Inconsistent Consonantal Effects on F0 in Cantonese and Mandarin

Qian Luo, Karthik Durvasula, Yen-Hwei Lin

Department of Linguistics and Languages, Michigan State University, USA
luoqian@msu.edu, durvasul@msu.edu, liny@msu.edu

Abstract

Previous research shows that aspiration and sonorancy can have inconsistent consonantal effects on vowel F0 across languages, and even within the same language in different studies: they are reported to either lower F0, raise F0, or have neutral effects. This paper is interested in such inconsistent consonantal effects on vowel F0 in Cantonese and Mandarin. The results of a series of production experiments show that: (1) consonantal effects on vowel F0 are language-specific; (2) the consonantal effects can be conditioned by lexical tones. The findings also provide insights for puzzles in historical tone change.

Index Terms: tone, consonantal effects, production, Cantonese, Mandarin

1. Introduction

Consonants can often influence F0 of adjacent vowels. Voiced obstruents are known to consistently lower F0 of the adjacent vowels across languages, while voiceless obstruents usually raise vowel F0. This effect has been argued to form the basis of tonogenesis and tone split [1], [2]. However, there is no consensus regarding the effects of aspiration and sonorancy on vowel F0. Some studies show that aspiration lowers pitch [3], [4], while others suggest that it raises pitch [5], [6]. Likewise, sonorants are pitch depressors in some studies [7], [8], but neutral in others [9], [10]. The current study is interested in such inconsistent consonantal influences in tonal languages and asks (1) what are the consonantal effects on vowel F0 in tonal languages? (2) What factors condition these consonantal effects on F0? The results may also provide insights into historical tone change in Chinese.

2. Background

This section provides the overview for the consistent and inconsistent effects on F0, as well as the reported interaction between consonantal effects and lexical tone. The background for the target languages is also reviewed.

2.1. Consistent Consonantal Effects on F0

Voiced obstruents consistently depress F0, and voiceless ones consistently raise F0 of the adjacent vowel. This is a robust observation found in many languages, especially those that have [voice] as a contrastive feature [11].

There have been two proposed sources for the voicing effects on F0: (a) the *automatic account* proposes that this effect is due to physiological and aerodynamic factors [12], [13]; (b) the *controlled account*, on the other hand, does not consider the consonantal effects as unintended side effects, but as an outcome of controlled articulation for enhancing the contrastive feature [voice] [11].

2.2. Inconsistent Consonantal Effects on F0

Unlike the voicing effects, there is no agreement in the literature on consonantal effects on F0 by aspiration and sonorancy. The inconsistency of the consonantal effects may be due to unstable articulation or aerodynamic properties when producing aspiration and sonorancy, while voicing is more stable. It has also been found that the aspiration effect on pitch perturbation varies among speakers [5], [6], which suggests considerable individual differences for the aspiration effects. Alternatively, as per a *controlled account*, the asymmetric pattern of consistent and inconsistent effects is probably due to different demands for enhancement of the features [voice], [spread glottis] and [sonorant].

2.3. Consonantal effects conditioned by lexical tone?

Some tonal languages are reported to have interactions between consonantal effects and lexical tonal categories. [3] reported a lowering effect of aspiration in Mandarin only in the mid rising T2(35)[1] and the low T3(214), but the effect was not observed in other tonal categories. [4] also found different magnitudes of consonantal effects in different Cantonese lexical tone contexts. However, there is no generalized account for what tonal category is most closely associated with consonantal pitch perturbation related to aspiration. As part of the research question (2), the current study asks whether lexical tonal contexts play a crucial role in conditioning consonantal effects on F0 in tonal languages.

2.4. Target Languages

The target languages are Mandarin and Cantonese, both of which have a two-way laryngeal contrast in obstruents: aspirated voiceless obstruents and unaspirated voiceless obstruents. Cantonese has six long contrastive tones: high-level T1(55), mid-rising T2(35), mid-level T3(33), low-falling T4(21), low-rising T5(23) and low-level T6(22). Mandarin has four lexical tones: high-level T1(55), mid-rising T2(35), low T3(214) and high-falling T4(51).

Mandarin and Cantonese are selected as target languages for the following reasons. Firstly, there are inconsistent reports of consonantal effects on F0 in these two languages (aspiration lowering: [3], [4]; aspiration raising: [5]);

Secondly, Mandarin and Cantonese have different historical tone merger processes in the rising tonal category. In Cantonese, low rising tones with sonorant onsets followed those with voiced obstruent onsets to merge with the falling tone category. On the other hand, Mandarin had an unexpected tone merger: low rising tones with sonorant onsets merge with high rising tones with voiceless obstruent onsets, instead of following voiced obstruent onsets to merge with the falling tone category. The different diachronic patterns provide the

[1] According to Chao tone numbers, the tone code ‘5’ indicates the highest pitch of the pitch range and ‘1’ the lowest [14]; e.g., T1(55) is the transcription for the high level tone.

background for a follow-up study, which can make use of our findings of consonantal effects on F0 in these two languages in the current study.

3. Production Experiments

This section presents the design of a series of production experiments for the purpose of investigating the consonantal effects on F0 in Cantonese, Cantonese speakers' Mandarin and monolingual speakers' Mandarin.

3.1. Participants

All participants who volunteered in the study are undergraduate or graduate students at Michigan State University. Eight Cantonese speakers (5 female, 3 male) are from Guangzhou city in China and speak standard Cantonese natively and native-like standard Mandarin. Six of the eight Cantonese speakers (3F, 3M) continued to participate in the Mandarin production experiment after the Cantonese experiment, to test any effect from bilingualism

Six monolingual Mandarin speakers (2 female, 4 male) participated in the Mandarin production experiment. They are all from Northern China and do not speak any other Chinese languages except Standard Mandarin.

3.2. Stimuli

Cantonese stimuli (n=86) include six Cantonese tones: T1(55), T2(35), T3(33), T4(21), T5(23) and T6(22), and Mandarin stimuli (n=86) cover four Mandarin tones: T1(55), T2(35), T3(214) and T4(51). All stimuli have a CV template. There are three types of initial consonants: aspirated obstruents [t^h, k^h, p^h], unaspirated obstruents [t, k, p] and sonorants [m, l, n]. The vowels in the Cantonese stimuli are [a, ɔ, e], and the vowels in the Mandarin stimuli are [a, i, u]. There are lexical gaps in the following categories in Cantonese (i.e. no stimuli in these Cantonese subgroups): sonorant onsets with T3(33), aspirated onsets with T6(22), and unaspirated onsets with T4(21) and T5(23).

3.3. Procedures

Stimuli were presented in a randomised order through PsychoPy [15] on a MacBook Air. There were instructions and a practice session before the experiment. Participants were asked to produce three repetitions of each sentence presented on the screen at a normal pace. The carrier phrase for Cantonese stimuli was [ŋɔ5 kɔŋ2 _______ tsi6 sam1 ts^{h}i3] 'I say ____ character for three times', and the carrier phrase for Mandarin stimuli was [wɔ3 ʂwɔ1 _______ tsɯ4 san1 ts^{h}i4] 'I say ____ character for three times'. Each sentence was displayed on the screen for 8 seconds and participants had to finish the three repetitions within that amount of time. A Logitech desktop microphone was used to record participants' data. Cantonese speakers participated in the Cantonese production experiment. Six of the Cantonese speakers then did the Mandarin production experiment. Monolingual Mandarin speakers did only the Mandarin production experiment

3.4. Measurements

Praat [16] was used to track the labeled target syllable and yield F0 values. The F0 values were extracted at every 5ms for the first 50 ms of the F0 trajectory. The F0 values were normalized into cents in R [17]. Following the findings in [4], where significant consonantal effects were mostly found within 20ms into the vowel in Cantonese, the mean normalized F0 values for the first 20ms following the voicing onset were used for analysis. Following [18], [19], a sonorant baseline is used to interpret the consonantal effects.

4. Results

Three factors (consonant type, lexical tone, vowel) are taken into consideration in the experiment. Three-way ANOVA tests were performed in R for data analyses. The presentation of the results only involves pooled analyses, which use mean normalized F0 values across the three repetitions of all participants as the dependant variables.

As shown in Table 1, neither consonant type nor vowel type significantly influences onset F0 within the first 20ms in any of the three language contexts. However, the lexical tone factor has a significant effect on onset F0 in all language contexts. No other crucial interaction is found.

Table 1. *ANOVA test on the onset F0 values in three language contexts, C=Consonant, T=Tone, V=Vowel*

Source of effects	**Cantonese**		**Cantonese Mandarin**		**Native Mandarin**	
	F	P	F	P	F	P
T	2.26	<.05*	7.71	<.01**	30.61	<.01**
C	0.59	>.05	0.82	>.05	0.29	>.05
V	0.23	>.05	0.20	>.05	1.12	>.05
T×C	0.32	>.05	0.12	>.05	0.24	>.05
T×V	0.25	>.05	0.25	>.05	0.06	>.05
C×V	0.17	>.05	0.37	>.05	0.16	>.05
C×V×T	0.16	>.05	0.16	>.05	0.11	>.05

4.1. Results: Cantonese

Figure 1 shows the F0 trajectories in each Cantonese tonal category: Within the first 20ms, F0 after both types of obstruents is generally higher than F0 after sonorants. Beyond 20ms, F0 after different types of consonants tends to converge. For T4(21), there is a cross-over between aspirated and sonorant around 15-20ms: F0 after aspirated dramatically drop below the sonorant baseline.

ANOVA tests (Table 2) show that, within the first 20ms, the consonantal effect is only significant in T1(55): where F0 following aspirated obstruents is significantly higher than after sonorants, and F0 after unaspirated obstruents is in intermediate between aspirated and sonorants. In other tonal contexts, no crucial consonantal effect is found. Two participants' data for T5(23) were excluded due to the unsuccessful recording of the aspirated stimuli.

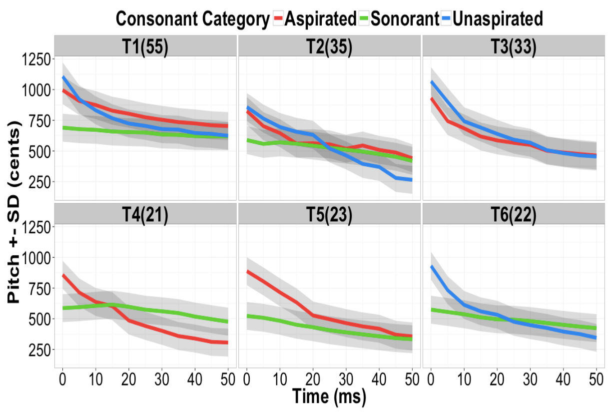

Figure 1: F0 (cents) in Cantonese (50ms). Shading indicates plus/minus one standard derivation.

Table 2. *Follow-up tests in each Cantonese tone ([TH, T, N] represents mean F0 values after aspirated, unaspirated and sonorant onsets respectively)*

Tone	F	P	t-test
T1(55)	[2, 14] 2.84	<.05*	TH ≈ T: t(7)=1.50, p >.05 TH > N: t(7)=3.67, p<.01** T ≈ N: t(7)=1.41, p >.05
T2(35)	[2, 14] .60	>.05	
T3(33)	[1, 7] .51	>.05	
T4(21)	[1, 7] 2.85	>.05	
T5(23)	[1, 5] .10	>.05	
T6(22)	[1, 7] 3.65	>.05	

4.2. Results: Native Mandarin

Figure 2 shows the F0 trajectories in each Mandarin tone produced by native Mandarin speakers. For T1(55) and T4(51), F0 after sonorants is lower than after both types of obstruents within the first 50ms. Using sonorants as the baseline, both aspirated and unaspirated onsets are shown to have a raising effect in T1(55) and T4(51). However, for T2(35) and T3(214), it is the F0 trajectory after aspirated obstruents that is lower than unaspirated obstruents and sonorants. This is consistent with the aspiration lowering effect in Mandarin reported by [3].

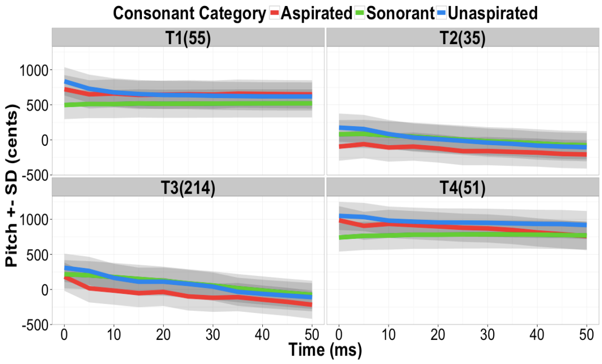

Figure 2: F0 (cents) in Native Mandarin (50ms). Shading indicates plus/minus one standard derivation.

ANOVA tests (Table 3) show that the consonantal effect within the first 20ms is significant in T1(55), T2(35) and T4(51). The follow-up t-tests show that in both T1(55) and T4(51), F0 after sonorant onsets is significantly lower than after both types of obstruents, supporting the raising effect of aspirated and unaspirated obstruents. F0 is significantly lower after aspirated obstruents than after unaspirated obstruents in T1(55), but such aspiration effect is insignificant in T4(51), showing that the raising effect of unaspirated voiceless onsets is greater than aspirated voiceless in T1(55), but not in T4(51).

In the mid-rising T2(35), F0 is significantly lower after aspirated obstruents than after unaspirated obstruents and after sonorants, whereas no significant difference is found between unaspirated obstruents and sonorants. Using the sonorant baseline, a siginificant aspiration lowering effect is supported in T2(35), as reported in [3].

Table 3. *Follow-up tests in each Native Mandarin tone ([TH, T, N] represents mean F0 values after aspirated, unaspirated and sonorant onsets respectively)*

Tone	F	P	t-test
T1(55)	[2, 10] 12.07	<.01**	TH < T: t(5)=2.83, p <.05* TH > N: t(5)=4.23, p <.01** T > N: t(5)=2.65, p <.05*
T2(35)	[2, 10] 7.00	<.05*	TH < T: t(5)=2.43, p <.05* TH < N: t(5)=2.71, p <.05* T ≈ N: t(5)=2.34, p >.05
T3(214)	[2, 10] .14	>.05	
T4(51)	[2, 10] 6.59	<.05*	TH ≈ T: t(5)= .85, p >.05 TH > N: t(5)=3.31, p <.05* T > N: t(5)=2.84, p <.05*

4.3. Results: Cantonese Speakers' Mandarin

Figure 3 shows the F0 trajectories in each tonal category in Cantonese speakers' Mandarin. Within the first 10-20ms, F0 after both types of obstruents is generally higher than after sonorants. Beyond 10-20ms, the F0 converge. More interestingly, F0 after aspirated obstruents are lower than after unaspirated obstruents and sonorants beyond 10-20ms in T2(35) and T3(214), which is similar to Native Mandarin in Figure 2. This suggests that Cantonese speakers' Mandarin is not Mandarin native-like in the first 10-20ms, but patterns with Native Mandarin beyond that time.

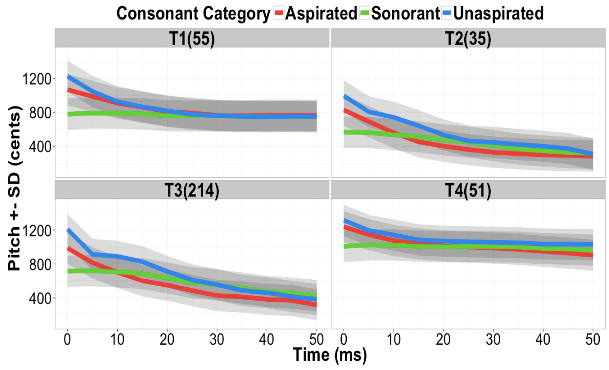

Figure 3: F0 (cents) in Cantonese Mandarin (50ms). Shading indicates plus/minus one standard derivation.

Table 4. *Follow-up tests in each Cantonese Mandarin tone*

Tone	10ms		20ms	
	F	P	F	P
T1(55)	[2, 10] 1.34	>.05	[2, 10] 1.93	>.05
T2(35)	[2, 10] 2.41	>.05	[2, 10] .51	>.05
T3(214)	[2, 10] 1.04	>.05	[2, 10] 1.41	>.05
T4(51)	[2, 10] 4.46	<.05*	[2, 10] 4.00	>.05

ANOVA tests (Table 4) show that the consonantal effect is not significant in any of the lexical tones within the first 20ms. However, the consonantal effect is significant within 10ms in T4(51). The follow-up t-test shows that F0 is significantly higher after unaspirated obstruents than after sonorants [t(5)=2.87, p<.05*], while no other comparison is statistically significant.

4.4. Results: Summary

Across the three language contexts, consonantal effects are not consistent. Comparing just the high-level T1(55) context, F0 after obstruents is significantly higher than after sonorants in both Cantonese and Native Mandarin. However, the mid-rising T2(35) context reveals the language-specific difference: within the first 10-20ms, F0 values after both types of obstruents are higher than after sonorants in Cantonese and Cantonese Mandarin (Figure 1 and 3), but a significant aspiration lowering effect is found in Mandarin mid-rising tone. Moreover, beyond 10-20ms, Cantonese Mandarin begins to pattern with Native Mandarin, suggesting that consonantal effects in Cantonese Mandarin may be influenced by Cantonese speakers' L1 and L2 phonology at different periods along the F0 trajectory.

Within the same language, consonantal effects on F0 are found in specific tonal categories, especially in tones that are associated with high pitch, such as Cantonese T1(55), Mandarin T1(55), Mandarin T2(35) and Mandarin T4(51).

5. Conclusion and Discussion

The results of our study show that the consonantal effects on vowel F0 can be language-specific and conditioned by lexical tone. For high-level T1(55) in Native Mandarin, F0 after both obstruents is above the sonorant baseline. For mid-rising T2(35) in Native Mandarin, however, an aspiration lowering effect is found. For T1(55) in Cantonese, a significant aspiration raising effect is found. For Cantonese speakers' Mandarin, no significant aspiration effect is found in any of the tonal contexts in our results, but Figure 3 suggests that Cantonese Mandarin behaves like Native Mandarin beyond 10-20ms.

It is interesting to see robust consonantal effects in tones associated with high pitch. From an automatic perspective, this could be due to articulatory factors: the range of allowable articulation variation for high pitch may be wider than mid or low pitch, which leads to more flexibility for consonantal effects to take place in high pitch. This can be parallel to vowel coarticulation reported by [20], which shows that low vowels (like high pitch) allow the greatest variability due to the fact that maximal articulation of low vowels is rarely reached[2].

The results of our study also provide some interesting insights into the diachronic changes of Chinese languages. In historical tone change from Middle Chinese to Modern Mandarin, low rising tones with sonorant onsets merge with high rising tones with voiceless obstruent onsets, instead of following voiced obstruent onsets to merge with the falling category. Our results partially explain this diachronic merger: it is possible that aspiration has a lowering effect on onset pitch in the rising category, and thus led to a lower rising contour shape, stabling the perception of rising tone. On the other hand, the onset pitch after unaspirated obstruents might be high and gave rise to a misperception of a falling shape. This only happened in Mandarin since such aspiration lowering effect is language-specific. This hypothesis still needs further support from other languages that have the same historical tone merger.

6. Acknowledgements

We gratefully acknowledge the valuable feedback from Christian DiCanio, Carlos Gussenhoven, James Kirby, the anonymous reviewers and the audience at the 5th TAL. We would also like to thank the Phonetics and Phonology Group at Michigan State University for their comments and support.

7. References

[1] Haudricourt, A. G., "De l'origine des tons en vietnamien," *Journal Asiatique*, no. 242, pp. 69–82, 1954.

[2] Kingston, J., "Tonogenesis," in *The Blackwell companion to phonology*, vol. 4, M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice, Eds. Oxford, UK: Blackwell Publishing, pp. 2304–2333, 2011.

[3] Xu, C. X. and Xu, Y., "Effects of consonant aspiration on Mandarin tones," *J. Int. Phon. Assoc.*, vol. 33, no. 2, pp. 165–181, 2003.

[4] Francis, A. L., Ciocca, V., Wong, V. K. M. and Chan, J. K. L., "Is fundamental frequency a cue to aspiration in initial stops?," *J. Acoust. Soc. Am.*, vol. 120, pp. 2884–2895, 2006.

[5] Zee, E., "The effect of aspiration on the F0 of the following vowel in Cantonese," *UCLA Work. Pap. Phon.*, vol. 49, pp. 90–97, 1980.

[6] Lai, Y., Huff, C. and Jongman, A., "The Raising Effect of Aspirated Prevocalic Consonants on F0 in Taiwanese," *Proc. 2nd Int. Conf. East Asian Linguist.*, 2009.

[7] Jeel, V., "An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants," *Annu. Rep. Inst. Phon. Univ. Cph.*, vol. 9, pp. 191–211, 1975.

[8] Maddieson, I., "The effects on F0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis," *J. Phon.*, vol. 12, pp. 9–15, 1984.

[9] Gandour, J., "Consonant types and tone in Siamese," *J. Phon.*, vol. 2, pp. 337–350, 1974.

[10] Kjellin, O., "Observations on consonant types and 'tone' in Tibetan," *J. Phon.*, vol. 5, pp. 317–338, 1977.

[11] Kingston, J. and Diehl, R. L., "Phonetic Knowledge," *Language*, vol. 70, no. 3, pp. 419–454, 1994.

[12] Whalen, D. H. and Levitt, A. G., "The universality of intrinsic F0 of vowels," *J. Phon.*, vol. 23, no. 3, pp. 349–366, 1995.

[13] Connell, B., "Tone languages and the universality of intrinsic F 0: evidence from Africa," *J. Phon.*, vol. 30, no. 1, pp. 101–129, 2002.

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

[15] Peirce, J. W. P., "PsychoPy—psychophysics software in Python.," *J. Neurosci. Methods*, vol. 162.1, pp. 8–13, 2007.

[16] Boersma, P. and Weenink, D., "Praat, a system for doing phonetics by computer," *Praat: doing phonetics by computer [Computer program]*, 2015. [Online]. Available: http://www.praat.org/.

[17] R Core Team, *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015.

[18] Hanson, H. M., "Effects of obstruent consonants on fundamental frequency at vowel onset in English," *J. Acoust. Soc. Am.*, vol. 125, no. 1, p. 425, 2009.

[19] Kirby, J. and Ladd, B., "Stop voicing and F0 perturbations: evidence from French and Italian," *Proc. Eighteenth Int. Congr. Phon. Sci. Glasg.*, 2015.

[20] Manuel, S., "Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts," in *Coarticulation: Theory, Data and Techniques*, W. J. Hardcastle and N. Hewlett, Eds. Cambridge University Press, pp. 179–198, 1999.

[2] We are grateful to Christian DiCanio for bringing discussion of vowel coarticulation to our attention.