The Phonetic and Phonological Features of Tone 3 in Taiwan Mandarin

Robert Sanders
University of Auckland

Abstract
This study examines the F0 pronunciation of 4 speakers of Taiwan Mandarin reading approximately 315 tone 3 characters, presented one-by-one, in random order, embedded in a total list of 1,275 characters. The three bilingual Mandarin/Southern Min subjects showed a strong tendency to pronounce tone 3 with a sharper, falling contour, while the fourth speaker, possessing only minimal proficiency in Southern Min, strongly preferred a dipping contour instead. Like Chiang (1999), we hypothesize that the emerging 31 tone 3 citation contour from the 214 citation contour in Taiwan Mandarin reflects, in part at least, the substratum influence of Taiwan Southern Min, whose corresponding tonal category is pronounced in citation form as 51. Additionally, however, considerable phonetic variation was observed within the token pool of each speaker, yielding occasional examples of level and rising contours as well. We hypothesize that from the speaker’s point of view this variation is not psychologically real, although some of it, when it begins to physically approximate a tone 2 contour, may be perceptually real for children in the acquisition process. And that the contour of Taiwan Mandarin tone 3 might rise enough to approximate the rising contour of tone 2 likely reflects the influence of tone 3 sandhi both in Mandarin and Taiwan Southern Min. Finally, observing sound change in progress of the Taiwan Mandarin tone 3 citation contour from 214 to 31, we find strong support for the basic assumption of the theory of lexical diffusion proposed by Wang (1969), that throughout most of the process of sound change, variation among lexical members of the same qualifying category is both random and robust, and that it is neither uniform nor exceptionless. From the point of view of what a speaker thinks s/he is saying, however, it may not always be abrupt.

1. Introduction
As a student of Chinese for over half the history of the Mandarin Training Center and as a teacher of that language for roughly two decades, I have been grappling with two themes about Chinese tones for most of my adult life, one of a somewhat theoretical nature, the other a very practical issue that up until recently never appeared to be especially relevant to my more theoretical question. In the case of that quasi-theoretical question, I have observed that when native speakers of a variety of Chinese other than prescriptive, broadcast Mandarin begin to speak their best Mandarin, it is often the pronunciation of their tonal contours, rather than the pronunciation of their segmental features, that betrays their status as non-native speakers of the prescriptive language. Given this extremely high frequency of negative transfer of tonal contours from native varieties of Chinese into non-native varieties of Chinese, one is compelled to conclude that the basic phonetic details of the tonal contours of one’s native variety of Chinese are almost psychologically set in stone, and thus remarkably impervious to external linguistic pressure even when transferred into a different variety of Chinese, cf. Huang (2004) for a study that demonstrates that when native speakers of one variety/dialect of Chinese listen to another variety/dialect of Chinese, their perception of the non-native tonal contours and patterns is indeed heavily influenced by their native tonal system.

But if the phonetic details of tonal categories are psychologically and articulatorily so stubborn, why then is it so rare to find any two subdialects of a given dialect that happen to share the same phonetic values for each of their shared tonal categories, e.g. despite their extreme geographical and linguistic closeness, why do Beijing and Tianjin Mandarin display such salient phonetic differences in tone? After all, because related dialects and subdialects are historically derived from a single, common ancestor that possessed an inherently stable system of tones, we would naturally expect to find more phonetic/tonetic commonality among closely-related tonal systems. And similarly, why is it also so common to discover in any given Chinese dialect examples of individual morphemes that have shifted in a seemingly random and unpredictable fashion from their expected historical tonal reflex categories into completely different tonal categories instead? How could any of this have happened in the first place, and what might the motivation(s) and mechanism(s)

---

1 This research is partially supported by the Grant-in-Aid for Scientific Research for the Japan Society for the Promotion of Science (No. 15401011). The author would like to gratefully acknowledge Professor Lien Chinfa (連金發) and his graduate students at Tsinghua University for helping to locate and record subjects, Professor Uehara Satoshi (上原聡) and his graduate students at Tohoku University for their assistance in converting the digital recordings into individual sound wave files, as well as Professor Shi Feng (石锋) and his graduate student Ran Qibin (冉啓斌) of Nankai University for their technical support with the acoustical analysis. Any and all errors are the fault of the author.
have been for such counter-intuitive developments to take place at all? And how does all of this relate to the competing views of sound change outlined by Wang (1969), that either sound change takes place so gradually and always to the same degree that it is imperceptible, affecting every member of the qualifying lexicon simultaneously without exception, or that it takes place abruptly, affecting potential candidates just one or a few at a time, and that it involves a period of intermediary variation, and is thus subject to exceptional results?

My ongoing practical question about tone, on the other hand, concerns how Mandarin tone 3 is traditionally presented to students. Specifically, despite the fact that native speakers only ever prescriptively pronounce tone 3 with a long, dipping contour under very limited circumstances, and hence, very infrequently, why is it that we almost universally beat it into our students’ heads that this long, dipping contour is in fact the basic, fundamental contour of tone 3? I have long been of the opinion that it would be much easier for students if they were simply told that the basic contour of the third tone is short and falling, which happens to become tone 2 directly in front of another tone 3 syllable, and that at the end of a phrase or in isolation it is often (but not always) pronounced with a long, dipping contour instead. Given the fact that foreign language teaching, especially at the most elementary level, is by nature a conservative and prescriptive enterprise, and given that linguistics is a field where an argument like “That is what I usually hear” lacks much persuasive power, it is not surprising that very few of my linguist or language teaching colleagues with whom I have discussed this issue have naturally taken to my proposal. In fact, anything short of “proving” that the tone 3 long, dipping contour is in fact dead and gone is not likely to induce most language teachers or linguists to embrace my proposal. It is only through my refamiliarization with Taiwan Mandarin pronunciation over the past year that not only have I found a location where the long, dipping contour might indeed be on the road to retirement, but I may also have found evidence in the variety of ways that this tone is pronounced there to serve as a means of unifying in part, at least, my two tonal themes and help establish a hypothetical understanding of one or more possible motivations and one or more possible mechanisms for how tonetic and tonemic tone changes might take place in Chinese over time.

2. Research Design

As part of an ongoing, comparative, acoustical study of tone and vowel pronunciation in Taipei and Beijing Mandarin, subjects in each location were asked to read a list of 1,275 uniquely different Mandarin syllables, presented as individual Power point slides in a random order, one character at a time, with each character’s pronunciation indicated either in Zhuyin Fuhao or Hanyu Pinyin. In this way, we could access the closest approximation of a speaker’s sense of correct pronunciation without needing to worry about the effect of the surrounding phonetic environment, including prosodic effects. Of this list of 1,275 different syllables, 315 belonged to the Mandarin tone 3 category.

The results presented below are based on data from four different subjects, two male and two female, all of whom were in their early to mid twenties, university educated and who grew up in Taipei. Of these, three (T15, T19 and T20) were proficient native speakers of Taiwan Southern Min, while one (T1) was only fluent in Mandarin. In the end, a total of 1,246 separate tone 3 tokens were examined. This amount is slightly less than the expected total of 1,260, and is the result of the poor recording quality of 14 tone 3 syllable tokens, resulting in those 14 tokens being eliminated from the total number under consideration here.

Subjects were digitally recorded reading each of the 1,275 individual characters. Next, each recorded syllable was converted into an individual soundwave file in preparation for analysis. For the purposes of this paper, this analysis is limited to impressionistic perceptual discrimination coupled with observation of F0 pitch track contours and tone letter calculations using acoustic software developed at Nankai University.

3. Results

The results presented below must be considered in light of what has been reported by Chiung (1999). In this study she measured the pronunciation of 22 different subjects pronouncing just one bisyllabic word that ended with the tone 3 morpheme wen3 (吻). In the end she reports that Taiwan Mandarin tone 3 “is becoming a falling (31) tone”. The data presented below certainly supports her claim that tone 3 in Taiwan Mandarin is in general evolving in the direction of a low, falling contour. However, because she limited her sampling to just a single pronunciation of a single, shared token morpheme across all subjects, Chiung failed to discover an important fact about the pronunciation of the tone 3 contour in Taiwan Mandarin— Although a falling contour is indeed the norm for many speakers, it is still occasionally possible for them to use either a level, rising or dipping contour as well. Given that Gandour (1984) reports that pitch contour is the single-most important perceptual clue that native Mandarin speakers use to identify a particular tone, the fact that tone 3 may take on a level, rising or falling contour has potential implications for perceptual miscues among listeners, some of whom may be children in the process of acquiring Taiwan Mandarin as a native language. Representative examples of each of the four possible contour categories realized in the pronunciation of the
four subjects is shown in Table 1 below.

Table 1: Representative Examples of the Different Contour Outputs for Taiwan Tone 3 [Insert soundwaves]

<table>
<thead>
<tr>
<th>Level</th>
<th>Rising</th>
<th>Dipping</th>
<th>Falling</th>
</tr>
</thead>
<tbody>
<tr>
<td>廣 guang3 (T19)</td>
<td>老 lao3 (T1)</td>
<td>馬 ma3 (T1)</td>
<td>把 ba3 (T19)</td>
</tr>
</tbody>
</table>

Although both the level and falling tone 3 contours above may sound in isolation as if they would be confused with tones 1 and 4 respectively, in reality, such confusion is not possible, due largely to the fact reported by Jongman et al. (in press) that “to identify tones differing only in F0 height, listeners must refer to their knowledge of the speaker’s F0 range.” That is, although in isolation the syllables with the level and falling contours may sound as if they are being pronounced as tones 1 and 4, once we contrast their pronunciations with their genuine tone 1 and tone 4 counterparts we can immediately hear the differences. This is shown in Tables 2 and 3 below.

Table 2: Contrast Between Level and Falling Tone 3 Contours and Tones 1 and 4 [Insert soundwaves]

<table>
<thead>
<tr>
<th>Subject</th>
<th>Level</th>
<th>Tone 1</th>
<th>Falling</th>
<th>Tone 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>T19</td>
<td>廣 guang3</td>
<td>光 guang1</td>
<td>把 ba3</td>
<td>獄 ba4</td>
</tr>
</tbody>
</table>

Table 3: Tone Letter Calculations for T19 Guang1/Guang3 and Ba3/Ba4

五度值曲線圖:

(T19 Guang1/Guang3)

五度值曲線圖:

(T19 Ba3/Ba4)

參考五度值:

一聲: 55
二聲: 33（有微升的趨勢）
三聲: 54
四聲: 51

On the other hand, it would appear that the rising contour of tone 3 may have the potential to be confused with tone 2. In our examination of the 1,246 tokens of tone 3 pronunciation we did discover a total of three tokens of tone 3 pronunciation that near-physically, at least, seem to have merged with tone 2. This is presented in Tables 4 and 5.

Table 4: Near-Physical Mergers of Taiwan Mandarin Tone 2 and Tone 3 [Insert soundwaves]

<table>
<thead>
<tr>
<th>Subject</th>
<th>Syllable</th>
<th>Tone 2</th>
<th>Tone 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>T20</td>
<td>Fu</td>
<td>扶</td>
<td>府</td>
</tr>
<tr>
<td>T1</td>
<td>Lao</td>
<td>牢</td>
<td>老</td>
</tr>
<tr>
<td>T15</td>
<td>Yao</td>
<td>搖</td>
<td>咬</td>
</tr>
</tbody>
</table>
Please note that in the first case tone 2 seems to have acquired a dipping contour, while in the second case tone 3 seems to have acquired a sharp rising contour. Of course, one could also argue that the initial falling segments of lao2 and lao3 are not linguistically significant, and are in fact only background noise.

Returning once again to Chiung’s observation that Taiwan Mandarin tone 3 is evolving in the direction of a falling 31 contour, we see confirmation below in Table 6 that the overall pattern for the 1,246 tokens is indeed heavily weighted toward a falling contour.

Table 6: Total Distribution of Taiwan Tone 3 Contour by Contour Type and Speaker

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T15</th>
<th>T19</th>
<th>T20</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>0</td>
<td>22</td>
<td>2</td>
<td>30</td>
<td>54</td>
</tr>
<tr>
<td>Rising</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Dipping</td>
<td>259</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>268</td>
</tr>
<tr>
<td>Falling</td>
<td>51</td>
<td>277</td>
<td>309</td>
<td>280</td>
<td>917</td>
</tr>
<tr>
<td>TOTAL</td>
<td>312</td>
<td>308</td>
<td>313</td>
<td>313</td>
<td>1,246</td>
</tr>
</tbody>
</table>

In Table 6 we can clearly see that with the exception of T1, the citation pronunciation of tone 3 appears very much to exhibit a falling contour and not a dipping one, although it is also possible to encounter among these speakers a level, rising or dipping contour from time to time as well. The only important difference in the sociolinguistic profile of T1 on the one hand and the other three subjects on the other, is that the latter three subjects possess a strong command of Taiwanese while T1 does not. The possible relationship between a falling tone 3 contour in Taiwan Mandarin and the possible influence that Taiwanese may have had on this will be discussed below in the following section.

4. Discussion

Certain questions arise concerning the phonetic and phonemic facts of Taiwan Mandarin tone 3. What are the possible motivations for the underlying citation contour of tone 3 to shift from the prescriptive 214 or even so-called descriptive half-tone 21(1) contour to 31? Why does the contour of this tone sometimes even rise, even to the point of physically approximating the tone 2 contour 35? And why is it possible for a sharply falling tone 3 contour not to be confused with the

---

2 With the exception of subject T19’s single pronunciation of guang3, what is being called “level” here is in fact physically not nearly as high or as level as a genuine tone 1 contour. Perceptually, however, they appear to listeners to be level. That a slightly falling contour might be perceived by a Mandarin-speaking listener as being level is completely consistent with Wang (1967), where he documents that Mandarin listeners perceive synthesized stimuli along a level to rising contour continuum (which could just as well be a level to falling continuum) in a categorical fashion. Therefore, we can assume that although these “level” contours exist in rather large numbers, especially in the speech of T15 and T20, because of their salient pitch height discrepancy with a genuine tone 1 contour, there is actually little genuine risk of them ever being confused with tone 1.
sharply falling tone 4 contour?

As to why the tone 3 contour in citation form generally falls in Taiwan Mandarin rather than preserve the prescriptive dipping pattern, one can point both to prescriptive Mandarin itself and to Taiwan Southern Min for possible motivations. On the one hand, the percentage of times even in prescriptive Mandarin that a tone 3 syllable happens to be spoken in a prosodic environment that allows it to be pronounced with a full 214 contour is rather minimal compared with the percentage of times that it occurs in an environment where it absolutely cannot be pronounced with a dipping contour. What this means from the point of view of first language acquisition is that of the total number of tone 3 input tokens heard by a child, clearly the falling 21(1) contour is the most frequently encountered. Because of this, some children, at least, may be inclined to hypothesize that this is the underlying contour for Mandarin tone 3. And it just so happens that the citation contour of the corresponding tonal category in Taiwan Southern Min is also falling, although it exhibits a much sharper fall of 51. And as we have seen in Table 6, it is indeed the three subjects who are fully proficient native speakers of Taiwan Southern Min who clearly favor the falling contour for Taiwan Mandarin tone 3.

This still leaves the question of why the falling tone 3 contour 21(1) in prescriptive Mandarin may be expanding in Taiwan Mandarin in the direction of 31. In part, this can probably be accounted for by the strong possibility that because the fall in Taiwan Southern Min is so sharp, the fall in Taiwan Mandarin needs to be made more prominent. However, an additional factor must also be considered for why tone 3 seems to be evolving toward 31 and not remaining short and steady at 21(1)—The duration of a so-called Mandarin half tone 3, although already slightly falling, is considerably shorter than that of an ordinary syllable uttered in isolation. In other words, an open syllable in Chinese naturally requires a longer duration than what the Mandarin half third tone can offer, so it is only natural for its duration (and hence, its range) to be expanded. Since the contour endpoint is already at 1, the lowest point in a speaker’s natural pitch range, the only possible place for expansion to take place is at the top, which can only be accomplished by raising the onset further up the speaker’s pitch range. In this way, the “natural” need to raise that onset point of falling tone 3 is reinforced by the Taiwan Southern Min substratum pressure that comes from the 51 citation reading of 上声 syllables.

If the trend in Taiwan Mandarin is for tone 3 to be pronounced with a 31 contour, why then is it not somewhat confusable with tone 4, which prescriptively has a contour of 51? Chiang (1999) claims that while Taiwan Mandarin tone 3 is moving toward 31, Taiwan Mandarin tone 4 is moving toward a 53 and away from 51. This would suggest that a register distinction might be forming in Taiwan Mandarin, with tone 4 occupying the higher falling register and tone 3 occupying the lower falling register. The obvious motivation for tone 4 to evolve into a 53 contour would be to minimize its phonetic overlap with tone 3. This kind of reactionary development exemplifies a push chain. Such a scenario, where two tone categories have acquired different phonetic details in just a few decades of time due to mutual interaction, has the potential to explain in part at least, why cross-dialectal, and even cross-subdialectal comparisons so frequently reveal profound differences in tonetic details despite their commonly-shared tonemic systems. Unfortunately, the very little tone 4 data examined here to date does not confirm the existence of any examples of a 53 contour for tone 4, although it must be noted that more than 99% of the tone 4 data has yet to be examined.

At the same time as we note possible motivations for phonetic changes within tonemic categories, we should also address the phenomenon revealed in Tables 4 and 5, that every once in a while a tone 3 in Taiwan Mandarin gets pronounced with a rising contour that appears close enough physically to perceptually overlap with the normal tone 2 contour. Despite the physical evidence to suggest that such an overlap takes place, it is highly doubtful that the subjects recorded here ever themselves imagined that these items belonged to the tone 2 category, and if asked to repeat their reading of these characters, would very likely not pronounce them with a rising contour at all. From the standpoint of production, then, these rising contours, despite their approximation with the tone 2 contour, are simply an alotonic variation of tone 3. From the point of view of perception, on the other hand, especially among children acquiring the language, it is possible that if the rising tone were heard enough times for a particular morpheme, it may end up shifting from the 上声 category into the 阳平 category in their individual lexicons.

And what might the motivation be for the occasional realization of a rising tone 2 contour for a syllable that the speaker clearly understands to be a tone 3 morpheme? Again, Mandarin itself can provide part of the possible explanation and part of the explanation can also be provided by Taiwan Southern Min. From Mandarin we note that when tone 3 is immediately followed by another tone 3 syllable, the speaker effortlessly, and more importantly, unconsciously, pronounces that first tone 3 syllable with a tone 2 rising contour. Of course it is not possible to know for certain what is going through the mind of a speaker when s/he is pronouncing a tone 3 character in isolation, but it is not completely out of the realm of possibility to assume that from time to time s/he may be imagining a full word that contains the isolated character in question, and that some of the time the imagined word is composed of two tone 3 syllables with the character in question occupying the first slot. Such a scenario would naturally trigger sandhi in the first syllable. Imagined Mandarin tone sandhi, then, might explain the occurrence of the rising contours documented in Table 3. But imagined Taiwan Southern Min might also be a motivating factor,
as the sandhi form of a 上聲 morpheme just so happens to also be a rising contour, in this case 24. And finally, in support of the idea that even when a Mandarin tone 3 syllable is pronounced with a rising contour it might go unnoticed, Huang (2004) shows that Mandarin tone 2 and tone 3 "are indeed perceptually more confusable than any other two tone pairs."

5. Conclusion

In the end, what does any of this have to do with issues related to the possible motivation, mechanism and nature of sound change as discussed by Wang (1969)? We have clearly identified one case of movement toward phonetic change in Taiwan Mandarin within a tonemic category (the shift in the phonetic value of tone 3 from 214 or 21(1) to 31), and have also noted one potential future possible tonemic shift across categories (the possible future permanent shift of isolated morphemes from the tone 3 category to tone 2). In all cases we note variation (and hence, exceptions) as the seed of change which remains present over much of the time that the process plays out. These facts support Wang’s lexical diffusion model of sound change which posits variation as a fundamental part of sound change. All of this contradicts the assumption of the Neogrammarian model that assumes sound change is uniform, imperceptible and consistent, affecting all qualifying members simultaneously and universally. But on the other hand, much of this variation in pronunciation is not abrupt in the sense that it is not psychologically abrupt, as it is highly doubtful that speakers are aware of the great range of physical variation in their own pronunciation of the tone 3 contour despite the articulatory evidence to the contrary. Rather, from the speaker’s standpoint, variation at this stage can likely only be considered physically real, not psychologically real, though for the next generation of speakers who encounter this input first as listeners, some of that variation may eventually end up being re-interpreted as being real in every sense of the word. That categorical sound change derived from variation which takes place under the psychological radar of speakers should first be implemented by listeners and not by speakers is fully consistent with the speculation of Ohala (1981) that diachronic sound change can occur due to the listener’s misperception and/or reinterpretation of certain sounds.

References
Chiung, Wi-Vun Taiffalo, 1999. The tonal comparisons and contrasts between Taiwanese and Taiwan Mandarin. Handout presented at the Sixth Annual UTA Student Conference in Linguistics, 8 pages. 