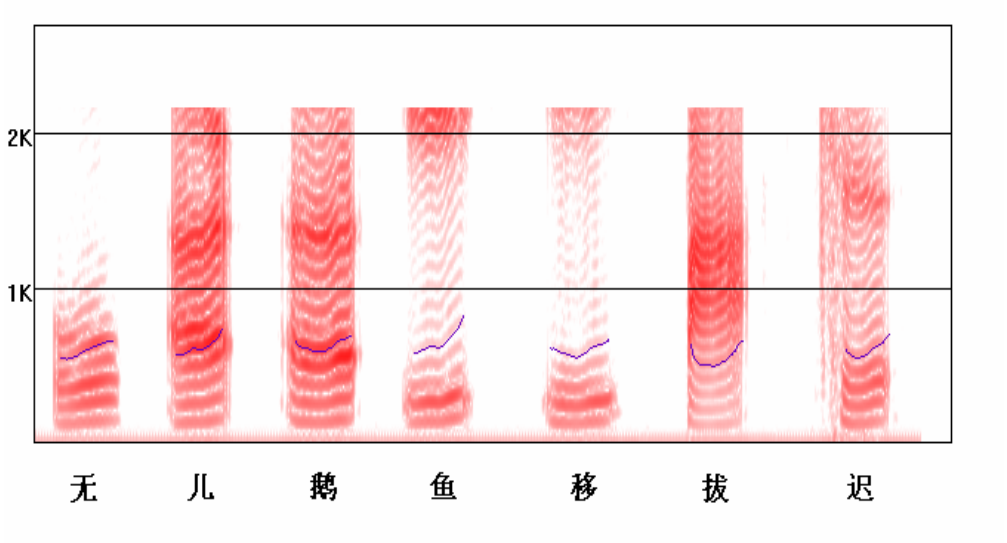
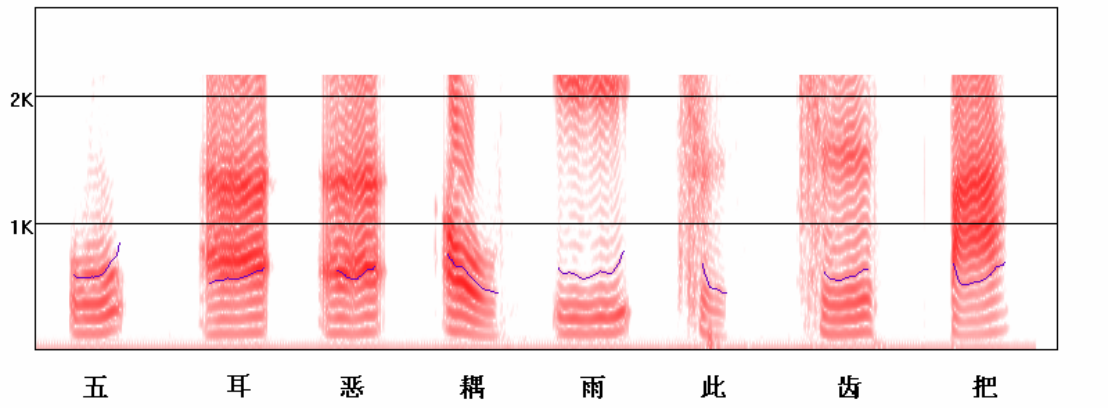


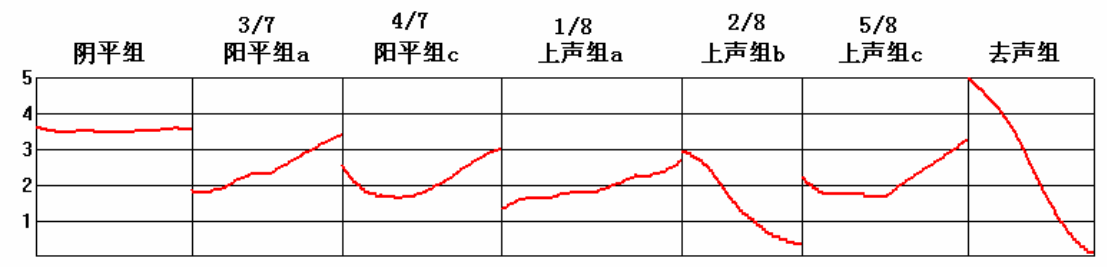
Figure 5: Pitch contour pronunciations of 8 Tone 3 syllables for subject B[4]



The internal variation within the pronunciation of this one speaker for both Tone 2 and Tone 3 is reflected in the 1-5 tone grid shown in Figure 6:

[4] The acoustic quality of the recording of the ninth Tone 3 character on the list, 椅, was too poor to allow for analysis by the software and therefore is not included in this person's data.

Figure 6: Average pitch contour for all four sets of Subject B's tonal tokens superimposed on a 1-5 tone grid (reflecting variation within an individual's pronunciation of both Tone 2 and Tone 3)



As one can see from inspecting Figures 4-6 above, for subject B, of the seven tokens of Tone 2, three were pronounced with a rising contour averaging 24, while four were pronounced with a dipping contour averaging 213. In the case of this informant's pronunciation of the eight Tone 3 tokens, one was pronounced with a rising 13 contour, two with a falling contour averaging 31, and five with a dipping contour averaging 113 or 114.

2.2. The Separate Issues of Stability and Conservatism

Comparing Figure 3 with Figure 6, it is apparent that two issues must be taken into consideration when evaluating any informant's production of a tonal category. On the one hand, we have to determine whether that informant consistently uses the same contour shape when pronouncing any isolated syllable belonging to the same tonal category, or whether instead there is a degree of variability in the contour shape. If the same contour shape is consistently used then we can call this speaker's production stable. On the other hand, if a speaker is observed to vary the tonal contour shape s/he uses when pronouncing isolated syllables belonging to the same tonal category, then we can characterize his/her production as being unstable. Clearly subject A, as shown in Figure 3, is much more stable in his production of Tone 2 and 3 contour shapes than is subject B, as shown in Figure 6.

However, a second revelation can be uncovered by carefully inspecting Figure 3—although a speaker may prove to be very consistent in his/her pronunciation of a particular tonal category, this fact alone is not indicative of whether or not the speaker's favored contour shape for that category is in fact prescriptively correct. What we see in Figure 3 is that the informant is very consistent in pronouncing Tone 2 with a dipping contour and Tone 3 with a falling contour, when in fact the prescriptively conservative contours for those two tonal categories are rising and dipping respectively. In other words, informant B's production can be characterized as being both stable and innovative in terms of the prescriptive standard, rather than stable and conservative.

The pronunciation of Tone 3 by subject B, as shown in Figure 6, provides yet another piece of evidence in support of the need to treat stability and conservatism as two separate and distinct parameters. In this case we see that of the eight Tone 3 tokens, one is pronounced with a rising contour, three with a falling contour and four with the conservative dipping contour. In other words, although this informant's pronunciation of this tone is very unstable, he marginally favors the conservative dipping contour over either of the other two innovative alternatives. This, then, contrasts with his unstable pronunciation of Tone 2, where he slightly favors the innovative dipping contour over the conservative rising contour. B's pronunciation of Tone 2 is therefore unstable and innovative, while his pronunciation of Tone 3 is unstable and conservative.

2.3. Monolinguals vs. Bilinguals

TSM is frequently cited as a major factor in the development of TM (Kubler 1985, Huang 1993). Therefore informants were divided into two groups—those who spoke only TM and those who were bilingual in both TM and TSM. Designation of membership into the appropriate linguistic group was determined on the basis of a brief written survey asking informants to identify the frequency s/he used TM and TSM in different formal and informal domains. Those who reported frequent use of TSM in domains outside the family were classified as bilinguals; those whose use of TSM was restricted to 'kitchen Chinese' inside the home, as well as those who did not speak any TSM at all, were classified as monolingual.

2.4. Age Grouping

The age range of all 33 informants was 17-64. In an effort both to reflect rough generational units of twenty years and to take into account differences in language use at the time ethnic Southern Min Taiwanese were acquiring TM, these 33 informants were divided into three age groups—50 through 64 (henceforth 'old'), 30 through 49 (henceforth 'middle-aged') and 17 through 29 (henceforth 'young'). Generally speaking, 'old' ethnic Southern Min Taiwanese would have as young children spoken TSM at home with siblings, parents and grandparents and would not have been extensively exposed to TM prior to starting kindergarten or primary school. On the other hand, 'middle-aged' ethnic Southern Min Taiwanese, the offspring of that 'old' group, had parents and siblings with whom they could speak TM at home if they so wished, while TSM was still the more practical language for communicating with grandparents. Young ethnic Southern Min Taiwanese, therefore, represent the first generation where active command of TM may have been sufficient alone for speaking with all generations of one's own family.

3. Results[5]

3.1. Raw Data

The simple picture of the citation contour pronunciation of each Tone 2 and Tone 3 character for all 33 informants is shown in Tables 2 and 3 below.

Table 2: Total Tone 2 contour distribution by individual character

Contour	無	兒	魚	拔	遲	鵝	移
rising	23	19	26	20	21	18	20
dipping	10	14	7	13	12	15	13
TOTAL	33	33	33	33	33	33	33

Table 3: Total Tone 3 contour distribution by individual character

Contour	五	耳	雨	把	齒	惡	椅	耦	此
dipping	24	23	20	22	20	24	20	22	18
falling	8	10	12	9	13	9	9	11	15
rising	1	0	1	1	0	0	3	0	0
TOTAL	33	33	33	32	33	33	32	33	33

Without taking into consideration the sociolinguistic variables of age, gender or linguistic background, the picture directly above looks very unstable, with variable pronunciation the norm across the board, though to slightly different extents depending on the particular character in question. In all cases, however, we note that the prescriptive/conservative tone contour (rising contour in the case of Tone 2, dipping contour in the case of Tone 3) is favored.[6]

Keeping in mind the picture revealed in Figures 3 and 6 above, that the pronunciation of a particular individual can be characterized as being stable or unstable, conservative or innovative, and that these two sets of binary label choices are independent of one another, let us now examine the raw data in Tables 2 and 3 above with reference to each of our three sociolinguistic variables of age, gender and language background to deter-

[5] The content of Section 3 is taken verbatim and in total from the corresponding section of Sanders (2008).

[6] This fact may be the effect of sampling bias, in that old informants, who constitute the clear majority of the total number of informants surveyed, happen to be the most conservative group as far as tone contour pronunciation is concerned.

mine whether any patterns can be teased out. For the purpose of this analysis, the pronunciation of an individual speaker is defined as being 'stable' if after examining the contour shape of every character belonging to a common tonal category and deciding which contour shape has been used the majority of the time by that speaker, no more than three exceptional tokens can be found. Four or more exceptions to the observed majority preference indicate that that speaker's pronunciation is unstable. A speaker whose observed preferred contour shape matches the prescribed contour shape for that tonal category is said to be conservative, while a speaker whose observed preference for a particular contour shape differs from the prescribed contour shape is said to be innovative.

3.2. Data by Age Group

Table 4: Relative group stability and collective conservatism by age group

Age 50-64 (20 informants)		**Tone 2**	**Tone 3**
Stability	Stable	16	15
	Unstable	4	5
Conservatism	Rising	**15***	-
	Falling	-	3
	Dipping	5	**17***
Age 30-49 (5 informants)			
Stability	Stable	4	4
	Unstable	1	1
Conservatism	Rising	**3***	-
	Falling	-	1
	Dipping	2	**4***
Age 17-29 (8 informants)			
Stability	Stable	6	7
	Unstable	2	1
Conservatism	Rising	**3***	-
	Falling	-	4
	Dipping	5	**4***

*Prescriptive contour shape

The picture revealed by Table 4 is striking in that it provides strong counter-evidence to the seemingly chaotic and unstable synchronic snapshot suggested by the raw data in Tables 2 and 3. In fact, there is strong evidence in Table 4 that not only are Tones 2 and 3 in TM changing respectively from a rising contour to a dipping one and from a dipping contour to a falling contour, but that these changes are accelerating over time. Furthermore, at any point along the change timeline the degree of collective, group instability remains about constant.

As shown in Table 4, in terms of conservative versus innovative tone contour pronunciation, the ratios of 15:5 and 16:3 for the old group, 3:2 and 4:1 for the middle-aged group and 3:5 and 4:4 for the young group, tell us that the older the age group is the greater its collective degree of conservatism, and hence, the younger the age group is the greater its collective innovation. This being the case, we have apparently captured evidence of tone contour change that has been accelerating over time.

At the same time that these changes have been accelerating, there is no evidence to suggest that group pronunciation itself has become more destabilized as a result, given the rather constant stable-to-unstable ratios of 16:4 (4:1) and 15:5 (3:1) for the old group, 3:2 and 4:1 for the middle-aged group and 6:2 (3:1) and 7:1 for the young group.

3. 3. Data by Gender

Table 5: Relative group stability and collective conservatism by gender

Male (12 informants)		**Tone 2**	**Tone 3**
Stability	Stable	8	8
	Unstable	4	4
Conservatism	Rising	**6***	-
	Falling	-	4
	Dipping	6	**8***
Female (21 informants)			
Stability	Stable	18	17
	Unstable	3	4
Conservatism	Rising	**15***	-
	Falling	-	5
	Dipping	6	**16***

*Prescriptive contour shape

There is nothing inherently surprising in what is shown in Table 5. Females are seen to have more stable pronunciation of both tonal categories than do males, with the former displaying a stable pronunciation of Tone 2 at a rate of 18:3 (6:1) versus a rate of 8:4 (2:1) for males, and females preferring a stable pronunciation of Tone 3 at a rate of 17:4 (~4:1) versus a 8:4 (2:1) rate for males. Additionally, females are seen to favor the prescriptive contour of each tone at much higher rates than do males, favoring it at a ratio of 15:6 (5:2) for Tone 2 and 16:5 (~3:1) for Tone 3, while males are seen to favor the prescriptive contour of Tone 2 and Tone 3 at a ratio of 6:6 (1:1) and 8:4 (2:1) respectively. Females have more stable pronunciation and are more conservative than their male counterparts.

3.4. Data by Language Background

Table 6: Relative group stability and collective conservatism by language background

Monolingual (11 informants)		**Tone 2**	**Tone 3**
Stability	Stable	10	11
	Unstable	1	0
Conservatism	Rising	**6***	-
	Falling	-	5
	Dipping	5	**6***
Bilingual (22 informants)			
Stability	Stable	16	14
	Unstable	6	8
Conservatism	Rising	**15***	-
	Falling	-	3
	Dipping	7	**19***

*Prescriptive contour shape

Table 6 reveals one expected result and one possible surprise. As expected, monolinguals are seen to display a significantly more stable pronunciation of the two contours than do the bilinguals. However, when it comes to which group is more conservative in its pronunciation of the two tones, it is the bilinguals, with ratios of 15:7 (~2:1) for Tone 2 and 19:3 (~6:1), rather than the monolinguals, with ratios of 6:5 (~1:1) for both Tone 2 and Tone 3, that show the higher rate of conservatism. This observation could, however, be the effect of the much larger sample size of old informants in

comparison to middle-aged and young informants, and these old informants tended to mostly be bilingual. Because the current sample size is not large enough to tease apart bilingualism and age, we are not in a position to argue convincingly that bilinguals are more conservative than monolinguals in terms of their tone contour pronunciation. Nevertheless, because of the striking discrepancy in the conservatism ratios between the two groups, it is something worth noting for future investigation.

4. Discussion

4.1. Possible Motivation for Tonetic Change

While scholars like Kubler (1985) and Huang (1993) have noted the substantial substratum influence of TSM on the development of TM, there is insufficient objective evidence to conclude that the changes in tone contour observed here are in fact primarily the products of this type of influence here. This is supported in Table 7 below:

Table 7: Comparison of Citation Contours of Etymological Tone 2 & Tone 3 in TM and TSM

Tonal Category	**Old TM**	**New TM**	**TSM[7] (Taipei)**	**TSM (Kaohsiung)**
Yangping (TM Tone 2)	Rising (35)	Dipping (325)[8]	Rising (13)	Rising (24)
Shang Sheng (TM Tone 3)	Dipping (213)	Falling (21)	Falling (53)	Falling (53)

4.1.1. Yangping

Given that the Yangping tonal category is pronounced with a rising contour in both prescriptive Old TM and TSM, then from a logical point of view, if TSM had actually held sway over the development of New TM Tone 2 then TM Tone 2 should never have changed its basic contour shape in the first place. After all, a rising contour in TSM should have served to reinforce a rising contour in New TM, not push it to take on a dipping contour instead. Obviously another explanation exists for why TM Tone 2 is changing into a dipping contour in citation pronunciation. The general mechanism to be suggested here is one that can account not only for how a so-called rising tone changed into a dipping contour, but also for how any tone shape might change into any other tone shape. This proposed process takes the form of reanalysis, whereby the original tonal contour, a continuous sound signal consisting of one linguistically meaningful segment set adjacent to one or two linguistically meaningless background noise segments, is then resegmented for linguistic meaning and background noise by a new generation of

[7] All TSM tonal descriptions come from Shi & Deng (2006)

[8] Li et al. (2006)

speakers[9]. In such cases, what was once partially or fully meaningless background noise may come to carry the linguistic load of the signal while the segment of that sound signal that originally carried all of the linguistic weight might come to be reinterpreted as mere background noise, thus becoming surplus to requirement. Subsequent adjustments can then be made to make the new linguistically meaningful part of the sound signal more salient and the newly insignificant portion(s) of the original sound signal less so. The details of how this process might have taken place in TM both for Tone 2 and Tone 3 will be addressed in more detail in section 4.2 below.

4.1.2. Shang Sheng

On the surface it can easily be argued that in the case of Shang Sheng (TM Tone 3), TSM, with a falling contour in citation form, must have played a significant role in turning the old dipping contour of TM Tone 3 into a falling contour. While we do not want to completely discount any possible substratum influence on the development of TM Tone 3 into a falling citation contour, there nevertheless exist other internal factors that just as well may have served to bring about this same change. First, Tone 3, unlike the other three tonal categories of Mandarin, has one contour specifically limited to citation pronunciation and another contour limited to use when it occurs in connected speech. Whereas the Mandarin Tone 3 citation allotone is dipping, its connected speech allotone is in fact falling. Additionally, in real life the connected speech allotone for Mandarin Tone 3 occurs in natural speech much more frequently than does the citation allotone. From a language acquisition perspective, then, each token of a particular input presented to a child acquiring his native language is a vote influencing the child's eventual choice of which allotone to consider as basic. The significantly greater number of 'votes' calling for a falling contour as opposed to a dipping contour could very well have helped to influence the next generation of TM speakers to adopt the falling contour as the basic contour.

There is also systematic pressure for discarding the existence of a special citation contour for Tone 3. After all, allotonic variation between citation pronunciation and connected speech pronunciation only exists for Tone 3 and does not exist for any of the other three tonal categories of the language. For reasons of systematic regularity alone, there seems to exist some degree of motivation to 'over-generalize' or 'simplify' the system, yielding the straightforward, exceptionless rule that 'speech contour = citation contour'.

Lastly, there is yet another potential mechanism in place, already touched upon in the section immediately above, which might also help facilitate a dipping contour to evolve into a falling contour, i.e. the possibility for a new generation of speakers to reanalyze the linguistically salient and linguistically meaningless components of a tone

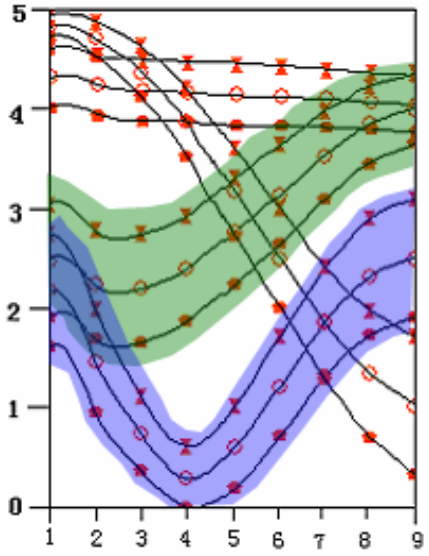
[9] The breaking down of a continuous signal into linguistically meaningful and linguistically meaningless segments is analogous to how cursive writing in Chinese is processed by native readers. While non-natives may have no idea which parts of the squiggly line embody strokes of the character and which parts are merely transitional movement between strokes, a native speaker can usually work out this information very easily.

category's specific contour shape differently from previous generations. The details of how this reanalysis may have taken place for Tone 2 and Tone 3 are discussed directly below.

4.2. How Resegmentation of the Continuous F0 Speech Signal Might Work

Figure 7 below shows the overall pitch range and average citation pitch contours of Tones 1-4 for 52 Beijing Mandarin speakers as reported in Wang (2006).

Figure 7: Distribution of the Four Citation Contours in Beiiing Mandarin (Wang 2006)



Of particular note here are the contour shapes of Tone 2 (shaded in green) and Tone 3 (shaded in blue). A careful inspection of the entire Tone 2 pitch track includes not only the high rise from level 3 to level 5 over the last six milliseconds of the syllable pronunciation, but also a slight downward descent over the two milliseconds preceding the starting point of this prominent rise. Putting the two segments together forms a physical dip at the beginning of the signal. Wang (personal communication) notes that this initial dip is consistently present in the pronunciation of all 52 speakers she studied. In terms of linguistic significance in BM, however, the first two seconds of downward descent is merely background noise; it is the six milliseconds of sharp rise that follow this segment that embody the linguistically meaningful portion of this eight millisecond sound signal.

From the perspective of a child learner of this input, however, there is no a priori reason to assume that this particular pitch track shape must always be interpreted in exactly this way. It is really up to the learner to determine which part of this continuous physical signal conveys meaning and which part of it does not. Once the decision is made by the learner which part of this pitch track is meaningful and which part is not, then s/he is free to do whatever is necessary to make the meaningful segment more prominent and the meaningless portion less so. Should a new generation of speakers decide that the

meaningful portion of this pitch contour includes the beginning portion of the signal, which when combined with the beginning part of the rise forms an initial dip, then this reanalysis could ultimately end up focussing instead on the initial dip as the perceived linguistically meaningful segment rather than the rise. In order to make this dip more salient over the full eight milliseconds of required sound signal, both the height and duration of the initial drop could then be lengthened and the height and duration of the final rise could simultaneously be shortened. Such a move would thus produce a dipping contour shape similar to what we see for Tone 3, though not necessarily of exactly the same pitch range.

In the case of BM Tone 3 above, the entire duration of the dipping contour is linguistically meaningful. However, once again there is no a priori reason to assume that this particular pitch track shape must always be interpreted in exactly this same way, and is therefore potentially a candidate for any type of reanalysis by a new generation of speakers. It is noted in Section 4.1.2 above that there exist system-internal reasons why a citation contour might be reinterpreted as being falling rather than dipping—in connected speech it takes a falling contour and the other three tones of Mandarin have citation contours that match their connected speech contours. It is also noted in that section that there exists one system-external factor, the falling citation contour of Shang Sheng in TSM, which could likewise serve as a substratum motivation for a new generation of TM learners to reinterpret the citation contour of Tone 3 as falling. Certainly the combination of these internal and external factors, coupled with the ease in which a new generation can potentially reanalyze a given pitch contour for linguistic meaning and background noise segmentation, and then modify that contour shape to make the linguistically meaningful segment more prominent and the linguistically meaningless portion less so. In the end it is possible for us to identify in the development of the TM Tone 3 contour shape several potential seeds of change and one very easy to manipulate mechanism to facilitate this change. The question can now be asked whether the changes observed in TM Tones 2 and 3 likely occurred independently of one another or whether it is more likely that the change in one category led to a change in the other.

4.3. Possible Sequencing of This Chain of Events

We believe that it is useful to view a lexical tone system, like the type found in Chinese, in much the same way as we view a vowel system. Each contains a fixed set of categories and, as Saussure (1983) notes in his discussion of phonemes, each individual member/category of that set is defined more in terms of its functional opposition to all other categories belonging to that same set than it does to each category's own individual phonetic particulars. Because of this, the status of a phoneme or toneme does not change even if its respective phonetic details do, just as long as the opposition it experiences with all other phonemic or tonemic categories in that set is maintained. The only time that problems arise in such a system is when the phonetic change to one member of the set is such that it ends up becoming homophonous or at least comes to approximate the pronunciation of another member of that same set. In this case one of two things will happen—

either two previously functionally distinct categories will merge as one or the pronunciation of the newly encroached upon member of the set will be forced to change as well in order to maintain the existing salient oppositional distance it has thus far experienced with the recently encroaching category. Readjustments in vowel pronunciation in order to maintain or optimize the perceived distance of categorical oppositions take one of two forms, one called a push chain and one called a pull chain.

Based on the distribution of conservative versus innovative pronunciation of Tone 2 and Tone 3 across the three different age groups contained in Table 4 and presented again below, we wish to speculate that the relative sequence of the two tonetic changes in TM reported in the present study possibly unfolded similarly to the manner that a push chain unfolds.

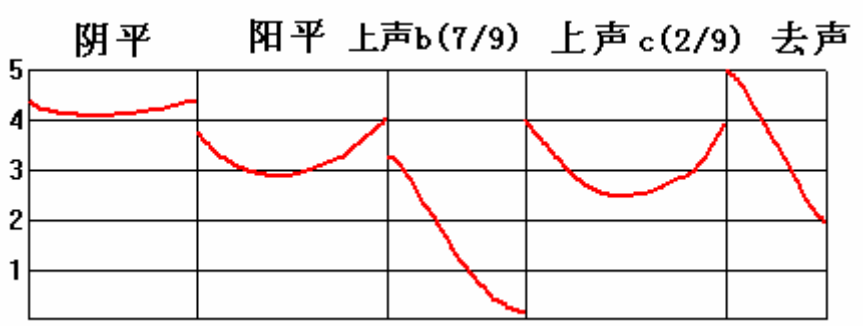
Table 4: Relative group stability and collective conservatism by age group

Age 50-64 (20 informants)		**Tone 2**	**Tone 3**
Stability	Stable	16	15
	Unstable	4	5
Conservatism	Rising	**15***	-
	Falling	-	3
	Dipping	5	**17***
Age 30-49 (5 informants)			
Stability	Stable	4	4
	Unstable	1	1
Conservatism	Rising	**3***	-
	Falling	-	1
	Dipping	2	**4***
Age 17-29 (8 informants)			
Stability	Stable	6	7
	Unstable	2	1
Conservatism	Rising	**3***	-
	Falling	-	4
	Dipping	5	**4***

*Prescriptive contour shape

In this table as we move from the old group to the middle-age group to the young group we see that in the case of both Tone 2 and Tone 3 that not only does innovation increase over apparent time, but also that for both the old and middle-aged groups, the change in Tone 2 is more common than the change in Tone 3. That we see more change in Tone 2 than in Tone 3 strongly suggests that Tone 2 began to change before Tone 3 did. If this is so then the first tonetic change created two dipping contours in the system. In their perceptual study of TM Tone 2 and Tone 3 discrimination based on the assumption that both of these tonemes in citation form are indeed pronounced with a dipping contour, Fon et al. (2004) note that the only useful acoustic cues remaining that allow native speakers to distinguish these two particular dipping contours are the degree of slope in the fall and degree of subsequent rise that ends the signal, because the other normal features used to distinguish one tone shape from another, pitch shape and pitch range, have been neutralized. as a result, Tone 2 is perceived to be encroaching upon Tone 3. In response to this encroachment, and supported by the internal motivation discussed in Sections 4.1.2 and 4.2 above, Tone 3 then likely began to evolve into a falling contour. This new falling citation contour for Tone 3 then created another problem in the system, as Tone 4 was already itself pronounced with a falling contour. However, it would appear that TM speakers are adjusting to this encroachment upon Tone 4 by Tone 3 by treating these two falling contours as belonging to two different registers (occupying different pitch ranges), which is exemplified for a single speaker in Figure 8 below:

Figure 8: Subject C's Four Citation Tones



In the above figure we see that Tone 2 for this particular speaker is consistently dipping and that Tone 3 is largely falling (7 out of 9 tokens). Additionally, we see that the falling contour of Tone 3 has a pitch range of 31, whereas the pitch range of Tone 4 is no longer the prescriptive 51, put has been shortened to merely cover the pitch range 53. By contracting the pitch range of Tone 4 and positioning it in a pitch sector comfortably above that of the normal pitch range occupied by the Tone 3 falling contour, a high-low register distinction has been created which can help to maintain the perceptual distance between Tone 3 and Tone 4.

4.4. What Changes in TM Tone Contours Might Tells Us about Tonetic Sound Change in Chinese in General

From both a diachronic and synchronic perspective, tonal categories (tonemes) seem to be remarkably stable, with regular correspondences easily established over time and distance between any one variety of Chinese and another. At the same time, however, the phonetic details of individual tonal categories vary hugely even over very short geographical distances, strongly suggesting that tonal contours are themselves inherently unstable and subject to frequent change in shape. The present study has managed to capture a series of such tonetic changes within TM that likely unfolded over a short period of just three to four decades. However, despite these tonetic changes, the tonemic system itself has been left intact, with the original four tonemic categories preserved. In the scenario proposed here, tonetic change initially came about as a result of reanalysis of a continuous sound signal by a new generation of speakers, who in the process of acquiring their native language reassigned linguistic meaning and background noise status to different segments of set tonal contour patterns they were hearing. In the process of identifying a new linguistic locus and new areas of background noise within that continuous sound signal, conditions were created allowing for easy manipulation of that pitch track to make the linguistically relevant portion more prominent and lessening the prominence of the background noise. Original minor features of the pitch wave, such as a small dip, could then be accentuated while linguistically less meaningful portions, such as a rise in pitch, could then be shortened or eliminated altogether. With this reanalysis and phonetic readjustment a new contour was then created for Tone 2 that began to encroach acoustically upon Tone 3, which under pressure from Tone 2 and with other factors as well pushing it to take on another contour shape, it changed into a falling contour. This new falling contour then began to encroach upon Tone 4, which reacted by shortening its full pitch range, thus maintaining a clear distance from the new Tone 3 pitch contour. This sequence of events for tonetic change closely replicates the unfolding of the commonly observed phenomenon known as a vowel push chain.

We wish to suggest that what has been uncovered here provides a window not only into what transpired over time in TM, but more importantly, it offers one very plausible mechanism by which lexical tone languages in general, including Chinese, evolve over time. It is not being suggested that this is the only route by which tonetic change takes place. However, it does represent a very plausible explanation for why we observe such a rich diversity of tonetic details across speech communities that are otherwise so genetically and geographically close, and hence, supports the impression that globally speaking, that tonetic details are relatively transitory while tonemic categories remain significantly more stable.

5. Conclusion

Returning now to the original research questions, we see now that the general trend over time in TM is for the Tone 2 citation contour to be dipping and for the Tone 3

citation to be falling. However, this remains a change in progress, meaning that it is still possible to observe variation between the original prescriptive contour shape and its innovative counterpart, both across speakers and within the speech of a single individual. This variation, however, is not completely random, as it can, to a certain degree, be correlated with specific sociolinguistic variables. Specifically, males are more likely than females and bilinguals are much more likely than monolinguals to display individual variation in contour shape pronunciation. At the same time, however, age does not appear to be a factor in determining the likelihood that variation within the pronunciation of a given individual will occur, although it is clearly the case that the older a speaker is the greater the likelihood that s/he will be employ a conservative contour shape for both Tone 2 and Tone 3 when it is pronounced in isolation.

A second key variable when considering the likelihood that a given speaker of TM will favor either a conservative or innovative contour is the speaker's gender, with females displaying a much stronger tendency than males to employ a conservative contour shape. It may also be the case that language background exemplifies a third important variable in the conservative/innovative opposition, although the sample size of this study does not allow us to tease apart language background from speaker age and thus discover a definitive answer to this question.

We speculate that the original motivation for the tonetic changes observed here was reanalyis of the citation contour of Tone 2 by a new generation of TM speakers. These new speakers took what was once a linguistically meaningless portion of the original Tone 2 citation pitch track, a slight downward slope at the very beginning of the pronunciation, and reanalyzed that downward slope, together with the rise that followed it, as being linguistically meaningful. Subsequent adjustments were then made to make the new linguistically meaningful dipping contour more salient and the less significant rise at the end of the original sound signal less salient. The result of this reanalysis and readjustment was a dipping contour that closely resembled the Tone 3 citation contour both in terms of contour shape and pitch range. Already under a certain amount of pressure from its more frequently used connected speech allotone and from TSM to be pronounced with a falling contour anyway, Tone 3 adjusted to the encroachment of the new Tone 2 contour and become falling. However, this new falling tone was in danger itself of encroaching upon the existing Tone 4 falling contour, so an adjustment was made to Tone 4 whereby it changed its pitch range of 51 by ceding the 31 range to Tone 3 and keeping the 53 range for itself.

The mechanism of pitch track reanalysis by a new generation of speakers, with subsequent phonetic readjustment to make the linguistically meaningful portion of the continuous sound signal more salient and rendering the other meaningless portion(s) of that same signal less salient provides both a credible motivation and a simple mechanism for realizing tonetic change. In the study presented here, change within the tonemic system proceeded in exactly the same manner as a vowel push chain, with tonetic change to one member of the toneme set initiating a chain of events that led to tonetic change in

other members further down the line. With Tone 2 changing from a rising contour to a dipping one it then began to encroach upon the pitch contour of Tone 3, thus motivating tone 3 to likewise change its contour shape in order to maintain its perceptual distance from Tone 1. In this particular case, there already existed additional internal and external pressures motivating Tone 3 to change into a falling contour. However, in so doing, this new contour came to encroach upon the falling contour of Tone 4, and in an effort to maintain their perceptual distance the pitch range was divided into two registers, one high, one low, with Tone 4 occupying the upper register and Tone 3 occupying the lower one. The scenario proposed here not only provides a window into what likely transpired in the tonetic evolution of TM, but more importantly, it also offers a workable model for how tonetic change might unfold in any lexical tone language. Such a scenario, in turn, provides a very simple and plausible explanation for why we observe in languages like Chinese such a rich diversity of tonetic details across speech communities that are otherwise so genetically and geographically close.

REFERENCES

CHIUNG, WI-VUN TAIFFALO. 1999. The tonal comparisons and contrasts between Taiwanese and Taiwan Mandarin. Presentation at the Sixth Annual UTA Student Conference in Linguistics.

FON, JANICE. 1997. *What are tones really like: An acoustic-based study Taiwan Mandarin tones*. Thesis. National Taiwan University.

FON, JANICE AND WEN-YU CHIANG. 1999. What does Chao have to say about tones? A case study of Taiwan Mandarin. *Journal of Chinese Linguistics* 27.1:15-37.

FON, JANICE, WEN-YU CHIANG and HINTAT CHEUNG. 2004. Production and perception of two dipping tones (Tone 2 and Tone 3) in Taiwan Mandarin. *Journal of Chinese Linguistics* 32.2:249-281.

HUANG, HSUAN-FAN. 1993. 語言，社會與族群意識. *Yǔyán, Shèhuì yǔ Qúnzú Yìshì*. (*Language, Society and Group Identity*). Taipei: Crane Publishing Co.

KUBLER, CORNELIUS. 1985. *The development of Mandarin in Taiwan: A case study of language contact*. Taipei: Student Book Co. Ltd.

LI, AIJUN, ZIYU XIONG and XIA WANG. 2006. Contrastive study on tonal patterns between accented and standard Chinese. *Proceedings of the 5th International Symposium on Chinese Spoken Language 2006*.

SANDERS, ROBERT M. 2008. Citation tone change in Taiwan Mandarin: Should teachers remain tone deaf? To appear in the *Journal of the Chinese Language Teachers Association* 43.2.

SAUSSURE, FERDINAND DE. 1983. *Course in general linguistics.* Eds. Charles Bally and Albert Sechehaye. Trans. Roy Harris. La Salle, Illinois: Open Court.

SHI, FENG and DAN DENG. 2006. 普通話與臺灣國語的語音對比. Pǔtōnghuà yǔ Táiwān Guóyǔ de yǔyīn duìbǐ (A phonetic contrast of Mainland Mandarin and Taiwan Mandarin), in Dah-an Ho et al. ed. *Linguistic studies in Chinese and neighboring languages: Festschrift in honor of Professor Pang-hsin Ting on his 70th birthday.* Academia Sinica Language and Linguistics Monograph Series Number W-6, 1:371-393.

SHIH, CHI-LIN. 1988. Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory* 3:83-109.

WANG PING. 2006. 北京话单音字调和元音的统计性研究. *Běijīnghuà dānyīn zìdiào hé yuányīn de tǒngjìxìng yánjiū (A statistical study of the tones and vowels of Beijing Mandarin monosyllables).* Dissertation. Nankai University.