An Acoustic Study of Zhajin Gan Tone

Yingyi Zhou[1], James Kirby[2]

[1]The Hong Kong University of Science and Technology, [2]The University of Edinburgh
yzhoubm@connect.ust.hk, j.kirby@ed.ac.uk

ABSTRACT

This paper presents the first phonetic study of the tone system of Zhajin Gan. While important for our understanding of tonogenesis in general and Chinese historical phonology in particular, the tone systems of Gan dialects have not been analysed acoustically. Zhajin Gan can be analysed as having 6 or 7 tones in non-checked syllables, with most having two allotones conditioned by historical differences in onset type. However, as synchronic cues to the laryngeal contrast have weakened or disappeared, the previously redundant onset f0 differences are being phonologized, setting the stage for additional tone splits. In Zhajin Gan, plosives and affricates from MC *Ciqing* and *Quanzhuo* series are synchronically realized as lax-voiced stops, correlating with lower onset f0, but without any evidence for synchronic aspiration. We discuss the possible role that non-modal phonation may have played in the evolution of this complex tone system.

Keywords: Zhajin Gan, tone, f0 lowering, prevoicing, phonation

1. INTRODUCTION

Middle Chinese (MC) is traditionally reconstructed as having a three-way contrast between voiced, voiceless unaspirated, and voiceless aspirated initial obstruents [18]. In most modern Chinese languages, this system is sometimes preserved, but often reduced to a two-way contrast between voiceless and aspirated due to a devoicing of the voiced series. The Gan dialects, along with some Xiang and Wu varieties, are unusual in that instead, the voiced and aspirated onsets are understood to have merged [3, 40, 41]. However, the synchronic realization of this merged initial category in Gan has been variously reported as voiced unaspirated [3, 22, 23, 24, 32, 37], voiceless aspirated [39], voiced aspirated [6, 33], or co-existence of all the phonetic realizations mentioned above in different generations [5, 21].

Understanding the realization of these segments is also important for understanding the evolution of tone in Gan and in Chinese languages more generally. In most Chinese languages, the MC aspirated stops pattern with voiceless stops in terms of their effect on the MC tones, but in Gan, the aspirated stops unexpectedly *lower* tone. It has been suggested that the aspirated stops went through a phase of 'voicing' which conditioned this lowering [28], but also that historical [5, 22, 23, 24, 37] and synchronic [21, 35, 37, 38] differences in aspiration have been responsible for further tone splits, in particular among the historical *Yin Qu* and *Yin Ping* tones.

Finally, while Gan dialects have been the object of considerable attention, no acoustic studies of them have been carried out to date. Accordingly, descriptions of the tone system vary from researcher to researcher. In the dialects of Xiushui County, for instance, the number of tonemes (MC *Ru* included) reported varies from 6 to 11.

In this paper, we present the first acoustic study of a Gan Chinese dialect. We focus on the non-checked tones of Zhajin Gan, a Xiushui Gan variety. We concentrate on three questions:

(1) How many tonemes are in Zhajin Gan?
(2) Is there evidence for a synchronic difference in aspiration that may be responsible for tone lowering?
(3) What is the synchronic phonetic realization of the historical voiced and aspirated initials?

1.1 Location and vernacular

Xiushui County, located in the northwest corner of Jiangxi Province, China, lies at the junction of Hubei, Hunan and Jiangxi provinces. Zhajin town is located around 30km west to Yining town, the city center of Xiushui. According to *Language Atlas of China* [14, 16], Xiushui Gan is categorized into Chang-Jing and Chang-Du variety of Gan respectively.

1.2 Terminology and notation

Conventionally, tones in Chinese languages are described with reference to their presumed historical origins in terms of onset and rime types. For reference, the traditional terminology and related reconstructed phonetic realizations of MC onsets are given in Table 1. There are 4 reconstructed tones in MC, namely the *Ping* (level),

Shang (rising), *Qu* (departing) and *Ru* (entering) tones. MC initials are divided into four major categories: *Quanqing* (voiceless unaspirated), *Ciqing* (voiceless aspirated), *Quanzhuo* (voiced unaspirated), and *Cizhuo* (sonorants). The reconstructed MC initials are illustrated using labial consonants (bold form) and the notations (italic form) are listed in Table 1, and 'b' and 'f' used in the notation stand for plosive/affricate and fricative-initial tones respectively. Conventionally, tones with *Qing* (voiceless) initials are called *Yin* tones and tones with *Zhuo* (voiced) initials are called *Yang* tones. The evolution of MC *Ru* in Gan is also affected by the coda (*-p/t vs *-k); however, due to the complexity of this system and space constraints, we do not treat the *Ru* tones here.

For brevity, we refer to *Quanqing* onsets as P_A and the (in Gan allegedly merged) *Ciqing-Quanzhuo* onsets as P_B.

Table 1: Notation. (P_B segments are in thick borders.)

MC initial \ MC tone		*Ping*	*Shang*	*Qu*
Qing	*Quanqing* ***pa,*fa**	*1a*	*2a*	*3a*
	Ciqing ***pha,*fha**	*1c*	*2c*	*3c*
Zhuo	*Quanzhuo* ***ba**	*1b*	*2b*	*3b*
	Quanzhuo ***va**	*1f*	*2f*	*3f*
	Cizhuo ***ma**	*1m*	*2m*	*3m*

2. METHODS

2.1. Participants and data collection

All data were obtained during a field trip to Xiushui, Jiangxi in January 2017. Here, we present data from three speakers (one male: M1, aged 64, and two females: F1, aged 73, F2, aged 47), all from Zhajin town, Xiushui county. 503 monosyllabic tokens were chosen from *Fangyan Diaocha Zibiao* [15] to cover a range of onset, rime and tone types. Onsets were from the set /b d f h j k m n ɳ ŋ p s t tɕ ts v w z/. Speakers were recorded using Praat [1] via an Avid Mbox Mini connected to a Beyerdynamic TG H55C headset microphone. Tokens were presented to speakers as Chinese characters, one per PowerPoint slide and isolation forms were recorded. All the sound files were annotated in Praat [1] and analyzed using VoiceSauce v1.31 [29] to obtain acoustic measurements at a 1ms resolution.

2.2. Analysis methods

F0 curves were modelled using generalized additive models with the R *mgcv* package [36] Smoothed estimates were then computed using the *itsadug* package [33]. The plots shown below are based on these smoothed predicted estimates.

3. RESULTS

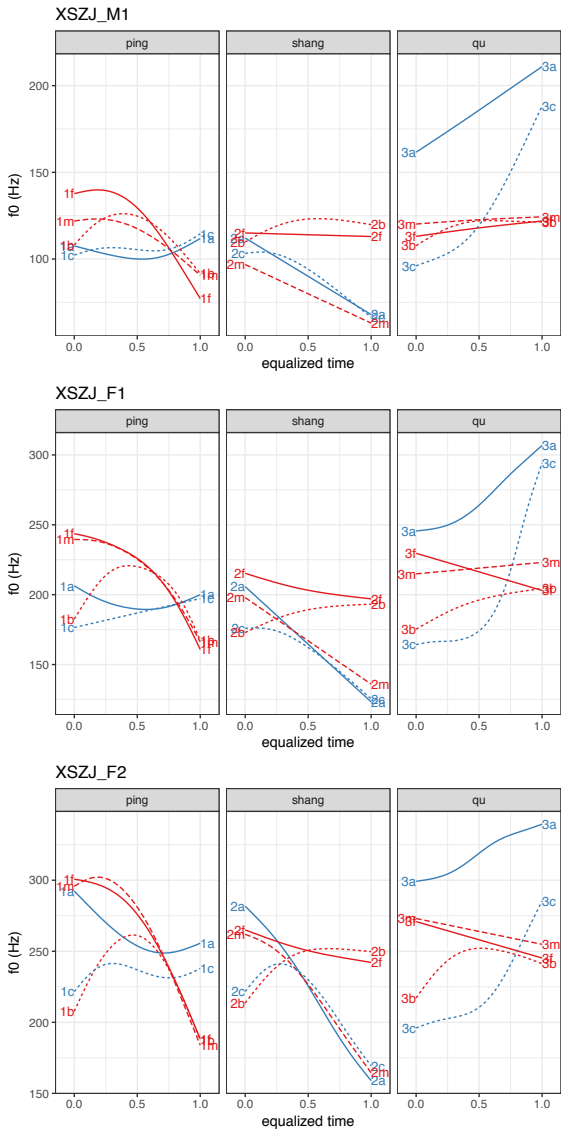
3.1 Tones

Fig. 1 shows smoothed estimates for each speaker. (i.) Broadly speaking, both historical aspirates (*-c)* and voiced obstruents (*-b*) are associated with lower onset f0. (ii.) Tones from historical *Shang* sonorants (*2m*) are unexpectedly merged with those of the *Yin* voiceless unaspirated series (*2a*); whereas in *Ping* and *Qu* they pattern with the *Yang* fricatives (*1c, 3c*). (iii.) The *Quanzhuo* tones in *Shang* and *Qu* (*2b, 2f, 3b, 3f*) appear to be merged.

The f0 curves can be analyzed into sub-groups based on initials. Basically, the MC *Yin* tone has allotones based on *-a* (MC unaspirated) and *-c* (MC aspirated) initials. The situation of the MC *Yang* tone is more complicated. Historical voiced plosives (*-b*) predictably act as depressors [7, 27, 30, 31, 42], while *-f* (MC voiced fricatives, now voiceless) and *-m* (MC sonorants) tend to pattern together. However, in the *Shang* series, MC *Cizhuo* sonorants (*2m*) have merged with *Quanqing* obstruents (*2a*), while the *Yang Qu* tones (*3m/3f*, *3b*) have merged with those of the *Shang* series, which is a well-documented sound change in Mandarin varieties, Gan and some Xiang varieties as well [11]. The *Yin Qu* tones, although behaving as expected with respect to historical aspiration, have developed much more extreme pitch divergences than the *Ping* or *Shang* series, and it is probably safe to say that tones *3a* and *3c* are distinct tonemes.

Interestingly, sonorant-initial tones pattern differently across the MC tonotypes. *1m* and *3m* pattern with *Yang* tones while *2m* patterns with *Yin* tone. Hirata [9] suggests that sonorants can be either lax or tense articulated. One possibility is that this is related to the pitch environment [26], that is, in high pitch environment, sonorants are articulated with rather tense vocal folds and tend to pattern with *Yin* tones, while in low pitch environment, they are produced with a rather lax vocal folds setting and tend to pattern with *Yang* tones. The extremely high pitch offsets of *3a* and *3c* would then be a very recently-arisen feature.

Figure 1: Predicted GAM smooths for tones based on historical initial and tone class by speaker. Red are *Yin* tones, blue *Yang* tones. Solid lines indicate tones with plain voiceless and fricative onsets; dashed lines indicate sonorant onsets; dotted lines indicate P_B onsets.



In terms of the absolute value and the temporal reach of the f0 differences, tones *-a*, *-c*, *-f*, and *-b* appear to be different tonemes for most of the speakers (except, perhaps, M1's *1a/c* and *2b/f* tones). However, we would first like to know if there are any phonetic differences in the realization of the onsets that might condition f0 lowering, which would support treating the *-c* and *-b* tones as allotones of the *-a* and *-f* tones, respectively.

Figure 2: VOT distributions of P_A and P_B by speaker. Blue are MC unaspirated, red are MC aspirated, and green are MC voiced.

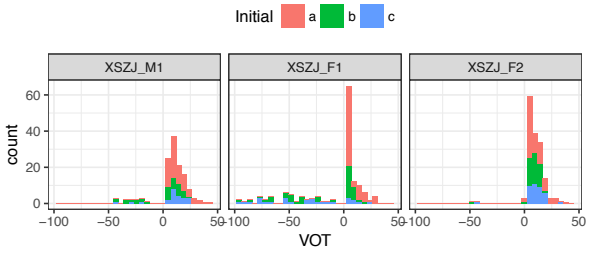
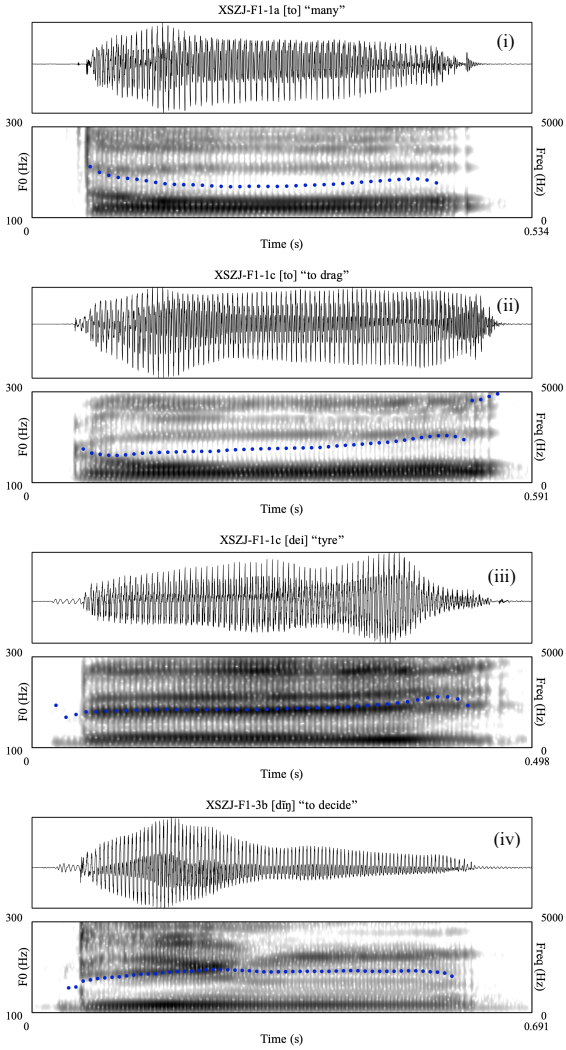


Figure 3: Spectrograms and waveforms of P_A (i) and P_B (ii, iii, iv) tokens produced by F1.



3.2 Synchronic aspiration and realization of P_B

It is clear from Fig. 1 that tones following historically aspirated stops (*-c*) have lower f0 onsets than those following historically voiceless unaspirated stops (*-a*). One suggestion [21, 35, 38, 39] is that this may be the result of a synchronic difference in aspiration. Fig. 2 shows the

distribution of VOT across the *-a* (voiceless), *-c* (aspirated) and *-b* (voiced) onsets. For M1 and F1, items with historical P_B onsets (*-c* and *-b*) are variably realized with closure voicing (22% for M1, 56% for F1); this is much rarer for the younger speaker, F2 (just 2%). Notably, there are no prevoiced velar onsets in the data.

When MC *Ciqing* and *Quanzhuo* are realized as voiceless, they are unaspirated, frequently with shorter (but not significant) VOTs than *-a* onsets. Fig. 3 shows some spectrographic examples comparing P_A onsets to P_B onsets realized with and without voicing lead. Conditionally, the initial portions of the vowels following *-c* and *-b* onsets can be breathy, with a rising amplitude envelope and some spectral degradation compared to those following *-a* onsets. Space does not permit a full analysis of the voice quality differences, but there is at least some evidence that syllables bearing P_B initials may involve breathiness to induce/reinforce the onset f0 lowering effect [8], which is believed to be indispensable for further tone splitting [2].

Furthermore, syllables with P_B initials are often realized as approximants or fricatives of the same articulatory place by M1, especially velar stops, consistent with the idea that voiceless P_B segments are produced with a lax vocal fold setting [3]. There also appears to be a relationship between prevoicing and f0 differences: M1 prevoices a fair number of P_B initials and has a relatively small difference in terms of f0 between *1a* and *1c* and *1b* and *1m/1f*, while F2 has fewer prevoiced initials but a larger onset f0 difference. This may be indicative that a further tone split is in progress.

4. SUMMARY

(1) How many tonemes are in Zhajin Gan?

Zhajin Gan could be acoustically analysed as having 6 or 7 tones (M1) in non-checked syllables, with most having two allotones conditioned by historical differences in onset type. However, in the speech of the female speakers (especially the younger speaker F2), the loss of the historical laryngeal contrast appears to be connected to further tone splits. Speaker F2 might be analysed as having 10 distinct tones on live syllables: *1a, 1c, 1b, 1f, 2a(2m), 2c, 3a, 3c, 3b(2b), 3m(2f,3f)*. However, perceptual experiments would be necessary to assess if these differences are perceptually salient.

(2) Is there evidence for a synchronic difference in aspiration that may be responsible for tone lowering?

No acoustic evidence shows a synchronic onset aspirated-unaspirated contrast. But tones on syllables with historically aspirated plosives (tone *-c*) have lower onset f0, and show some evidence of breathy phonation at vowel onset.

(3) What is the synchronic phonetic realization of the historical voiced and aspirated initials?

Both historical voiced (*-b*) and aspirated (*-c*) initials are realized with and without prevoicing by the older speakers (M1 and F1), but in the speech of the younger female (F2), prevoicing is rare. While voice quality differences might distinguish the *-c* and *-b* onsets from the plain *-a* onsets, they appear to be otherwise acoustically similar to one another.

Acoustic cues other than pitch are often used to signal tonal contrast [10] and the concomitant cues can turn into phonologically distinct features in certain circumstances. As for the older speakers in our sample, they seem to employ multiple cues (vocal fold vibration, breathy phonation, and f0 difference) to distinguish certain tones/onsets. Breathiness may also play a role in the merging of MC *-c* and *-b* initials as well as the reinforcement of the onset f0 lowering characteristic of the P_B onsets.

It is also noted that MC tone *3a* and *3c* have an extremely high f0 offset (also found in MC *4a* and *4c, 4m* not analysed here). Synchronic aspiration does not appear to be the cause of this tone split; rather, it appears that the historical voiceless aspirated onsets first for some reason became voiced/breathy, and merged with MC voiced onsets, which may account for the appearance of 'aspiration-conditioned tone lowering' [28]. The prevoicing [11, 13, 18, 19, 20, 25] and breathiness in syllables with MC *-c* and *-b* may have induced f0 lowering of the tonal onsets on those syllables, and further lead to the tone splitting along with the ongoing "devoicing" process in Gan.

5. REFERENCES

[1] Boersma, P., Weenink, D. 2016. Praat: doing phonetics by computer. Version 6.0.19. Retrieved on 14 June, 2016. http://www.praat.org.

[2] Cao, J., Maddieson, I. 1992. An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, *20*(1), 77-92.

[3] Chao, Y. R. 1936. Types of plosives in Chinese. In *Proceedings of the 2nd international congress of phonetic sciences*, 106-110, UK: Cambridge University Press.

[4] Chao. Y.R., Ting. S., Yang, S. F., Wu, T. S., Tung. T. H. 1948. *Report on a survey of the dialects of Hupeh*.

Taipei: Institute of History and Philology of Academia Sinica.
[5] Chen.C.Y. 1991. *A summary of Gan dialect (Gan Fangyan Gaiyao*). Nanchang: Jiangxi Educational Press.
[6] Chen.L. 2017. *The phonetic research of the 'voiced area of Gan (Gan Fangyan Zhuoyin Zoulang Yuyin Yanjiu)*. Beijing: Guangming Daily Press.
[7] Downing, L. J. 2009. On pitch lowering not linked to voicing: Nguni and Shona group depressors. *Language Sciences*, *31*(2-3), 179-198.
[8] Gordon, M., Ladefoged, P. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics*, *29*(4), 383-406.
[9] Hanson, H.M., 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *The Journal of the Acoustical Society of America*, *125*(1), 425-441.
[10] Henderson, E. J. ,1965. The topography of certain phonetic and morphological characteristics of South East Asian languages. *Lingua*, 15, 400-434.
[11] Hirata S.1997. The "lax" and "tense" difference in north and south Chinese (Hanyu Fangyan Yinji Leixing 'Song-Jin'de Nanbei Chayi). *Chinese Languages and Linguistics*, 4, 291-322.
[12] Ho, D. A. 1994. 'Zhuoshang Guiqu' and contemporary Chinese. *Shengyun LunCong.* 2, 267-292.
[13] Hombert, J. M., Ohala, J. J., Ewan, W. G. 1979. Phonetic explanations for the development of tones. *Language*, 37-58.
[14] Institute of Linguistics, CASS, Australian Acad. of the Humanities. 1987. *Language Atlas of China*. Hong Kong: Longman Group (Far East)
[15] Institute of Linguistics, CASS. 1981. *Fangyan Diaocha Zibiao (Revised Version).* Beijing: Commercial Press.
[16] Institute of Linguistics, CASS. 2012. Institution of Ethnology and Anthropology, CASS, Language Information Sciences Research Centre, Hong Kong City University. *Language Atlas of China (Second Edition): Chinese Dialects.* Beijing: Commercial Press.
[17] Karlgren, B., Chao,Y.R, Lo, C.P., Li,F.K. 1994. *Études sur la phonologie chinoise*. *(Zhongguo Yinyunxue Yanjiu).* Beijing: The Commecial Press.
[18] Kirby, J. P., Ladd, D. R. 2015. Stop voicing and F0 perturbations: Evidence from French and Italian. In *Proceedings of the 18th International Congress of Phonetic Sciences,* Glasgow.
[19] Kirby, J. P., Ladd, D. R. 2016. Effects of obstruent voicing on vowel F 0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America*, *140*(4), 2400-2411.
[20] Li, F.K. 1977. *A Handbook of Comparative Tai.* Honolulu: The University Press of Hawaii, 15.
[21] Li, R.L., Chang, S. H. 1992. *A report on a survey of the Kejia and Gan dialects*. Fujian: Xiamen University Publishing House.
[22] Liu, L. X. 1999. *A comparative research of Kejia and Gan dialects (Ke Gan Fangyan Bijiao Yanjiu)*. Beijing: China Social Science Press.
[23] Lu, J. F. 2016. *A phonological study of the Chang-du group of Gan dialects*. Doctoral Dissertation, Hunan Normal University.
[24] Lu, J. F., Liu, L.X. 2013. Internal phonological difference of Gan dialects in Xiushui. *Journal of Longyan University,3,*27-36.
[25] Matisoff, J. A. 1973. Tonogenesis in Southeast Asia. *Consonant Types and Tone*, 71-95.
[26] Mazaudon, M. 1977. Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman Area*, 3 (2), 1-123.
[27] Rose, P. 2002. Independent depressor and register effects in Wu dialect tonology. *Journal of Chinese Linguistics*, *30*(1), 39-81.
[28] Sagart, L. 1984. How did the aspirated stops become voiced. *Computational Analysis of Asian and African Languages*, *22*, 87-94.
[29] Shue, Y.L., Keating.P., Vicenik, C., Yu, K. 2011. VoiceSauce: A program for voice analysis, Proceedings of the ICPhS XVII, 1846-1849. VoiceSauce v 1.31. Retrieved on 31 May 2017. http://www.phonetics.ucla.edu/voicesauce/index.html
[30] Steed, W. 2012. *Qingtian Wu lexical tone sandhi: Voiceless depressors and allotones*. Munich: LINCOM.
[31] Strazny, P. 2003. Depression in Zulu. *The phonological spectrum: Segmental structure*,1,223.
[32] Sun,Y.Z. 2007. *The phonetic research of the 'Gan' dialect in Jiangxi Province*. Beijing: Language and Culture Press.
[33] van Rij. J., Wieling. M., Baayen. R., van Rijn. H. 2017. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.3. https://CRAN.R-project.org/package=itsadug
[34] Wan. B. 2009. *The research on the initial historical strata of Gan*. Beijing: Commercial Press.
[35] Wang, L.N. 2010. Phenomena of voiceless aspiration initial turning into voiced and tonal distribution of aspiration in Gan dialect. *Studies in Languages and Linguistics,* (3), 56-61.
[36] Wood, S.N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association. 99:673-686. https://CRAN.R-project.org/package=mgcv
[37] Xie, L.W. 2006. Re-classification/Re-distribution of Gan group. *Dialect*, *3*, 267-217
[38] Xu, Y. 2013. Tone-split by aspiration in Chinese dialects. *Studies of the Chinese Language*, (3), 265-272.
[39] Yan, S. 1986. Re-classification/Re-distribution of Gan group. *Dialect*, 1,19-38.
[40] Yang, S. F. 1971. The tone classes of the Kiangsi dialect. *Bulletin of the Institute of History and Philology Academia Sinica, Volume 43, Part 3*. Taipei: Institute of History and Philology of Academia Sinica
[41] Yang, S. F. 1974. *Report on a survey of the dialects of Hunan*. Taipei: Institute of History and Philology of Academia Sinica
[42] Zhu, X. S., Rose, P. 1998. Tonal complexity as a dialectal feature: 25 different citation tones from four Zhejiang Wu dialects. In *Proc. Fifth International Conference on Spoken Language Processing,* Sydney.