Implosives and the Inherent F0 Perturbation in Chinese Dialects

CUN Xi
Department of Chinese, The Hong Kong Institute of Education, Hong Kong
cunxi@ied.edu.hk, citysun79@gmail.com

Abstract
It is gradually noticed that in a tonal language such as Chinese, phonation types are the phonetic cause of tonogenesis, different phonation modes can define three tonological registers, and the initial consonants of a syllable and the consonants at the end of the syllable would perturb the pitch contour [1]. In the traditional Chinese phonology, the “voiced” (actually, in most cases, they are sounds with slack voice, not real voiced consonants) initial consonants correspond to Yang tonal category, a lower tonal category, while the voiceless initial consonants correspond to Yin tonal category, a higher tonal category. However, implosives in Chinese dialects are contradicted to this rule. The voiced implosives are corresponding to the higher tonal category.

Previous studies discussed about the consonantal perturbation on F0 of the following vowels. The physiological mechanism of F0control is refined. However, the tonal perturbation of implosives in Chinese dialect is rarely mentioned. Current study is going to explore two questions: 1) which tonological register do implosive initial consonants correspond to? High, Mid, or Low? 2) How do implosive initials perturb the pitch contour?

It is found that generally, implosive initials correspond to the Mid tonological register, just as the voiceless initials do. But there are a few cases in which implosive initials can be produced with slack voice and creaky voice, when the syllable is produced with these phonation types. Also, it is found that implosive initials have its inherent F0. It ranges from 140-200Hz for male speaker; and 200-270Hz for female speaker. The inherent F0 for implosives produced with slack voice and creaky voice could be lower.

Index Terms: Implosive, Pitch contour, Consonantal Perturbation, Tonological register, Chinese dialects

1. Introduction
In a tonal language such as Chinese, the feature [Voiced] of the initial consonant is crucial to the category of tone. Such as in Cantonese, the “voiced” initial consonants in Middle Chinese is lost, but the role of which is substitute by a lower tonological register, which is called Yang Tonal category. Also, in Wu dialect, the slack voice initial consonants are paralleled with the low tonological register[2],[3].

Therefore, in Chinese dialects, the “voiced” (actually slack

voice)initial consonants corresponding to a lower tonological register is well acknowledged. However, it is found that in some Chinese dialects, a group of “Voiced” initial consonant is corresponding to the higher tonal category, or more appropriately the Mid tonological register[4],[5]. Previous studies have proved that these sounds, which were described as pre-glottalized voiced stops, are actually bilabial and alveolar implosives[6],[7]. So, there are two questions to be answered: 1) how to explain physiologically the “voiced” initial consonants associated with low tonological register; and 2) how do the voiced implosives initials behavior in the tones in Chinese dialects? What is the relationship between implosives initials and the tonological register, and the relationships between implosive initials and the pitch contour.

Since these sounds are rarely talked in Chinese, some people thought they were the relics of the language contact with the non-Chinese languages in the history [8]. But now, the phonetic characteristics of the sounds are clearer, when there are more reports about the newborn implosives in Chaozhou, Xiang and Gan dialects[9]. It is interesting to explore how the implosives influence the tonal category in Chinese dialects.

2. Literature Review
Studies concerning the consonant perturbation of F0 found that the influence of the voiced stops is different from that of the voiceless stops and implosives. Hombert [10] sampled F0 at 20 ms intervals from vowel onset to 100 ms after vowel onset, and found that the onset F0 of the following vowel was slightly higher after consonants such as voiceless stops and implosives; and was lower after voiced stops. Hombert and Ohala [11] investigated the physiological cause of the consonant perturbation. They discussed two hypotheses: the “aerodynamic” hypothesis and the “vocal cords tension” hypothesis. The “aerodynamic” hypothesis arguing that the perturbation is result of the different velocity of the airflow after the consonants, i.e. the velocity of airflow is higher after the voiceless stops than the voiced stops, and thus would produce a higher F0. However, this hypothesis could not explain how the effect can sustain for 100 ms or even longer. As for the “vocal cords tension” hypothesis, it ascribed the effect to the placement of the larynx, i.e., the larynx is actively lowered by the cricothyroid (CT) muscle. This hypothesis could not explain why
when produce implosives, the larynx is even lower, but F0 might be higher.

It is later known that fundamental frequency (F0) is controlled by several factors affecting vocal cords tension (VCT). For example: the position of the larynx, the sub-glottal pressure (P_s); the nervous system; and the activities of laryngeal muscles [12].

The VCT is directly proportional to the effective stiffness of the vocal folds [12]. The activity of the tissue layer and the muscle layer can be separated. When the cricothyroid (CT) muscle contracts, thyroarytenoid (TA) muscle is inactive. The vocal folds will be lengthened, and the effective stiffness of all the tissue layers (not the muscle layer) of the vocal folds will increase, and F0 will increase. The increase of F0 is by the increase of the passive longitudinal tension. This case is suitable in the control for high F0 with low intensity, such as falsetto. Therefore, the phonation type falsetto corresponds to the High tonological register [1].

As for the control of low to intermediate fundamental frequencies, both the CT muscle and the TA muscle will work. The contraction of TA muscle, which stiffens the vocal cords, will increase the active longitudinal tension of the vocal cords, and will attribute to the increase of F0. This tonal control mechanism can be applied to the Mid to Low tonological register.

As the control of F0, then, how do the initial consonants perturb the onset F0? It is found that the perturbation of initial consonants is always generated when initiate the vocal cords vibration. Different mechanisms used in initiating vocal cords vibration require different conditions of P_s and the position of the larynx, and thus have different effects on F0.

In order to initiate and later to maintain the vibration, it is required to maintain an adequate trans-glottal pressure drop (P_tr). To achieve this, one can “puff out the cheeks, lower the velum a little to allow some nasal escape, widen the pharynx in the anterior-posterior dimension or not narrow it in the media-lateral dimension, lower the larynx, slacken the vocal folds or hold the vocal folds a little apart” [13].

For the voiced stop, the vocal cords vibration is a damped oscillation. Because when the vocal cords are vibrating, P_s diminishes and equals quickly by the airflow through the glottis. Besides nasalization, to slacken the vocal cords (decrease VCT by release TA) is the strategy to maintain voicing. Titze [12] found that the phonation threshold pressure increase with F0. If the stiffness of the vocal cords (the muscle portion) is lowered, the system has less resistance to overcome. As we mentioned above, F0 is proportional to the VCT. So F0 will decrease with the damped oscillation of the vocal cords. The slackness of the vocal cords might last for a period of time even after oral release. In Chinese dialects, the voiced stop initials do not cause the low tonological register, because the perturbation on F0 will not last for the whole syllable. The low tonological register is caused by the slackness of the vocal cords and even hold the vocal cords a little part which are called slack voice and breathy voice, because the activity of the laryngeal muscles can sustain throughout the whole syllable [1].

As for implosive, it has a growing oscillation. The energy, which sustains the growing oscillation, comes from P_s. Since the larynx is actively moved down, P_s increases with the downward movement of the larynx. In this system, there is no physical factor that restrains the control of F0. The stiffness of the vocal cords can be adjusted freely to achieve the target F0. Therefore implosives do not raise or depress F0 of the following vowel.

Also, the downward movement of the larynx for an implosive is not by the action of CT, but by the sterno-hyoid (ST) muscle [12]. The larynx is lowered as whole box by the external muscle, and would not influence internal muscles of the larynx. Although the position of the larynx is lowered, F0 is not lowered.

In the above, the physiological cause of the tonal perturbation from phonation types and initial consonants has been investigated. In the following part of this paper, how do the implosives behave in tones in Chinese dialects. Which tonological register do implosive initials correspond to? How do implosives perturb the pitch contour?

3. Method

This study is based on my PhD thesis [7] about the phonetic properties of implosives in Chinese dialect. The description of the initial consonants, finals and tones had been finished. The aerodynamic data about the implosives were made which illustrated the negative intra-oral pressure of the sounds used in this study. Current study is focusing on to which tonological register the implosive initials correspond, and the inherent perturbation on F0.
Words selected for current study are minimal pairs with the same finals but different tones. The speakers are asked to read the tokens in a pair, so that the pitch value of the tokens in a pair can be compared without normalization, because the speakers have already read them contrastively.

The recordings of three speakers (a male ZM and two female YH and JY) in Chaozhou, two speakers (a male LX and a female LM) in Wuyang, and one male speaker (GSF) in Songjiang were selected for study.

First, is to identify the phonation types of each syllable. Second, the measurements of F0 on the initial consonants and the onset of the finals are made. For the initial consonants, three points are measured, the first (c1), middle (c2) and last (c3) points. For the onset F0 of the final, four points (f0, f1, f2, f3) on the first 1/3 duration of the final were selected to measure the F0. Among them, F0 of the first point of the final (f0)is equal to the last point of the consonant (c3). There are totally six points (c1, c2, c3, f1, f2, f3) to be measured in each syllable. The duration of the initial consonants (Dur_c) and the finals (Dur_f) are also measured.

In Songjiang dialect, the tonal perturbation of implosives, slack voiced stops and voiceless aspirated stops was compared. Syllables with the same final, but different initials (implosive, slack voiced stop, and voiceless aspirated stop) were chosen, e.g. 碰 [be52], 盤[be31]. There were 9 sections of bilabial stimuli, 9 sections of alveolar stimuli, and a group of palatal-alveolar stimuli of which 7 words were pronounced as palatal implosive. Each section was recorded two times in isolation, once in phrases, and once in a carrier sentence. The carrier sentence is 这个字读 __/, ko ko zì tu?__/. (this character is __).

In Wuyang dialect, syllables with the same final (or with the same main vowel as that in Ru Sheng), with the implosive initial and different tones were chosen, e.g. 刀[dou55], 僚 [dou24], 到[dou11]. (The voiceless aspirated stops were not compared with the implosives, because they are not in the same tonal categories.) There were 21 stimuli of the syllable [ba], 9 stimuli of the syllable [bei], 3 stimuli of the syllable [bat], 9 stimuli of the syllable [dou], 3 stimuli of the syllable [dou], and 1 stimuli of the syllable [dok]. Each section was read twice in isolation, once in phrases.

In Chaozhou dialect, syllables with the same final (or with the same main vowel as that in Ru Sheng), with implosive initial, but different tones were chosen, including, [ba] (4 stimuli), [bi] (9 stimuli), [bou] (10 stimuli), [bo] (2 stimuli), [be?] (2 stimuli), [di] (3 stimuli), [dou] (3 stimuli), [dou] (3 stimuli), [di] (7 stimuli), [da] (5 stimuli), [mua] (3 stimuli), [mua] (4 stimuli), [mua] (2 stimuli). Each stimuli was recorded two times in isolation and two times in the carrier sentence 这个字读 __/tsi a zi ʈa __/. (This character is __).

All the recordings were made by a laptop with an M-Audio Mobile pre USB preamp with a built-in audio interface that is designed to improve the quality of the sound. It is connected with a one-point stereo microphone Sony ECM-MS907. The acoustic recording and analyze were made by Praat.

4. Results

4.1 The relationship between implosive initial and the tonological register

The relationship between implosive initial and the tonological register is one of the most important questions to be investigated.

Data from Songjiang Wu dialect shows that implosive initial is corresponding to the Mid tonal register, which is contrasted with the Low register that is characterized by slack voice. As shown in the tonal category and register of Songjiang in Table 1, there are two tonological register Mid and Low in Songjiang dialect. The Yin and Yang tonal categories in Songjiang dialect correspond to the Mid and Low tonological registers, respectively. In each register, there are four tonal categories, Ping, Shang, Qu, Ru. The voiceless initials and implosive initials are in the Mid register. The Low register corresponds to the syllables with slack voice.

This pattern is just as what Chao [3] and Li [4] described, the “voiced” implosive initial correspond to Yin tonal category3, just as the voiceless stop initials do.

In Wuyang Yue dialect, implosive initials present in five Yin tones, Yin Ping, Yin Shang, Yin Qu, Yin Ru1, and Yin Ru2, as shown in Table 2. But the five Yin tones belong to two tonological registers. Yin Ping and Yin Ru1 belong to the Mid register. Yin Shang, Yin Qu and Yin Ru2 belong to the Low register, because when pronouncing these tones, speaker would produce creaky voice and/or slack voice parallel to the low pitch.

In Chaozhou Min dialect, implosive initials present in eight tones, Yin Ping, Yin Shang, Yin Qu, Yin Ru, Yang Ping, Yang Shang, Yang Qu, and Yang Ru, as shown in Table 3. Among the eight tones, Yin Qu and Yin Ru belong to the Low tonological register, because there are slack voice and creaky voice presenting with the low pitch.

It is found that the slack voice and creaky voice are parallel with the low pitch in all the cases. And, implosive initials can present in the Low tonological register. However, the question is whether the implosive initials are produced with slack voice or creaky voice, or not. If the implosive initials were produced with

3The Yin and Yang tonal categories in Songjiang dialect correspond to the Mid and Low tonological registers respectively.
modal voice, even though they appear in the low tonological register, they would not influence the tonological register of the following segment. If the implosive initials were produced with slack voice or creaky voice, then they might have influence the tonological register of the following segment inherently.

The results show that the implosive initials can be either produced with modal voice as well as with other phonations such as creaky voice and slack voice. In most cases, implosive initials are produced with modal voice, even though they are in the Low register, they are produced with different phonations with the following segment.

And there are a few cases, the implosive initials are produced with the phonations same as the following segment. For example, [ddɛu\(^{11}\)] in Wuyang is a low tone, as shown in Figure 1. The syllable initial is an alveolar implosive with increasing amplitude of the sound wave. The irregular sound pulses in the sound wave, and the unstable pitch contour shows that this pronounced with creaky voice. The creaky voice predominant the whole syllable, including the implosive initial. There are several irregular pulses in the middle of the initial consonant, where the pitch is as low as 92 Hz, which is very low for a female speaker. Since the vocal folds vibrate irregularly, the pitch changes abruptly from 92 Hz to 187 Hz in the middle of the initial consonant. These are showing that the implosive initial is not produced with modal voice, but with creaky voice as the following segment.

![Figure 1](image1.png)

Figure 1. [dœu\(^{11}\)] in Wuyang, the syllable is pronounced with creaky voice. The waveform and spectrogram are shown from up to down. The pitch contour is shown on the spectrogram with the dotted-dashed line.

Implosive initials can also be produced with slack voice, for example, [bbei\(^{11}\)] in Wuyang dialect, as shown in Figure 2. The sound with slack voice is with low pitch, 130-140 Hz for female speaker. There is random noise shown in the spectrogram. The amplitude of the harmonics in the middle and upper parts of the spectrogram is decreased. The energy above 1000 Hz is greatly decreased, as shown in the spectra (0-5 kHz) of the first few periods (50 ms).

![Figure 2](image2.png)

Figure 2. [bei\(^{11}\)] in Wuyang which is pronounced with slack voice. The waveform, spectrogram, the pitch contour and the spectra slice of the first 50 ms after release, are shown from up to down.

### Table 1. Implosive initials and the tonal category and tonological register of Songjiang dialect

<table>
<thead>
<tr>
<th>Register</th>
<th>Ping</th>
<th>Shang</th>
<th>Qu</th>
<th>Ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid</td>
<td>52</td>
<td>44</td>
<td>335</td>
<td>55</td>
</tr>
<tr>
<td>Low</td>
<td>31</td>
<td>113</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initials</th>
<th>Ping</th>
<th>Shang</th>
<th>Qu</th>
<th>Ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>碼 [bɛ]</td>
<td>擋 [bɐ]</td>
<td>当 [dɐ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>低 [dɛ]</td>
<td>补 [bʊ]</td>
<td>担 [dɛ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>滤 [pʰɛ]</td>
<td>土 [tʰu]</td>
<td>皮 [pʰɛ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>京 [ɕi52]</td>
<td>九 [jʊu44]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>低 [bɛ]</td>
<td>提 [dɛ]</td>
<td>对 [ʒʊ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>低 [dɛ]</td>
<td>提 [dɛ]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
higher than that of the male speaker. And it will be different between speakers.

It is found that, for the modal voice, the implosive initials have an inherent F0. For male speaker, it ranges from 140 to 200 Hz, and for female speaker it ranges from 200 to 270 Hz. The inherent F0 of implosive initial is related to the physical properties of the vocal cords. So, the inherent F0 is correct.

The result of T-value for Z1 is 5.569 (P<0.001, N=96), which means for these sample recordings, Z1>0 is significant at the level of 0.001. The hypothesis that the implosive initials have an inherent F0 is correct.

Since the implosive initials have an inherent F0, it is hypothesized that, if the F0 of point f1 (F0_f1) is higher than the mean F0 of c1 (F0_c1), then there will be rising curve on the pitch contour of the initial consonant (F0_c3>F0_c1); if the F0 of point f1 is lower than the mean F0 of c1, then, there will be a falling curve on the pitch contour of the initial consonant (F0_c3 < F0_c1).

\[ Z_1 = \frac{(F0_f1 - F0_c1)}{(F0_c3 - F0_c1)} \]

The result of T-value for Z1 is 5.569 (P<0.001, N=96), which means for these sample recordings, Z1>0 is significant at the level of 0.001. The hypothesis that the implosive initials have an inherent F0 is correct.

As for the male speakers, the average pitch value is lower than the inherent F0 of the implosive initial. The implosive initial will generate a falling curve on the pitch contour of the syllable. This is similar to the voiceless stop initials. As for the female speaker, since the pitch value of some tones can be even higher than the inherent F0 of the implosive initials, a rising curve is usually observed before high level and high falling tones; a falling curve is found before low level and low rising tones.

In the above, I discussed about the inherent F0 of the implosive initial which is produced with modal voice. As for those produced with other phonation types, such as slack voice and creaky voice, the F0 will be much lower.

### 5. Discussion

The relationship between implosive initial and the tonological register can be explained from two points. First is the physiological explanation. As mentioned in literature review, when pronouncing an implosive, there is no physical constrain on the change of VCT, so the vocal cords are in a default situation just as that when pronouncing voiceless consonants. Therefore, implosive initials correspond to Mid tonological register, same as the voiceless initials.

Second, this phenomenon can be explained from the historical sound change. Implosive initials historically derived from the homo-organic voiceless stop initials [14]. The sound change from stiff voiceless stops to implosives is ongoing in Chaozhou Min dialect [15]. It is also found that implosives are the phonetic variation of voiceless stops in Zhangzhou, Quanzhou Min dialect, northern Gan dialect, Guibeituhua, Xiangtan Xiang dialect, Liancheng Kejihua, and Zaomin Yao language [1]. Implosive initials derived from the voiceless stop initials and substitute the role of voiceless initials in the tonal system. Therefore, it is naturally to observe that implosive initials correspond to the same tonological register as the voiceless stop initials do.
Generally speaking, implosive initials correspond to the Mid tonological register. However, there are a few cases in which implosive initials are pronounced with creaky voice or slack voice. Since slack voice and creaky voice usually parallel with the Low tonological register, implosive initials may also correspond to the Low tonological register. Therefore, implosives may be produced by different laryngeal activities when the vocal cords are closed. If the vocal cords are closed by stretch TA, then there will be a high F0; if the vocal cords are closed by compress the vocal cords by the media pressure, or the vocal cords are kept apart to facilitate vibrating, then a lower F0 will be generated. This is how implosive initials related to the Low tonological register.

6. Conclusion
By investigating the implosive initial in Songjiang Wu dialect, Wuyang Yue dialect, and Chaozhou Min dialect, it is found that generally, implosive initials correspond to the Mid tonological register, just as the voiceless initials do. But there are a few cases in which implosive initials can be produced with slack voice and creaky voice, when the syllable is produced with these phonation types. Also, it is found that implosive initials have its inherent F0. It ranges from 140-200Hz for male speaker, and 200-270Hz for female speaker. The inherent F0 for implosives produced with slack voice and creaky voice could be lower. As for the male speaker, implosive initials usually generate a falling curve on the pitch contour.

7. Reference