Possible explanation of Chinese misidentified tones

Guohua Hu
Department of Languages and Literatures, Gothenburg University, Gothenburg, Sweden

Abstract

Even though F_0 plays an essential role in tone perception other facts (e.g. different consonant types, vowel quality, see Hombert, 1978) also effect tone perception. This article tries to find some evidence of influence between initial consonants and misidentified tone patterns in terms of the acquisition of Chinese tones. The main results show that voiced sounds, e.g. [l], effect that (T)one 2 (high) is misinterpreted as T3 (low), and that aspirated stops, e.g. [p^h], cause that T3 (low) is misidentified as T2 (high).

Introduction

Stops are universal phenomena. Different languages use various stop systems. Chinese, on one side, has only a voiceless stop system; there the features [±aspirated] play an essential role and voiced sounds include approximants e.g. [l], [w] and the nasals (Lin, 2007:45–65).

On the other side, [±voiced] are distinguishing features in Swedish, see also (Hu, 2012). Apart from the segments, suprasegments in languages give another view. This article only focuses on how Swedes have perceived Chinese tones so a presentation of Swedish suprasegments is irrelevant here.

Chinese is a tone language, which has four lexical tones (T1, T2, T3, and T4, see Figure 1) and one neutral tone (or toneless), e.g. 吗 ma *question mark* (eh?). A classical example often given in literatures is 妈 mā *mother*, 麻 má *hemp*, 马 mǎ *horse*, and 骂 mà *scold*. The words distinguish their meanings by different tones.

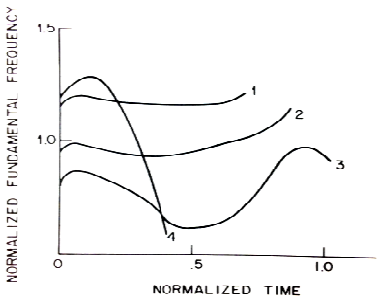

Figure 1: F_0 contours with four Chinese tones (from: Chuang et al., 1972)

Figure 1 shows the F_0 (fundamental frequency) contours of the four Chinese tones. In short, T1 is high and relatively level over most of the duration of the tone. T2 begins with a relative lower pitch compared to T1, and the onset of the rise occurs in the middle portion and ends almost as high as T1. T3 contour displays the lowest region of the F_0 range, although extending at least to the midpoint of the range by the offset. F_0 in the beginning of both T2 and T3 are quiet close to each other. T4 starts high and robustly falls to the bottom during its duration. Phonologically, T1 is high level, T2 is high rising, T3 is dip low and T4 is low falling.

Earlier studies (e.g. Klatt, 1973; Zee, 1985) have paid their attention to the pitch of the vowel (F_0). Many scholars (e.g. Shen & Lin, 1991; Chuang et al., 1972; Kiriloff, 1969) have conducted perception tests outgoing from different native speakers since different languages vary in their pitch patterns and functions. Their main results have shown that T2 is often misidentified as T3 and vice versa, but T3 is not so frequently misidentified as a T2.

The explanation for a misconception of this tone pair is that "neither the falling and rising contour nor the position of the dip point in tone-3 alone can be the perceptual cues to discriminate tone-3 from tone-2" (Chuang et al., 1972:299), see also *Figure 1*.

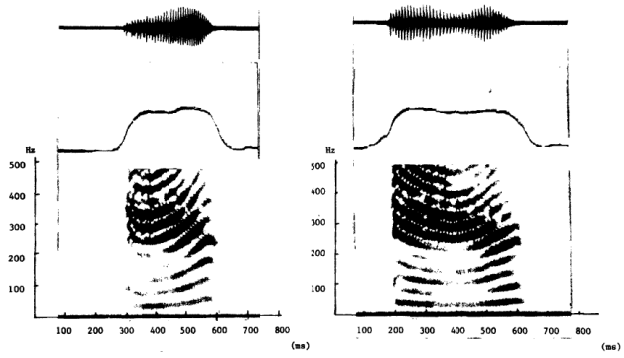

Figure 2: Turning points between T2 and T3 (from: Shen & Lin, 1991).

Shen and Lin (1991) have reported that the time of the turning points is the perceptual cue that differentiates T2 from T3. Moreover, a shift in vowel duration has been shown to affect the perception of T2 and T3, see *Figure 2*.

Significant is that they have chosen test words only with voiced initial consonants e.g. lóu *a storied building* and lǒu *basket*, máo *hair* and mǎo *riveting,* and ní *mud* and nǐ *you*, etc. They have, however, not explained why they used test words only with a voiced initial, so the question is if this is enough for stating that the T2 is confused T3. We lack an explanation for the misperception between T3 and T2.

In a study of tone perception of adult Swedes Hu and Lindh (2010) have shown the similar result, i.e. that T2 and T3 and vice versa are most frequently confused. However, they haven't illustrated *why* this tone conception occurs.

The hypothesis of this article is that the initial consonants affect the tone perception since the feedback from the informants, after the experiment, has revealed that they experienced that the consonants preceding the vowels disturbed the tone perception.

Furthermore, Hombert (1978) has claimed that F_0 perturbation of initial stops does disturb the tone perception. However, he has only analyzed the acoustic features between English and French stops. There was no evidence how English speakers perceived French stops and vice versa.

This article pays its attention to the tone perception based on Hombert. Firstly the perception experiment will be presented. Secondly, the results are illustrated. Next, the initial stops and tone confusions patterns are discussed.

The current study

Twenty-five different unhandled disyllabic words were selected from a textbook for the beginners of Chinese at Gothenburg University. No syllable has a nasal initial. The test words are disyllabic ($\sigma_1\sigma_2$) including 15 (of the possible 19) tone combinations. The test word 劳动 láodòng [^{35}lau^{51}tɔŋ] 'work', to take an example, consists of two syllables. The initial consonant of σ_1 is a lateral approximant (voiced), the initial consonant of σ_2 one is an unaspirated stop. The tone combination of this word is T2+T4.

One male and one female native speaker of Chinese pronounced the words in isolation. Each speaker repeated the words twice with a pause of 1 second in between and there is 2 seconds pause before the following new word. The audio was presented in high quality headphones in the student language lab of the university.

Listening tests have been conducted every year between 2007 and 2013. Each year there have been 25–40 participants (bilingual Chinese immigrants are excluded from the results). The actual group consisted of 18 students who were admitted in the autumn semester of 2007. The subjects had the possibility to listen to the words as many times as they needed for noting correct transcription and tones. The listening test was also an exam, which pushed the participants to perform well.

The students almost correctly transcribed the sounds by using Romanization (Pinyin system). The wrong spellings are so few that they are ignored in the present analysis that focuses only on tone identifications.

Results

The 18 participants had to identify 25 disyllabic words. The total responses were 900 (= 18 × 25 × 2). 278 tones were misidentified (30.88%). The matrix of tone responses is shown in *Table 1*, where **bold** indicates the correct answer.

Table 1: Matrix of tone responses (%)

	Response																			
	Type I				*Type II*					*Type III*										
	σ_1				Σ_2					σ_1				σ_2						
Stimulus (T)	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*0*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*0*		
1	**75**	13	3	9	**82**	7	0	7	4	**75**	11	8	6	**82**	9	0	9	0		
2	16	**60**	11	13	2	**57**	33	5	3	5	**60**	24	11	0	**58**	33	9	0		
3	10	33	**52**	5	1	23	**69**	5	2	13	13	**54**	20	7	19	**69**	5	0		
4	3	17	1	**79**	3	10	3	**78**	6	5	11	5	**79**	7	9	3	**78**	3		
0					0	0	0	47	**53**					0	0	0	47	**53**		

Depending upon where the misidentified tones of the syllables in the words occur they were divided as follows (note that T_0 never can be distributed in σ_1). The highest percentages of misidentifications are:

- Type I, σ_1: T3→T2 (33%)
- Type II, σ_2: T2→T3 (33%)
- Type III: the tones of both σ_1 (T2→T3, 24%) and σ_2 (T2→T3, 33%) are wrongly perceived

Table 2 below shows the influence of the initial consonants on the incorrect tone responses.

Table 1: Correlation between misidentified tones and the initial consonants of three types (%)

	Sounds	Features	Type I	Type II	Type III	
			σ_1	σ_2	σ_1	σ_2
Voiceless	Stop	Unaspirated	12	17	8	29
		Aspirated	15	13	16	12
	Fricative		31	13	35	22
	Affricate	Unaspirated	6	6	6	14
		Aspirated	7	11	4	2
Voiced	Central Approximant		4	33	4	20
	Lateral Approximant		24	7	27	0
	∑		100	100	100	100

- Type I: The fricatives can be suspected to play a role for tone confusions (31%). Thereafter follow the stops (unaspirated 12% and aspirated 15%). The stops have a stronger connection with misidentified tones than the affricates. Also the lateral seems to have an impact on the tone perception (24%).
- Type II: One third of the central approximant affricates are seen together with tone confusions. The stops have a similar tendency as in Type I. However, fricatives decrease to 13% (31% in Type I). Aspirated affricates tie up stronger (11%) with tone misperception than the unaspirated (6%).
- Type III: If both syllables have a fricative as their initial consonant it looks plausible that they influence the tone perception (35% and 22%). The initial stops of the σ_2 are possibly involved with the wrong tone identifications (29%). Unaspirated affricates (14%) have a substantial relation to misidentified tones. Concerning voiced initials, the results show that the central approximants have a strong association with the wrong tone responses of σ_2 (20%), whereas the lateral (27%) seems to have its impact on σ_1.

A plotting schedule, showing the patterns for tone confusions related to the initial consonants was developed. The space of this article is too small for reproducing it here, but it can be requested from the author. Since the data was very little, we pick only the results according to the analyses above. A suspicion was found that only some initial consonants influence the tone confusion patterns T2→T3 and T3→T2.

- Type I, σ_1: In 20% of the cases the aspirated initial stops seem to have a connection with the tone confusion T3→T2. The initial lateral approximant has a close link to T2→T3 (12%).
- Type II, σ_2: The aspirated initial affricates seem to be involved when the tone confusion T2→T3 occurs in 38% of the cases. Unaspirated stops might have a relationship to the wrong tone pattern T3→T2 (15%), it is however not as frequent as in the case of the aspirated affricates. In 25% a central approximant as initial ties up with this same tone confusion. The same kind of initial even shows a connection with T3→T2 (15%)

- Type III: When a lateral approximant is the initial of σ_1 there often occurs a T2→T3 confusion, 26%. Concerning a possible pattern for the T3→T2 confusion the initial consonants are so scattered that no strong connections can be suspected.

Conclusion

It may be concluded that the initial consonants do affect tone perception. Aspirated initials seem to be linked to cause that T3 is misinterpreted as a T2. Unaspirated stops and approximants show a connection of the wrong perceived tone pattern T2 as a T3. It might be added that this pattern has repeated itself in the tests of 2008–2013 (although the numbers in this article refer to the test 2007).

The department of Chinese has started a perception coaching system. It includes all Chinese disyllabic phonotactic and tonotaxical patterns in order to observe closely the relationship between initial consonants and misidentified tone patterns. Meanwhile, the acoustic data of F_0 disturbance by both Swedish and Chinese consonants preceding vowels should be collected.

References

Chuang, C. K., S. Hiki, T. Sone & T. & Nimura 1972. The Acoustical features and perceptual cues of the four tones of standard colloquial Chinese *Proceedings of the 7th International Congress of Acoustics,* 3297−300.

Hombert, J.-M. 1978. Consonant Types, Vowel Quality and Tone. In: V. A. Fromkin (ed.), *Tone − A Linguistic Survey*, 77–111.

Lin, Y.-H. 2007. The Sounds of Chinese. Cambridge: Cambridge University Press.

Hu, G. 2012. Chinese and Swedish Stops in Contrast. Paper presented at the *Proceedings of Fonetik 2012 Gothenburg Department of Philosophy, Linguistics and Theory of Science*, 77–80.

Hu, G. & J. Lindh. 2010. Perceptual mistakes of Chinese tones in 2-syllable words by Swedish listeners (Abstract). Paper presented at the *Proceedings of the Fourth European Conference on Tone and Intonation* (TIE4). Stockholm, 121.

Kiriloff, C. 1969. On the auditory perception of tones in Mandarin. *Phonetica* 20:63–67.

Klatt, D. 1973. Discrimination of fundamental frequency contours in synthetic speech duplicaitons for models of pitch perception *Journal of the Acoustical Society of America* 53, 8–16.

Shen, X. S. & M. Lin. 1991. A Perceptual Study of mandarin Tones 2 and 3. *Language and Speech* 34:145–156.

Zee, E. (1985). Sound change in syllable final nasal consonants in Chinese *Journal of Chinese Linguistics* 13(1):291–330.

Proceedings of Fonetik 2013

The XXVIth Annual Phonetics Meeting 12–13 June 2013, Linköping University Linköping, Sweden

Studies in Language and Culture
no. 21

Robert Eklund, editor

Linköping University

Conference website: www.liu.se/ikk/fonetik2013
Proceedings also available at: http://roberteklund.info/conferences/fonetik2013

Cover design and photographs by Robert Eklund
Photo of Claes-Christian Elert taken by Eva Strangert on the occasion of his 80th birthday

Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference
held at Linköping University, 12–13 June 2013
Studies in Language and Culture, no. 21
Editor: Robert Eklund
Department of Culture and Communication
Linköping University
SE-581 83 Linköping, Sweden

ISBN 978-91-7519-582-7
eISBN 978-91-7519-579-7
ISSN 1403-2570

Printed by LiU-Tryck, Linköping, Sweden, 2013