Tonal Confusions in Cantonese at Different Signal-to-noise Ratios

Cheung Pak-pei

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, April 30, 1992.

ABSTRACT

The effects of noise and tone groups on tonal perception and confusions in Cantonese are examined in this study Two tone 1 variants and five other tones, tone 2 to tone 6, of the monosyllable /fu/ were presented to 35 normal subjects monaurally between the range of +20 dB to -20 dB signal-to-noise ratios (*S/N*s). Results show that tone identification accuracy diminishes when the *S/N* decreases. When the S/N is below -15 dB, performance drastically decreases The relative difficulty in perceiving level tones, rising tones and falling tones in noise are reported. The tonal confusion patterns parallel with the findings in literature. Finally the results will be discussed with reference to the fundamental frequency shapes and perceptual distance of tones in Cantonese.

INTRODUCTION

The effects of noise on speech communications have been studied for a long time. Miller et al (1951) studied the intelligibility of speech as a function of the size of vocabularies at different noise levels. Later Miller and Nicely, (1955), and Wang and Bilger (1973) investigated the effects of noise on the transmitted information of articulatory and phonological features of English consonants. However, the effects of noise in lexical tone perception in tonal languages, such as Cantonese, have been rarely explored.

Tones in Cantonese are lexical or phonemic that pitch variations contrast the meaning of a word (Pike, 1948). During the past several decades, the number and classification of tones have been a topic of interests among linguists and phoneticians. Modern instrumental investigations have now agreed that there are six contrastive tones in Cantonese (Fok, 1974). The present study was conducted to examine the effects of noise on lexical tone perception in Cantonese.

Noise

Noise interferes speech perception by masking the speech signal which results in reducing the contrast between signal and noise, and in turn, the speech intelligibility. That is, the higher the noise level, the more poorly we can comprehend the speech.

The relative intensity between signal (speech) and noise is specified as signal-to-noise ratio (*S/N*). This is defined as the difference in decibels between the signal and the noise when presented to the same ear (ears) simultaneously. For instance, when the speech is 10 dB weaker than the noise, the *S/N* is -10 dB; when they are both at the same intensity, the *S/N* is 0 dB, and so on.

In this study, white noise was used for masking the speech signal because of its resemblance to everyday noise (Zwicker & Fastl, 1990). White noise is also called broad-band noise, which has approximately equal energy in each frequency and covers a relatively broad range of frequencies up to 20 kHz.

Tones in Cantonese

Cantonese comprises both level and gliding (or contour) tones. Pike (1948) defined level tone as one in which the pitch of a syllable does not rise or fall during its production within the perceptual limits. Gliding tone is one in which the pitch of a syllable has a rise or fall or some combination of rise and fall.

Chao (1947) accounted that Cantonese has three main inflections (even, rising and going) with two levels of pitch (upper and lower). This constitutes a total of six tones: upper even, upper rising, upper going, lower even, lower rising, and lower going. Fok (1974) assigned tone numbers from one to six to each lexical tone. The "entering tones" which occur only with syllables ending in /-p/, /-t/, or /-k/ are regarded as the short forms of tones 1, 3 and 6 respectively, with tone numbers assigned as seven to nine (table 1).

In this study, Fok's tone number, Chao's tone values and Kao's (1971) qualitative description are adopted because they are widely used.

Table 1. Tones in Cantonese.

Fok's Tone	Kao's Qualitative Description	Pitch Value	Word	Root	Meaning
1	high falling/level	55/53	'夫'	/fu_1/	'husband'
2	high rising	35	'苦'	/fu_2/	'bitter'
3	mid level	33	'富'	/fu_3/	'wealthy'
4	low falling	21	'符'	/fu_4/	'symbol'
5	low rising	23	'婦'	/fu_5/	'woman'
6	low level	22	'負'	/fu_6/	'negative'
7	high entering	5	'福'	/fuk_7/	'bliss'
8	mid entering	3	'✕'	/fuk_8/	no character
9	low entering	2	'服'	/fuk_9/	'to obey'

Tse (1973) reported that the two variants of tone 1 are probably two variants of the upper even tone. These two forms, tone 1 (55) and tone 1 (53), have distinct pitch forms and they are distributed rather evenly throughout the corpus both in isolation and in context. Whether these two tones are perceived differently is investigated in this study.

Tonal perception

Tonal perception in Cantonese was extensively studied by Fok (1974). She reported that fundamental frequency (F_o) variation is the prime carrier of tonal information and confusions often occurred in tones with similar F_o shapes. Subjects often confused between tone 2 and tone 5; tone 3 and tone 6; and tone 4 and tone 6. They also labelled tone 1 (53) as tone 3 or tone 6. Falling tones and rising tones were found to be identified more accurately.

Gandour (1981) used a multidimensional scaling model to reanalyze Fok's data to explore the perceptual dimensions in Cantonese tones. The results revealed three dimensions: contour, direction and height. Contour differentiates level tones from gliding tones. Direction distinguishes the falling tones from the rising tones. Height corresponds to the average fundamental frequency. Further analysis suggested that contour and direction dimensions (the F_o shapes) are more salient than the height dimension.

In Yiu's study (1989) on tone disruption in Cantonese aphasics, tonal confusion patterns parallel with Fok's (1974) finding. The data also suggested that both contour and height dimensions were disrupted in aphasics; but contour appeared to be more resistant to tonal impairment.

Whitfield and Evans (1965) reported that our auditory system appears to be more efficient at detecting dynamic changes in F_o than differences in levels between two F_o signals of the same magnitude. Additionally, Anderson (1978) agreed that gliding tones are more perceptually salient than level ones. Klatt (1973) concluded from experiments using speech-like synthetic stimuli that it is far more difficult to detect an average frequency difference or a difference in rate-of-change of F_o if the change occurs in a F_o with falling slope. However, Ching (1984) reported that the falling tone, tone 4, was best identified by the children in her experiment. From these findings, we may hypothesize that gliding tones are perceived more accurately than level tones. In gliding tones, it is difficult to predict the relative ease of perceiving between falling tones and rising tones.

Speech perception in noise

Miller and Nicely (1955), and Wang and Bilger (1973) investigated the effects of noise on the transmitted information of articulatory and phonological features of consonants in English. They reported that both nasal and voice features are well perceived in noise.

Objectives of the study

The aims of this study were to investigate:

(i) the effect of white noise on tone identification in Cantonese;

(ii) whether any tones are more difficult to perceive than the others, such as level tones being more difficult than gliding tones;

(iii) whether tone 1 (55) is more difficult to perceive than tone 1 (53); and

(iv) whether the tonal confusion patterns in white noise parallel with the previous findings.

METHOD

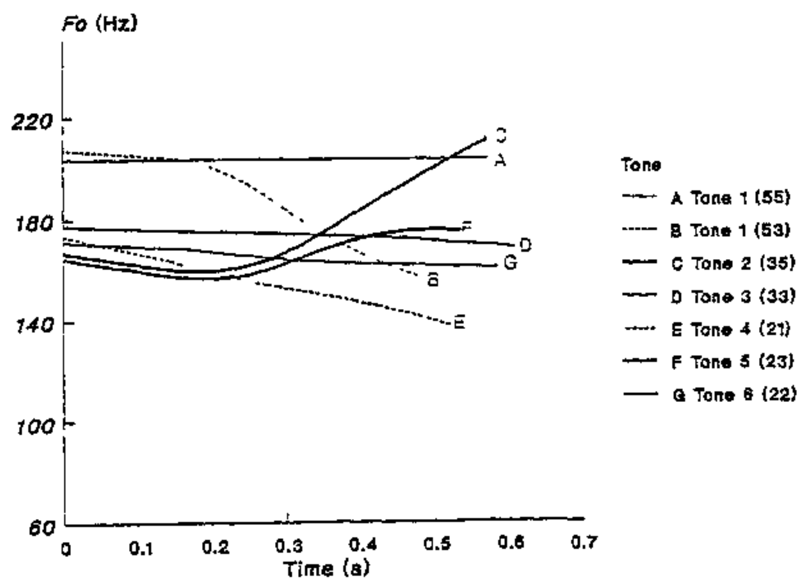
Subjects

Thirty-five Cantonese-speaking subjects (twenty-two males and thirteen females) with age range from nineteen to twenty-six (mean: 23.57, standard deviation: ±0.982) participated in the tone identification experiment. All of them passed a screening procedure, which consisted of an audiological screening and a tone identification screening. Their Cantonese was judged not to have any characteristics of non-Cantonese accent.

Stimuli

Seven tones ([55, 53], 35, 33, 21, 23, 22) of the monosyllable /fu/ were produced by an adult male native speaker of Cantonese. The fundamental frequency tracings of these seven tones, which were analyzed by KAY Visi-Pitch 6087DS, are illustrated in figure 1. Each stimuli were recorded with the same intensity by peaking at the same point of zero on the volume unit (VU) meter on Sony TC-D5M cassette recorder in a sound-treated room. These seven recorded tones were then duplicated. Seven trial items, which served as screening procedure and trials, will be duplicated first. Then seven blocks of 14 test items, two repetitions of each tone, were recorded in random order within each block. Thus, a total of 98 items for seven *S/N* were prepared for each subject. At the beginning of the test tape, a pure tone (1 kHz), peaking at zero on the VU meter, generated from REVOX C279 Mixing Console built-in 1 kHz tone generator was recorded for calibration prior to the listening experiment.

Figure 1. Fundamental frequency tracings of the seven tones.



Pilot study

A pilot study was conducted to determine the *S/N*s of near chance and near perfect recognition, and hence, the intervals of the *S/N* to be used in the listening experiment. Three subjects who are native Cantonese speakers, aged 21, 22 and 24, were involved. Finally the following seven *S/N*s were determined: +20 dB, +10 dB, 0 dB, -5 dB, -10 dB, -15 dB and -20 dB.

Screening Procedure

A screening procedure was prepared to select subjects who (i) had normal hearing ability and (ii) could identified all the seven Cantonese tones correctly in the absence of white noise.

(i) In the audiological screening, a 20 dB HL pure tone at 250 Hz, 500 Hz, 1 kHz, 2 kHz and 4 kHz were tested in each ear (adapted from American Speech-Language-Hearing Association, 1990) from the Madsen OB822 Clinical Audiometer in a sound-treated Audiology laboratory. Those who failed to respond to any one of the ten tones were excluded in the listening experiment.

ii) Those who passed the audiological screening then received the tone identification screening. Only those who identified all of the seven Cantonese tones correctly proceeded to the tone listening experiment. This was done to exclude people who had tonal confusion in the absence of noise and to guard against the possibility that some of the characters might have variant pronunciations in some people.

Listening Procedure

Each subject, who passed the hearing screening, was tested individually in the Audiology laboratory. The test tape were played on REVOX B710 MK II Cassette Recorder, which provided output to the Madsen OB822 Clinical Audiometer The speech signals were calibrated by the 1 kHz pure tone and were presented at a fixed sound pressure level of 60.5 dB SPL. This was measured by the Bruel & Kjaer sound level meter Type 2235 and Brüel & Kjaer artificial ear Type 4125.

The white noise generated from the audiometer together with the lexical tones stimuli were presented monaurally through TDH-39P headphones mounted in MX41/AR cushions. The possible effect of ear advantage was counterbalanced between left and right ear.

Subjects were asked to identify each stimulus by circling the appropriate word on their response sheets. A forced-choice paradigm was used so that they were asked to guess if they missed any of the items. They listened to seven blocks of tone stimuli at different *S/N*s, in which the presentation order of the seven *S/N*s were counterbalanced by a 7x7 Latin square design. The entire experiment lasted approximately 30 minutes.

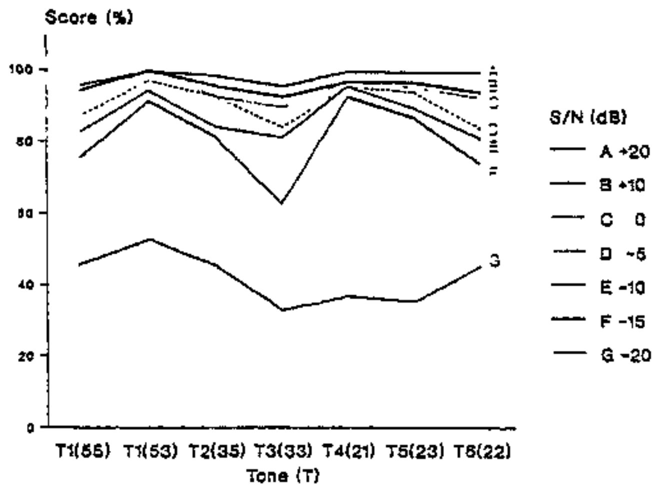
RESULTS

The effects of white noise and types of tones on tone identification in Cantonese were examined by a two way analysis of variance (ANOVA) with repeated measures. An error analysis was carried out to study the tonal confusion patterns.

Tone identification in noise

Referring to figure 2, the identification scores clustered in three groups between +10 and -15 dB *S/N*s. Falling tones appeared to be identified more accurately than rising tones, and rising tones were identified more accurately than level tones. Thus the seven Cantonese tones were collapsed into three groups for analysis according to their F_o shapes: level group of tone 1 (55), tone 3 (33) and tone 6 (22); rising group of tone 2 (35) and tone 5 (23); and falling group of tone 1 (53) and tone 4 (21). Three sets of their tone identification scores were obtained by averaging the scores within each group at each *S/N*.

Figure 2. Mean tone identification scores as a function each tone at different *S/N*s.



The tone identification scores were then subjected to a two-way analysis of variance with repeated measures, having seven levels of *S/N*s (+20 dB, +10 dB, 0 dB, -5 dB, -10 dB, -15 dB, or -20 dB) and three levels of tones (level, rising, or falling).

Table 2. Results of two-way analysis of variance with repeated measures on *S/N*s and tone groups.

Source of variation	*SS*	*df*	*MS*	*F*	*p*
Subjects *S/N*	9.43	34	0.28		
Tone	96.20	6	16.03	207.21	0.001
S/N x Tone	4.08	2	2.04	9.31	0.001
Error	2.03	12	0.17	2.54	0.003
	57.85	680	0.09		
Totals	169.59	734	18.61		

All main effects and interaction effect were found to be statistically significant at 95% confidence level, $p < 0.05$ (table 2). The mean tone identification scores were summarized in table 3.

Table 3. Mean tone identification scores as a function of *S/N*s and tone groups

Tone	*S/N* (dB) +20 (1)	+10 (2)	0 (3)	-5 (4)	-10 (5)	-15 (6)	-20 (7)	$\mathbf{M_T}$
Level(L)	1.943	1.876	1.829	1.705	1.638	1.419	0.829	1.606
Rising(R)	1.986	1.929	1.886	1.871	1.743	1.686	0.814	1.702
Falling(F)	2.000	1.971	1.971	1.929	1.900	1.843	0.900	1.788
$\mathbf{M_{S/N}}$	1.976	1.925	1.895	1.835	1.760	1.649	0.848	

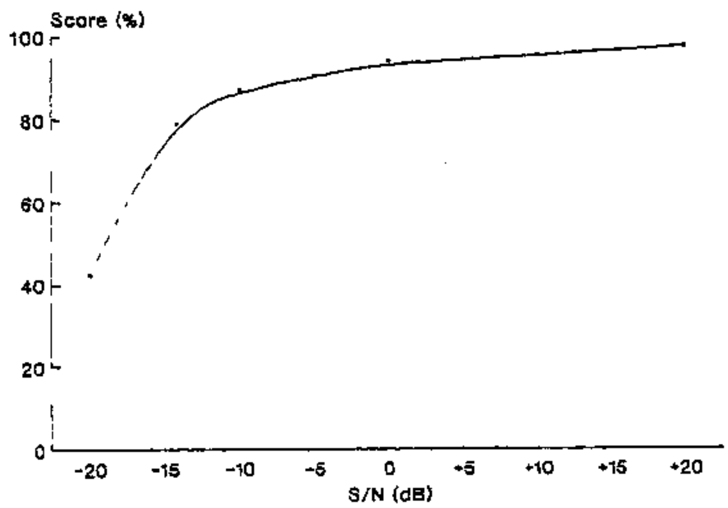
Key: $M_{S/N}$ = mean of the main effect due to S/N (M_1 ... M_7)
M_T = mean of the main effect due to tone group (M_L, M_R & M_F)

The main effect of *S/N* was such that decrease in *S/N* produced diminishing change of tone identification scores, $F(6,680) = 207.21$, $p < 0.001$. A Tukey's honest significant difference (*HSD*) test with the critical difference ($CD = 0.180$) revealed that tone identification was the worst at -20 dB *S/N* ($M_7 = 0.848$) and its mean significantly differed from the rest. The mean score at -15 dB *S/N* ($M_6 = 1.649$) was not significantly different from its precedent *S/N*s until it reached 0 dB *S/N* ($M_3 = 1.895$). The mean score at -10 dB *S/N* ($M_5 = 1.760$)

did not differ significantly from its precedent *S/N*s until it increased up to +20 dB *S/N* (M_I = 1.976).

The tone identification scores in percentage averaged across seven Cantonese tones at different *S/N*s were presented in figure 3. The tone identification accuracy decreased gently between +20 dB and -15 dB *S/N*s. At -20 dB *S/N*, there was a sudden drop of accuracy over 40%.

Figure 3. Mean identification scores averaged across seven tones as a function of *S/N*s. Each point represents 490 trials.



The nature of the main effect of tone groups, F(2, 680) = 9 31, $p < 0.001$ was determined by the Tukey *HSD* test (*CD* = 0.148). Results showed that the falling group (M_F = 1.788) was identified better than both the rising group (M_R = 1.702) and the level group (M_L = 1.606), but that the latter two groups did not significantly differ.

The interaction between *S/N*s and tone groups, $F(12,680) = 2.54$, $p < 0.005$, was analyzed using simple main effects analysis, followed by the Tukey *HSD* test and graphical analysis. The relevant means were presented in table 3. For the simple main effect analysis, seven one-way repeated measures ANOVA was performed at each level of the *S/N*s. Significant main effects were found at 0 dB, -5 dB, -10 dB and -15 dB *S/N*s, $p < 0.05$. The *p* values were presented in table 4. Tukey *HSD* test and graphical analysis (figure 4) on these four *S/N*s levels were used to study the nature of the interaction. The test ($CD = 0.163$) failed to show the nature of interaction occurred at 0 dB *S/N*.

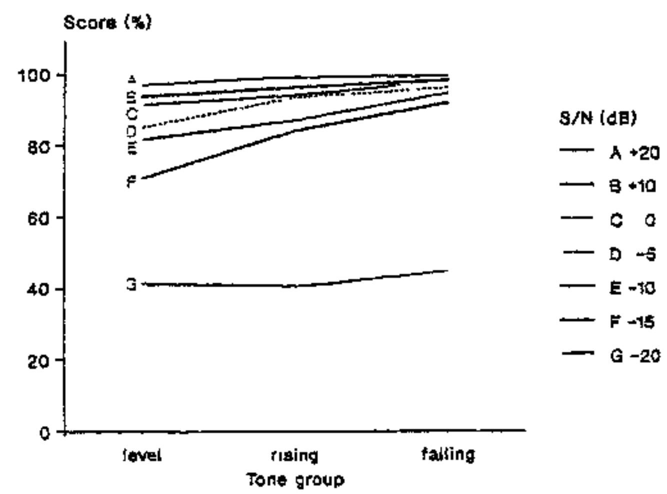
Table 4. The *p* values in the simple main effects analysis

	S/N (dB)						
	+20	+10	0	-5	-10	-15	-20
p	0.090	0.096	0.011	0.002	0.018	0.001	0.672

At -5 dB and -15 dB *S/N*s, both falling ($X_{F4} = 1.929$ and $X_{F6} = 1.843$) and rising groups ($X_{R4} = 1.871$ and $X_{R6} = 1.686$) produced significantly more change in identification score than the level group ($X_{L4} = 1.705$ and $X_{L6} = 1.419$). However, the means of falling group and rising group did not significantly differ. At -10 dB *S/N*, only falling group ($X_{F5} = 1.900$) produced significantly more reduction in the score than the level group ($X_{L5} = 1.638$). The rising group (X_{R5}

= 1.743) did not significantly differ with either of them. The tone identification scores in percentage as a function of tone groups at each *S/N*s (figure 4), showed concordant pattern with the results from the Tukey test. Between the -5 dB and -15 dB *S/N*s, the level tones produced a catalytic effect on impairing tone identification with noise, reducing more score than the falling and rising tones did.

Figure 4. Mean tone identification scores of three tone groups at each *S/N*s.



Tonal confusion pattern

Eight confusion matrices, table 5 to table 12, were constructed to study patterns of tonal confusion in Cantonese. Whether the number of incorrect responses indicated significant tonal confusion or not was determined by the following equation (Fok, 1974):

$$\frac{N-n}{5}+1.96\sqrt{\frac{N-n}{5}\left(1-\frac{N-n}{5N}\right)}$$

where n = number of correct identifications

N = number of trials = 70

The stimulus tones were listed in the top second row and responses in the column on the left in each matrix. Significant confusion processes were denoted by (*) in the confusion matrices.

Table 5. Tone confusion matrix at +20 dB S/N.

				Stimulus			
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	67	70					
2			69				
3	3*			67			
4					70		
5			1			70	
6				3*			70
% correct	95.71	100.00	98.75	95.71	100.00	100.00	100.00

Table 6. Tone confusion matrix at +10 dB S/N.

				Stimulus			
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	66	70					
2			67		1	2*	
3	4*			65			3*
4					68		1
5			3*			68	
6				5*	1		66
% correct	94.29	100.00	95.71	92.86	97.14	97.14	94.29

Table 7. Tone confusion matrix at 0 dB S/N.

	Stimulus						
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	64	70					
2			65			2	
3	6*			63		1	3*
4					68		2
5			5*			67	
6				7*	2		65
% correct	91.43	100.00	92.86	90.00	97.14	95.71	92.86

Table 8. Tone confusion matrix at -5 dB S/N.

	Stimulus						
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	61	68			1		
2		1	65			4*	
3	9*	1		59			6*
4					67		5
5			5*			66	
6				11*	2		59
% correct	87.14	97.14	92.86	84.29	95.71	94.29	84.29

Table 9. Tone confusion matrix at -10 dB S/N.

	Stimulus						
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	58	66		1			
2		1	59		1	6*	1
3	9*	2		57			6*
4	1		1		67		6*
5		1	10*			63	
6	2			12*	2	1	57
% correct	82.86	94.29	84.29	81.43	95.71	90.00	81.43

Table 10. Tone confusion matrix at -15 dB S/N.

				Stimulus			
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	53	64				1	
2		1	57	1		8*	1
3	14*	3		44	1		4
4		1		1	65		11*
5	1	1	13*	2		61	2
6	2			22*	4*		52
% correct	75.71	91.43	81.43	62.86	92.86	87.14	74.29

Table 11. Tone confusion matrix at -20 dB S/N.

				Stimulus			
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	32	37	7	7	8	10	6
2	4	10	32	9	6	13*	4
3	19*	8	6	23	13	7	12
4	1	2	3	5	26	8	10
5	5	10	16*	9	8	25	6
6	9	3	6	17*	9	7	32
% correct	45.71	52.86	45.71	32.86	37.14	35.71	45.71

Table 12. Tone confusion matrix with data pooled across the seven S/N ratios.

				Stimulus			
Response	1 (55)	1 (53)	2 (35)	3 (33)	4 (21)	5 (23)	6 (22)
1	401	445	7	8	9	11	6
2	4	13	414	10	8	35*	6
3	64*	14	6	378	14	8	35*
4	2	3	4	5	431	8	35*
5	6	12	53*	11	8	420	8
6	13	3	6	77*	20*	8	401
% correct	81.84	90.82	84 50	77.14	87.96	85.71	81.84

Between +20 dB and -15 dB *S/N*s, the earliest significant confusions occurred in level tones of tone 1 (55) and tone 3 at +20 dB S/N, then in rising tones of tone 2 and tone 5 at +10 dB S/N. Significant tonal process in falling tone of tone 4 did not appear until it reached -5 dB S/N.

Between +20 dB and -15 dB *S/N*s, error responses clustered at cells where significant confusions were found. The identification performance decreased gently as S/N decreased. At -20 dB *S/N*, incorrect responses fell into all error cells and the identification accuracy decreased abruptly.

So as to examine the overall pattern of tonal confusion, all the data in each cells across each levels of speech to noise ratio were summed (table 12). Significant and consistent confusions were found in all tones except tone 1 (53). Tone 1 (55) was misidentified as tone 3, [1->(55)3]. Tone 2 and tone 5 was confused with each other, [2<->5]. Tone 6 was misidentified as either tone 3 or tone 4, [6->3/4] Other confusion processes occurred in tone 3 [3<->6] and tone 4 [4<->6]. No significant tonal confusion occurred in tone 1 (53).

Table 13 summarizes the tonal confusion patterns in terms of tone groups.

Table 13. Summary table of tonal confusion pooled across each *S/N*s.

Tone group	Significant confusion	Number of occurrence	Proportion (%)
Level	#tone 1 (55) --> tone 3	64	20.06
	#tone 3 (33) --> tone 6	77	24.14
	#tone 6 (22) --> tone 3	35	10.97
	tone 6 (22) --> tone 4	35	10.97
Rising	#tone 2 (35) --> tone 5	53	16.61
	#tone 5 (23) --> tone 2	35	10.97
Falling	tone 4 (21) --> tone 6	20	6.27

In the level group, four significant tonal confusions were found, three level-level confusions and one level-falling confusion. Tone 3 and tone 6 were confused with each other. Tone 1 (55) was only significantly confused with tone 3, but not with tone 6. The total proportions (66.14%) and the number of significant confusions were higher than the other two groups.

In the rising group, two rising-rising confusions were evident between tone 2 and tone 5. The total proportion (27.58%) and the number of significant confusions were lower than level group, but higher than the falling group.

In the falling group, no falling-falling confusion was noticed. Instead, one falling-level confusion of tone 4 to tone 6 occurred. For tone 1 (53), no significant confusion was found at all. The total proportion (6.27%) and the number of significant confusion was the lowest.

DISCUSSION

Tonal confusion patterns

The present finding on tonal confusion (c.f. table 13) agrees with the previous perceptual studies (Fok. 1974; and Yiu, 1989) There are seven significant confusion processes which occur consistently across each *S/N*.

In examining the tonal confusion patterns, the similarity of F_o shapes and the perceptual space of the fundamental frequencies between two tones along production, are found to be applicable.

Referring to figure 1, the F_o tracings in the initial portions of high rising tone 2 and low rising tone 5 are very close to each other and they do not diverge until in the halfway of the production. The discrimination of these two tones relies heavily on differentiating the perceptual distance at the rising end portions and/or the offset F_o. Therefore, the [2<->5] confusion processes appear.

The F_o tracings of mid level tone 3 and low level tone 6 run parallel along the entire production and are very close to each other. Hence, level-level confusions of [3<->6] occur. In spite of the wide perceptual space between tone 1 (55) and tone 3, [1(53)->3] process occurred frequently. This supports that level tones are relatively difficult to perceive (Whitfield & Evans, 1965) and the height dimension is perceptually less salient (Anderson, 1974).

Conversely, [3->1(53)] process was not found. This can be explained by the fact that the perceptual distance between tone 3 and tone 6 is narrower than that between tone 3 and tone 1 (55). When tone 3 is heard, people are more likely to misidentify it as the nearby low level tone 6 rather than high level tone 1 (55).

No falling-falling confusion was found between tone 1 (53) and tone 4 may be due to different F_o falling shapes and their onset F_o portions being separated by very wide perceptual distance.

The only inter-group confusions, the [4<->6] processes, may be explained by the fact that the slope of F_o tracing of tone 4 is rather gentle as compared to other gliding tones. Additionally the perceptual space between them are close. Tone 4 appears to be perceptually similar to level tones. Although the perceptual space between tone 1 (55) and tone 3 is much wider than that between tone 6 and tone 4, much higher proportion of [1(55)->3] confusions occur than [6->4] process. This reflects that the direction and contour dimensions are more perceptually salient than the height dimension.

Furthermore, over 80% of the total confusion processes are within group (with # in table 13) and the rest are within group. Such high proportion of processes, which occurred in tones with similar F_o shapes, supports the importance of the direction and contour dimensions in the perception of lexical tones (Gandour, 1981; Yiu, 1989); otherwise more inter-group confusions would have appeared.

The F_o shape and the perceptual space are applicable to explain the tonal confusion processes. The F_o shape, which corresponds to the direction and contour dimensions, predicts that only tones with similar F_o shapes will lead to significant confusion. The perceptual space determines whether tones with similar contour will lead to significant confusions or not. It is important to distinguish tones with similar contours (Vance, 1977). The closer between two tones of similar contour, the more likely confusion will occur. To sum up, the finding in this study supports the claim that direction and contour dimensions are more perceptually salient than height dimensions.

Tone identification

Like other auditory experiments in noise, the lexical tone identification deteriorates with decreasing *S/N*, that is, the higher the noise level, the more difficult it is to identify tones.

Noise drastically impedes tone identification when the S/N is below -15 dB *S/N*s. Between +20 dB and -15 dB *S/N*s, performance drops gently and steadily. As compared to consonant identification, the accuracy decreases steeply and steadily as *S/N* diminishes. Besides, the 50% identification accuracy for consonants occurs between +5 dB and 0 dB *S/N*s whereas, in this study, that for tones occurs between -15 dB and -20 dB *S/N*s. Tones are more resistant to the masking effect of noise than consonants are.

Pooling the performance across each *S/N*, falling tones are perceived better than the rising tones and level tones.

The perception of tones in noise cluster into groups at certain range of *S/N*s. Between 0 dB and -15 dB *S/N*s, the interactive effect of tones and *S/N* emerges that falling tones are identified more accurately than rising tones, and in turn, more accurately than the level tones. Below -15 dB *S/N*, the noise level is probably too high that all three groups of tones are very difficult to perceive. Conversely, all three groups of tones are identified without much difficulty above 0 dB *S/N*.

Tone identification are influenced by several factors which may explain the results obtained in this study:

(i) types of contour (level, rising, falling or their combinations) - gliding tones have a high degree of perceptual salience that human perception is better for tones with movements (Whitfield and Evans, 1965; Pollack, 1968; Anderson, 1978; Gandour, 1978);

(ii) number of tones with the same contour - the more number of tones with the same contour, the more misidentification and confusion are expected; and

(iii) perceptual space between tones with similar contour - the closer the F_o tracings between tones with similar contour, the more confusions are expected.

The level group comprises three tone members of high level tone 1 (55), mid level tone 3 (33) and low level tone 6 (22). In the rising and falling group, each has only two tone members: high rising tone 2 (35) and low rising tone 5 (23); and high falling tone 1 (53) and low falling tone 4 (21). The first two factors, contour and number, predict that the level group are more prone to misidentification than the falling and rising group. This agrees with the finding.

Although falling tones are more difficult to perceive than rising tones at acoustic level (Klatt, 1973), the perceptual distance between high falling tone 1 (53) and low falling tone 4 is much greater than that between high rising tone 2 and low rising tone 5 (c.f. figure 1). Especially, the distance between the initial portions of tone 2 and tone 5 are very close. Therefore, rising tones are misidentified much more than the falling tones in Cantonese. The importance of distribution of tones, i.e. the tone space, is indicated (Hombert et al, 1979).

Tone 1 variants

The high falling tone 1 (53) is best identified and not significantly confused with other tones whereas considerable misidentification and significant confusion occurs in high level tone 1 (55) [1(55)->3]. Although they do not contrast lexically, tone 1 (53) and tone 1 (55) are perceptually different.

Further investigation

The monosyllable of /fu/ was used in this study to examine the tone perception, it is difficult to generalize the results to daily conversation situation. During conversation, tones are produced shorter than in isolation and their F_o tracings are affected by adjacent tones (Vance, 1977). Further study on tonal confusions beyond syllable level will indicate how these variables interact with tone perception in Cantonese.

CONCLUSION

Both noise and tones produce effect on tone identification in Cantonese. The relative ease of perceiving the lexical tone groups, in descending order, is falling tones, rising tones and level tones. Tones are found to be more resistant to the masking effect of noise than consonants are.

There are seven significant tonal confusion processes. Five of them are within-group confusions and two are inter-group confusion. The causes of tonal confusions can be explained by the F_o shapes and the perceptual space between tones. Besides these two factors, the tone identification accuracy can be explained by the number of tones with similar contour.

The claim that direction and contour dimensions are more perceptually salient than height dimension is supported. High falling tone 1 (55) is perceptually different from low falling tone 4 (21). Research on tonal perception beyond syllable level is suggested.

ACKNOWLEDGEMENT

I am very grateful to my supervisors, Professor Ann Zubrick and Dr. Godfrey Harrison, for their guidance and support throughout my study and at various stages of my research. I wish to thank also Mrs. Lydia So and Mr. Dennis Au for their valuable advice. Additionally I would like to thank the technicians in the department for their technical support. Lastly, I have to express my gratitude to those who participated in the experiment.

REFERENCES

American Speech-Language-Hearing Association. (1990) Guidelines for screening for hearing impairments and middle ear disorders. *Asha,* **32,** (Supplement 2)., 17-24.

Anderson, S.R. (1978). Tone features. In V. A. Fromkin (Ed.) *Tone: a linguistic survey*. New York: Academic Press.

Chao, Y.R. (1947). *Cantonese Primer*. Cambridge: Harvard University Press.

Ching, T.Y.C. (1984). Lexical tone pattern learning in Cantonese children. *Language Learning and Communication,* **4,** 317-33.

Fok, C.Y.-Y. (1974). *A perceptual study of tones in Cantonese*. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Gandour, J. (1978). Tone Perception. In V.A, Fromkin (Ed.) *Tone: a linguistic survey*. New York: Academic Press.

Gandour, J. (1981). Perceptual dimensions of tone: evidence from Cantonese. *Journal of Chinese Linguistics,* **9,** 20-36.

Hombert, J.-M, Ohala, J.J., & Ewan, W.G. (1979). Phonetic explanations for the development of tones. *Language,* **33,** 37-58.

Kao, D. (1971). *Structure of the syllable in Cantonese*. The Hague: Mouton.

Klatt, D. (1973). Discrimination of fundamental frequency contours in synthetic speech. *Journal of the Acoustical Society of America,* 53, 8–16.

Miller, G.A., Heise, G.A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology,* **41,** 329-352.

Miller, G.A., & Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America,* **27,** 338-52.

Pike, K.L. (1948). *Tone languages: a technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion.* Ann Arbor: University of Michigan Press.

Pollack, I. (1968). Detection of rate of change of auditory frequency. *Journal of Experimental Psychology*, 77, 535-531.

Tse, J: K.-P. (1973). *The upper even tone in Cantonese: an instrumental investigation.* Taipei: Chia Hsin Foundation.

Vance, T.J. (1976). An experimental investigation of tone and intonation in Cantonese. *Phonetica*, **33**, 368-392.

Vance, T.J. (1977). Tonal distinctions in Cantonese. *Phonetica*, **34**, 93-107.

Wang, M.D., & Bilger, R.C. (1973). Consonant confusions in noise: a study of perceptual features. *Journal of the Acoustical Society of America*, **54**, (5), 1248-66.

Whitfield, I.G., & Evans, E.F. (1965). Responses of auditory neurons to stimuli of changing frequency. *Journal of Neurophysiology*, **28**, 655-72.

Yiu, M.L. (1989). *Tonal disruption in Cantonese aphasics.* M.Phil. Thesis. Hong Kong: University of Hong Kong.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: facts and models.* Berlin: Springer-Velag.