LINGUISTICS, PSYCHOLOGY,
AND THE CHINESE LANGUAGE

edited by

Henry S. R. Kao

and

Rumjahn Hoosain

Centre of Asian Studies

University of Hong Kong

1986

9 Lexical Tone Perception in Mandarin Chinese: Evaluation and Integration of Acoustic Features

Chiu-Yu Tseng
Institute of History and Philology
Academia Sinica, Taiwan

Dominic W. Massaro and Michael M. Cohen
University of California, Santa Cruz

A traditional concern in speech research is to determine the acoustic cues that are utilized in perception. The goal is to establish the psychological relationship between perception and properties of the speech stimulus. Continuing in this tradition, the purpose of the present investigation is to determine the acoustic characteristics utilized in the perception of lexical tone in Mandarin Chinese. It has been assumed in linguistic theory that the principal phonetic features of tone are found in the domain of pitch, whose primary acoustic correlate is the fundamental frequency (F0). The term "tone" refers to a particular way in which pitch is utilized in language. A tone language is a language that utilizes pitch to contrast individual lexical items or words (McCawley, 1978). This definition includes the traditional tone languages of Africa, the Americas, and Asia as well as the marginal tone (or "pitch accent"[sic]) languages of Europe and Japan; and excludes intonation languages, like English, in which pitch is used to signal syntactic and/or semantic distinctions at the phrase or sentence level (Gandour, 1978).

There are four lexical tones in Mandarin Chinese. According to the overall F0 patterns, these tones can be described as high-level, mid-rising, mid-falling-rising and high-falling (Chao, 1968). They are traditionally called Tones 1, 2, 3 and 4 respectively. For example, the syllable *ma* with Tones 1, 2, 3 and 4 would mean 'mother,' 'hemp,' 'horse' and 'reproach.' Although linguistic theory has stressed the feature of pitch in the description of tone, it is possible that other characteristics are used in perception. Besides the F0, its onset frequency and direction; some other possible acoustic characteristics specifying differences in tones are vowel duration, amplitude and vowel quality.

Studies of tone production (Chuang, Hiki, Sone, and Nimura 1971; Howie, 1976; Tseng, 1981) have found that certain acoustic characteristics vary with changes in tonal variants for Mandarin words produced in isolation. For example, the intrinsic duration differed for different tones. The falling-rising tone (Tone 3) was the longest in vowel duration, while the falling tone (Tone 4) was the shortest. The intrinsic amplitude also seemed to vary. The falling-rising tone (Tone 3) was produced with the lowest amplitude, while the falling tone (Tone 4) was the

English). Also, the FO pattern for a particular tone appears to change in different vowels (Howie, 1976). If these observations reflect reliable differences in the acoustic characteristics of Chinese tone, then listeners might use the information in the perceptual recognition of tone. In addition, the exact property or properties of FO pattern important in tone perception have not been determined. Our present aspiration is to determine the acoustic features responsible for the perception of Chinese tones, the relative importance of these features, and how they are combined during perceptual recognition.

Several previous studies have attempted to investigate the perception of Mandarin tones. Chuang et al. (1975) presented sets of real-speech Chinese monosyllables differentiated by tones and found that Chinese listeners can readily identify the tones. Most errors involved confusions between Tone 2 and Tone 3. Howie (1972, 1976) conducted a series of perception tests in which Chinese listeners were asked to identify the four tones under various conditions. When the four different synthetic tones were imposed on a real-speech syllable, listeners identified the tones correctly 95% of the time. Howie also generated a set of real-speech citation syllables and replaced their original FO contours with a constant FO of 128 Hz. The listeners did poorly in identifying the tones under this condition. Howie concluded that the FO pattern is critical in the identification of the four Mandarin Chinese tones. However, Howie's studies did not further address the question of which characteristics of the FO pattern were important.

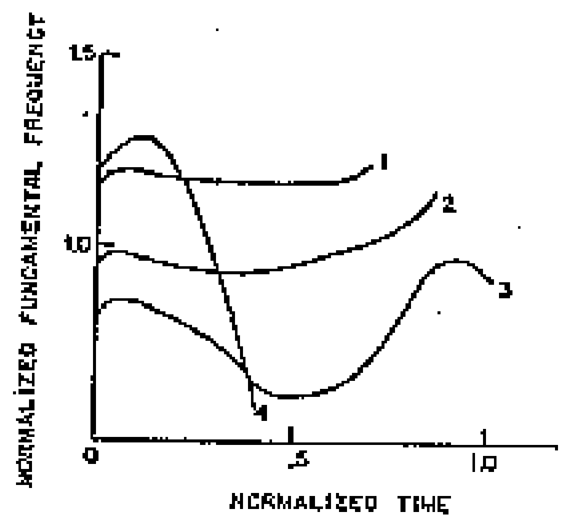
The perception of tone has also been studied using the procedure of multidimensional scaling. Gandour and Harshman (1978) synthesized 13 variants of the syllable /wa/ differing in both duration and FO frequency. The speech sounds were either long or short in duration, and had FO frequencies that were level, rising, and falling. The overall FO frequency was also varied as was the slope of the linear trajectories of the falling and rising tones. Thai, Yoruba, and American English subjects rated the dissimilarity between all possible pairs of the 13 speech sounds on an 11-point scale. Five dimensions best described the dissimilarity ratings. The dimensions were in decreasing order of importance: overall FO frequency, direction of FO change, duration of the syllable, the extremity of the FO frequency at the end of the syllable, and the slope of the FO change during the syllable. There were also interesting cross-language differences in the relative importance of the five dimensions.

EXPERIMENT 1

FO Pattern and Duration

As we have stated previously, the motivation of conducting an experiment on the role of FO pattern and vowel duration stems from observations of available production data. Figure 1 illustrates average FO contours and durations of the four lexical tones in Mandarin Chinese. It should be noted that only the averages are plotted in Figure 1 and there are large differences as a function of vowel context and syllable position. The patterns used in the present experiments are modelled after more specific occurrences of the patterns and, therefore will differ significantly from these average patterns. Both FO pattern and duration differ among the four tones. The FO pattern follows, at least roughly, the high-level, mid-rising, mid-falling-rising, and high-falling description of the four tones. The four tones also differ somewhat in duration. Howie (1974) observed that in Mandarin, vowels or syllables were longer with rising tones than falling tones. Tseng's (1981) data demonstrated that the falling-rising tone had the longest duration across all Mandarin vowel phonemes and the falling tone the shortest. The mean duration for the vowel /i/ was over 200 msec. longer in falling-rising tone than in falling tone. Given these observations, the first experiment evaluates the role of FO pattern and duration in the perception of the two tones.

Figure 1 Average Fundamental Frequency Contour and Duration (one speaker) of the Four Chinese Tones for Monosyllables in Prepausal Position. (Adapted from Gandour, 1978)



Method

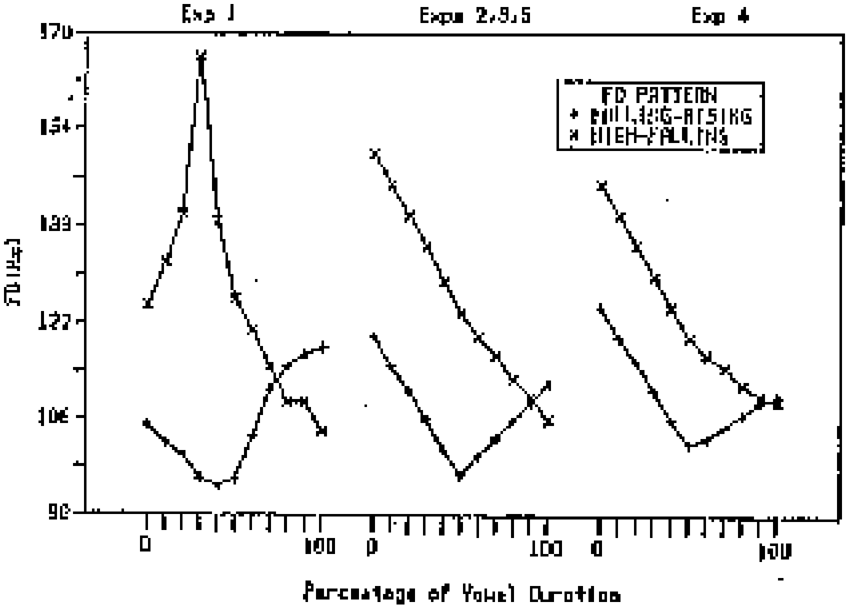
Subjects. The present experiments tested relatively few subjects for a relatively long period as opposed to testing many subjects for just a short period. At the practical level, it was difficult to find many native Chinese-speakers in our community. However, this strategy enabled us to look at individual results concerning the questions of interest. Individual subject data provide a stronger test of the fuzzy logical model since averaging across subjects will sometimes neutralize important trends in the results.

Six Chinese subjects, four male and two female visiting scholars at University of California Santa Cruz, participated in the experiment. All of the subjects selected were native speakers of Mandarin Chinese who had resided in the Peking area for most of their lives prior to their recent arrival in the U.S. Prior to their arrival in the U.S. a few months before the experiment, these subjects had little or no contact with foreigners. The subjects were paid $5.00 per hour for their service. Up to four subjects could be tested simultaneously in separate sound-attenuated rooms.

Stimuli. The stimuli were 49 vowels produced by the factorial combination of seven level of FO pattern and seven levels of vowel duration. The vowel used was a high front unrounded [i]. To synthesize the vowel, we began with the formant frequency values given by Howie (1976). These values were modified to derive an ideal Mandarin Chinese /i/, as judged by one of the authors. The values of F1, F2, F3, F4 and F5 were 255, 2190, 3130, 3730 and 4500 Hz respectively

The F0 pattern ranged between falling-rising tone and falling tone in Mandarin Chinese in equal steps. The F0 patterns of the two extreme tones were set equal to the average F0 patterns for the vowel /i/ given in Howie (1976). The F0 pattern was specified at 11 points during the vowel, and the F0 was linearly interpolated between adjacent values. The left panel of Figure 2 gives the F0 values of the endpoint stimuli for the falling-rising to falling tone continuum.

Figure 2 Endpoint F0 Patterns Used in Experiments 1 Through 5



The duration of the vowel stimuli ranged from 100 ms. to 400 ms. in equal steps of 50 ms. The F0 pattern was simply compressed or expanded as duration was changed. The amplitude parameter (AV) was set at 0 dB at the onset of the vowel, then moved linearly to 55 dB at 30 ms. after the onset. The amplitude then decreased linearly to 50 dB at 30 ms. before the offset, and to 0 dB at the offset. All stimuli used in Experiment 1 had this basic form, with systematic variation of the F0 pattern as well as the duration.

The speech sounds were synthesized using the Klatt (1980) software speech synthesis program implemented on a Digital PDP-11/34A computer. Time values were specified and parameters were calculated in 5 ms. increments. The output of the speech synthesizer was recorded with a sampling rate of 10 KHz and stored in files on a disc. During the experiment, the stimuli were played back at 10 KHz by a 12 bit D-to-A converter (Data Translation Model 1711). The output of the D-A converter was filtered at 20-4800 Hz by a Krohn-Hite Model 3500R bandpass filter. The stimuli were then amplified (McIntosh model MC-50) and presented to subjects at approximately 68 dB SPL over Koss Pro-4AA headphones.

Procedure. Each trial began with the presentation of one of the 49 vowels selected randomly without replacement in blocks of 49 trials. The subjects were told to identify the tonal pattern of the stimuli, i.e., falling-rising or falling, traditionally called Tone 3 and Tone 4 in Mandarin Chinese. Each subject responded by pressing either the key "3" or "4" on a computer terminal keyboard (TeleVideo Model 950). Subjects had 3 sec. to make their response. The next trial began immediately after the response interval. There were 25 practice trials followed by 394 experimental trials in each experimental session. The subjects did not know that the first 25 trials were practice and would not be analyzed. Before the first session of the first day, a practice session of 150 trials was also administered. The subjects participated in two sessions on each of four different days. The subjects took about a 10 minutes break between sessions. A total of [illegible] observations were collected for each subject for each of the 49 test sounds.

Results

Figure 3 shows the proportion of falling judgments as a function of the F0 pattern; duration of the vowel is the curve parameter. The likelihood of a "falling" judgement increased systematically with changes from falling-rising to high-falling F0 pattern, F(6,30)=319, p<.001. Although the main effect of vowel duration was not significant, vowel duration did interact with F0 pattern, F(36,180)=2.77, p<.001. Overall, the likelihood of a falling judgment decreased with increases in vowel duration and this effect was largest at the ambiguous levels of the F0 pattern (levels 3 and 4 of the F0 pattern in Figure 3). Table 1 gives the proportion of falling judgements for the six subjects. The overall effect of duration indicates a decrease in falling judgments with increases in vowel duration from 100 msec to 250 msec.

Figure 3 Proportion of Falling Identifications as a Function of the F0 Pattern. Vowel Duration in ms. is the Curve Parameter (Experiment 1)

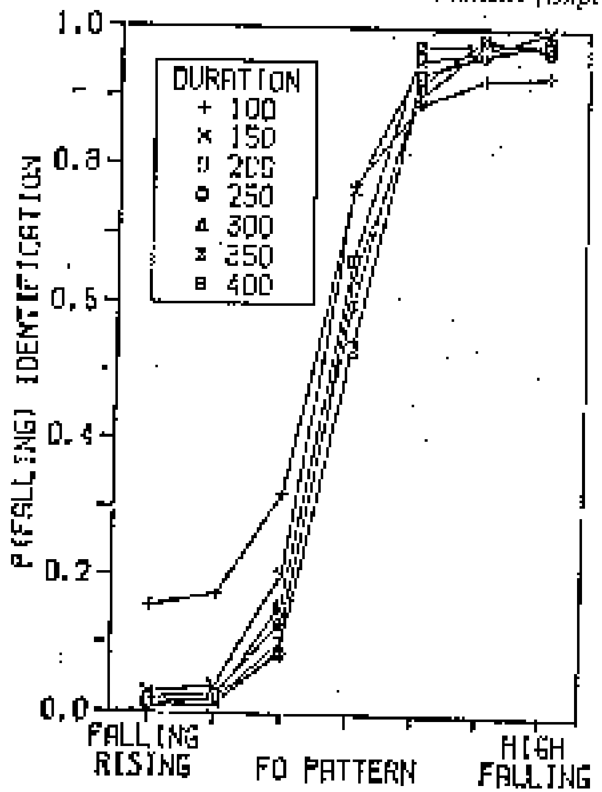


Table 1 Proportion of Falling Judgments for the Six Subjects as a Function of Duration of the Vowel

	Duration						
Subject	100	150	200	250	300	350	400
1	.75	.59	.52	.45	.47	.48	.51
2	.45	.53	.55	.48	.49	.51	.53
3	.53	.54	.52	.50	.50	.50	.50
4	.59	.56	.57	.55	.56	.57	.57
5	.37	.53	.53	.55	.50	.54	.56
6	.86	.67	.59	.50	.50	.50	.47
Average	.60	.57	.54	.51	.50	.52	.52

In addition to the role of duration as a cue for the discrimination of lexical tone, another result shown in Figure 2 is of interest. Subjects were less consistent in their tone judgments given a 100 ms. vowel. The result indicates that the discrimination of FO pattern was relatively degraded when the duration of the vowel was 100 ms. This result suggests that optimal processing of the FO patterns used in this study requires a vowel that is over 100 ms. long.

One might argue that the 100 ms. duration level was responsible for the interaction of duration and FO pattern seen in Figure 3. To test this idea, we eliminated this level in an additional analysis. The duration effect was now significant, $F(5,25)=2.78$, $p<.04$, and the interaction between duration and FO pattern remained statistically significant, $F(30,150)=1.63$, $p<.03$. Accordingly, there is still a significant effect of duration and especially when the FO pattern is ambiguous.

Although the effect of duration was statistically significant, only two of the six subjects showed a large and systematic effect. Their linguistic background did not seem to have much relevance to their sensitivity to duration as a cue to lexical tone. For these two subjects, FO pattern was still the most influential cue and the influence of duration was strongest when the FO pattern was relatively ambiguous.

Discussion

The results of Experiment 1 demonstrate that the subjects' perception of tone was mainly influenced by the FO pattern, as other previous studies have reported (Gandour, 1978, 1981). Duration had two different effects. First, when the vowel duration was very short (100 ms.), subjects' judgment of the FO pattern was degraded. This suggests that in order for the FO pattern to be perceived accurately, the vowel portion that bears the FO pattern needs to be longer than 100 ms. Second, 2 out of the 6 subjects did use vowel duration as a cue in perceiving tones. Linguistic background of the subjects did not seem to offer any explanation for these differences. So far, we can only conclude that some speakers utilize duration as a cue for tone perception, as has been suggested in some previous studies (Gandour, 1978).

EXPERIMENT 2

FO Pattern and Vowel Quality

The goal of the second and third experiments was to evaluate the influence of vowel quality on the perception of lexical tone. It has been observed that the FO pattern for tone varies with different vowels (Howie, 1976). Howie's (1976) measurements indicate that both the starting FO and the FO contour for a given tone differ for different vowels. For example, in all four Mandarin tones, the syllables with the high vowel /i/ were significantly higher in FO than corresponding syllables with the low vowel /a/ (Wang, 1977). The positive correlation between vowel height and pitch height has been observed for English (Lehiste, 1970; Peterson and Barney, 1952), and other languages as, for example, Danish (Petersen, 1976) and Korean (Kim, 1968). It might be the case that listeners evaluate the FO pattern relative to the vowel context. Hombert (1978) conducted an experiment in which the vowels /i/, /a/ and /u/ were synthesized, and three FO frequencies (115, 120, and 125 Hz) were superimposed on each vowel. American subjects were given two-vowel sequences in which the FO of the second vowel was either 5 Hz below, equal to, or 5 Hz above the FO of the first vowel. The subjects' task was to judge whether the first or the second vowel was higher in pitch. The low vowel /a/ was judged higher in pitch than the high vowels /i/ or /u/, although their FO frequencies were in fact equal. Similar results were obtained by Chuang and Wang (1978). If these results generalize to lexical tone judgments, a given FO pattern might be interpreted differently in different vowel contexts. However, given that the perceived variation of FO with vowel quality can be expected to be small relative to the changes in lexical tone, any effect of vowel context should only occur with ambiguous values of FO pattern.

In Experiment 2, a vowel continuum was created between the vowels /i/ and /y/. A continuum between / / and /a/ was used in Experiment 3. In both experiments, the same continuum of FO pattern going between falling-rising and falling tone was used. Subjects identified both the vowel quality and lexical tone but we restrict our analysis to just the tone judgments.

Method

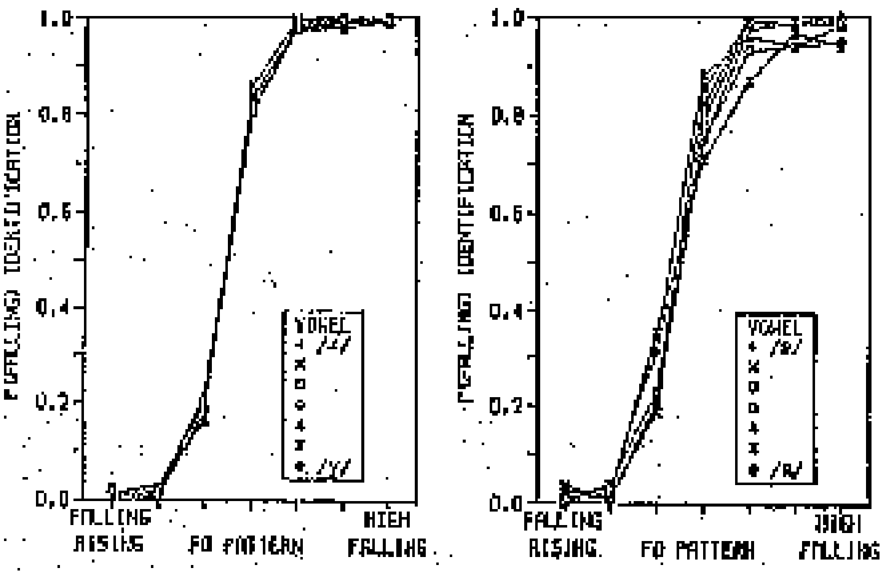
In the second experiment, the two variables manipulated were the FO pattern and the formant pattern. New FO Patterns were chosen for the ends of the falling-rising to falling continuum based on the averages across many different vowel contexts (see the middle panel of Figure 2 for these new values of the FO pattern continuum). These values of FO pattern were again specified at 11 points during the vowel. As in Experiment 1, the continuum was divided in 6 equal steps to generate a set of 7 patterns, ranging from falling-rising tone to falling tone. For the formant pattern, a continuum of seven vowels from /i/ to /y/ was generated. Both vowels are high front vowels in Mandarin Chinese which differ in only rounding, i.e., /y/ is a rounded vowel while /i/ is unrounded. The values of F1, F2, F3 and F4 for vowel /y/ were 262, 1973, 2542, and 3442 Hz respectively. The values for /i/ were 252, 2082, 2846 and 3586 Hz. The values were based initially on Howie (1976) and later revised slightly. The seven stimuli along the continuum were equally spaced in terms of their formant frequencies. The values for F5 remained constant at 4500 Hz for both vowels. The 49 vowels representing the factorial combination of FO pattern and formant pattern were synthesized as in Experiment 1. The duration of each vowel was 250 ms.

The same 6 Chinese subjects and identical testing procedures were employed. The response choices were words corresponding to the vowels /i/ and /y/ in both the falling-rising tone (Tone 3) and high-falling tone (Tone 4). Four keys on the computer terminal keyboard were labeled with the four Chinese words. The subjects participated in two 30-minutes sessions on each of four different days. A total of 48 observations were collected for each of the 49 stimuli from each subject.

Results

The left panel of Figure 4 shows the proportion of falling responses as a function of the FO pattern continuum moving from falling-rising to high-falling; vowel formant pattern is the curve parameter. As can be seen, the likelihood of falling judgments increased systematically with changes in the FO pattern, $F(6,30)=97.7$, $p<.001$. Although FO pattern appears to change with changes in vowel quality in real speech, vowel quality had no influence on the tone judgments.

Figure 4 Proportion of Falling Identifications as a Function of the FO Pattern; Vowel Quality is the Curve Parameter (Experiment 2 in the Left Panel, Experiment 3 the Right Panel)



EXPERIMENT 3

FO Pattern and Another Vowel Quality

Method

To provide another test of the contribution of vowel quality, a different vowel continuum, i.e., vowels /a/ to /ə/, was created to replicate Experiment 2. The values of the formant frequencies for /a/ were 635, 1363, 2488, 3359 and 4500 Hz for F1, F2, F3, F4 and F5 respectively. The values of the formant frequencies

for vowel /ə/ were set at 702, 1136, 2533, 3291 and 4500 Hz. All other parameters of the 49 stimuli were equivalent to Experiment 2. The same testing procedure was employed with four response choices corresponding to /a/ and /ə/ in both falling-rising tone and falling tone. The same group of Chinese subjects participated in this experiment. Again, we restrict our analysis to the judgment of lexical tone. A total of 24 observations per stimulus was collected for each subject in four experimental sessions.

Results

The right panel of Figure 4 gives the percentage of "falling tone" judgments as a function of the FO pattern going from falling-rising to falling tone; vowel quality is the curve parameter. The "falling tone" judgments changed systematically with FO pattern, $F(6,30)=1.01$, $p<.001$. Although the main effect of vowel quality was not significant, $F(6,30)=.68$, the interaction of FO pattern and vowel quality was significant, $F(36,180)=3.06$, $p<.001$. The functions are steepest for the vowels heard most like /a/ and more gradual for the vowels heard as /ə/. Thus, the perception of FO pattern appears to be more consistent with the vowel /a/ than with the vowel /ə/. The result is difficult to explain and contrasts with the absence of any effect of the /i/ — /y/ continuum in the previous experiment.

Given that the same FO patterns were used in Experiments 2 and 3, it is possible to ask whether the change in tone judgments differed as a function of the two different vowel continua. This is a very strong test of the contribution of FO pattern as a function of vowel quality because the vowel difference between the two continua is greater than the difference within each of the continua. However, an analysis of variance on the tone judgments across the two experiments revealed no interaction of FO pattern and the two different vowel continua, $F(6,30)$ $p<1$.

Discussion

The results of Experiments 2 and 3 demonstrated that subjects' judgments of tone changed systematically with changes in FO pattern. However, vowel quality did not appear to provide a cue for the perceptual recognition of lexical tone.

EXPERIMENT 4

FO Pattern, Amplitude, and Vowel Quality

We have stated previously that differences in amplitude are found with different tonal patterns in Mandarin Chinese (Li, 1964; Tseng, 1981). When produced in citation forms, the falling-rising tone (Tone 3), was produced with the lowest amplitude, whereas the falling tone (Tone 4) was the highest. If listeners utilize amplitude as a cue to the perceptual recognition of tone, they should tend to hear louder vowels as Tone 4 rather than Tone 3. Abramson (1975) found that amplitude differences contributed to the perception of tones in Thai, even though FO pattern was the dominant cue. We are thus interested in whether amplitude is utilized as a cue in perceiving tones, and if it is, how it is integrated with the FO pattern.

Method

Three variables were manipulated, namely, vowel quality, FO pattern and amplitude. Forty-five stimuli were generated by combining five levels of ...

pattern, three levels of FO pattern and three levels of amplitude. All stimuli were 250 ms. long. The five levels of formant pattern were equivalent to levels 2 through 6 from Experiment 2. The values of the formant frequencies F1, F2, F3 and F4 for /i/ were 254, 2064, 2795, and 2562 Hz respectively, and for /y/ 261, 1991, 2593 and 3455 Hz respectively. F5 remained constant at 4500 Hz. The FO pattern continuum was created by choosing steps 3 and 4 from the falling-rising to falling continuum in Experiments 2 and 3 as the two endpoints. (See the right panel of Figure 2 for the values of the chosen endpoints of the FO pattern continuum.) This continuum was then divided in 2 equal steps to yield three FO patterns. The three levels of amplitude were approximately 4 dB apart. The corresponding average amplitudes at the subjects' ears were 58, 55 and 52 dB-A SPL respectively.

The same 6 Chinese subjects participated in two 30-minutes sessions on two different days of experimentation. A total of 24 observations were collected for each of the 49 stimuli from each subject.

Results

Figure 5 shows the proportion of falling tone judgments as the FO pattern varies from falling-rising to high falling; amplitude is the curve parameter. The FO pattern had the expected effect, $F(2,10)=340$, $p<.001$. The main effect of amplitude was significant, $F(2,40)=13.97$, $p<.001$, as was the interaction of FO pattern and amplitude, $F(4,40)=13.05$, $p<.001$). The results indicate that "falling" judgments increased somewhat with decreases in amplitude when the FO pattern was ambiguous. This result is surprising in that falling tones are usually louder than falling-rising tone (Li, 1964). Replicating Experiment 2, vowel quality did not influence the tone judgments, $F<1$.

Discussion

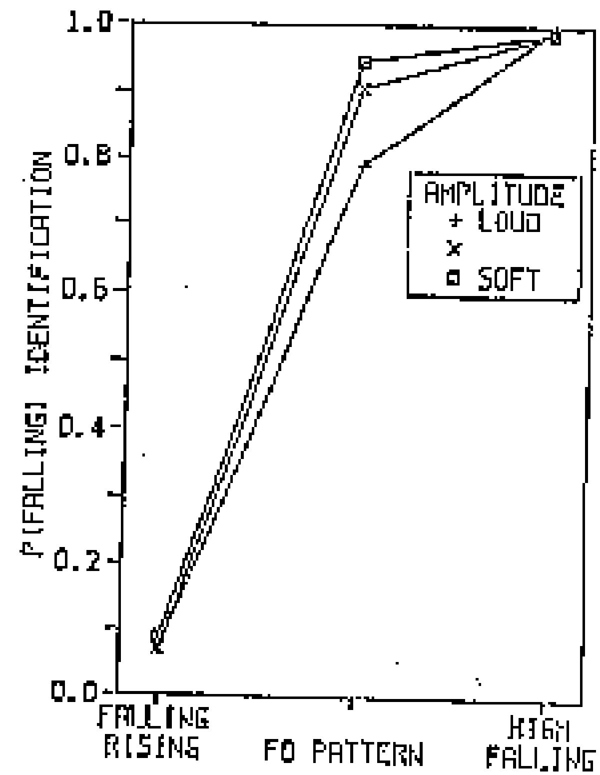
The results of Experiment 4 demonstrated once more the expected effect of the FO pattern. We also found significant effects of amplitude in that the subjects' judgment of tone did change slightly with changes in amplitude, but the amplitude effect is in the opposite direction from what might be expected given the amplitude difference observed in natural speech. The results show that the falling judgment increased slightly with decreases in amplitude, contrary to the larger amplitude of the falling tone in natural speech.

EXPERIMENT 5

FO Height and FO Contour

The purpose of the final experiment is to evaluate the contribution of two components of the FO pattern feature. The FO pattern in real speech and in the previous experiments is characterized by differences in frequency height and differences in contour. Either or both of these properties might be responsible for the perceptual importance of FO pattern in the recognition of lexical tone. In the fifth experiment, the height of the FO patterns was increased or decreased while maintaining the same absolute change during the speech sound.

Figure 5 Proportion of Falling Identifications as a Function of the FO Pattern; Amplitude is the Curve Parameter (Experiment 4)



Method

The same seven FO patterns used in Experiment 2 were employed at each of five different frequencies, giving a total of 35 stimuli. The test vowel was /i/ with a duration of 250 ms. Each FO pattern occurred either at the same frequencies used in Experiment 2, raised 10 or 20 Hz, or lowered 10 or 20 Hz. If contour is the critical acoustic feature, then frequency should have no influence on perceptual recognition. However, an effect of frequency would show that subjects are sensitive to and use FO height as a cue to lexical tone.

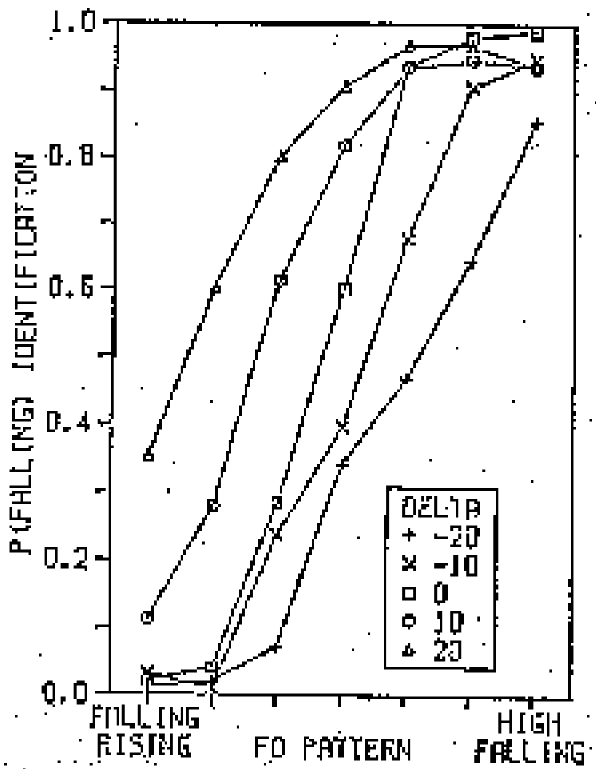
Six Chinese subjects, three native speakers of Mandarin Chinese from the Peking area and three native speakers of Mandarin Chinese from Taiwan served as subjects. Ss 1, 2 and 3 from China had resided in Peking prior to their recent arrival to the U.S. S1 and S2 had participated in the earlier experiments.

The same testing procedures were employed for this experiment. The subjects were asked to identify the tones either as the falling-rising tone or the falling tone, labeled Tones 3 and 4 on their computer terminal keyboard. The subjects participated in 2 sessions of the experiment. Twelve observations were collected for each subject for each of the 35 test sounds.

Results

Figure 6 gives the proportion of falling judgments as a function of FO pattern; the change in frequency is the curve parameter. The FO pattern had a strong effect on perceptual recognition, $F(6,30)=7.5$, $p<.001$. Differences in FO height also had a strong effect in that the proportion of falling judgments decreases with decreased in FO height, $F(4,20)=9.7$, $p<.001$. The interaction between the two variables, $F(24,120)=5.6$, $p<.001$, shows that the effect of each variable was largest at the more ambiguous level of the other variable.

Figure 6 Proportion of Falling Identifications as a Function of the FO Pattern; Relative FO Height (Delta) is the Curve Parameter (Experiment 5)



Discussion

The results of this experiment demonstrated a significant effect of both the FO contour and FO height on the perception of mid-falling-rising and high-falling tones. It has been well documented that various tone languages, for example, Thai, Cantonese, Yoruba, etc., possess register tones, that is, tones that are distinguished on the basis of FO frequency (or pitch height). As an example, Abramson (1975) found that the height of flat FO patterns cued the distinction among the low, mid and high tones in Thai, but not between the falling and rising

tones. What is of most interest is that both the FO contour and FO frequency are utilized by the listeners as cues to perceive tones in Mandarin which does not possess true register tones. The interaction between these two effective cues is such that the contribution of one cue is increased when the other cue is more ambiguous.

GENERAL DISCUSSION

We have reported 5 experiments in which we simultaneously manipulated two or more acoustic variables to assess their role in tone perception in Mandarin Chinese. The variables were: FO contour, FO height, vowel duration, vowel quality, and amplitude. As expected, the results across all the experiments demonstrated a strong effect of FO pattern on the perception of tones. The contribution of duration to the perception of tone was assessed in Experiment 1. Two of six listeners seemed to use duration as a functional cue in tone perception. The effect of vowel duration for the listeners who utilize vowel duration as a cue for tone perception was strongest when the other cue (FO pattern) was more ambiguous. The results from Experiment 1 also suggested that the vowel duration for the correct perception of the FO pattern without additional contextual information should be longer than 100 ms.

The results from Experiments 2 and 3 demonstrated that listeners were not sensitive to vowel quality as an effective cue for tone perception. However, the perception of FO pattern was degraded with the vowel /a/ relative to the vowel /u/ (cf. the right panel of Figure 4). An analogous effect was not observed for the /i/ — /y/ continuum.

The results from Experiment 4 showed the subjects' judgment of tone did change slightly with changes in amplitude. However, the amplitude effect is in the opposite direction from the amplitude difference observed in natural speech.

It was observed that the FO pattern contains both overall frequency and contour information. To test for the relative importance of these properties, they were varied in Experiment 5. The results demonstrated that both FO contour and FO height influence tone perception. The FO height and the FO contour are integrated together for perceptual recognition. The integration is such that the least ambiguous cue has the most impact on perception. An unambiguous FO contour reduces the contribution of FO height, whereas the FO height has a larger influence when FO contour is ambiguous. The descriptive labels for the lexical tones have both height and contour information. For example, Tones 3 and 4 are called mid-falling-rising and high-falling, respectively. The influence of these cues indicate that the labels are also perceptually valid.

FOOTNOTES

This research was supported by a grant from the National Institute of Mental Health (MH-35334) to Dominic W. Massaro.

REFERENCES

Abramson, A.S. (1975). The tones of Central Thai: Some perceptual experiments. In J.G. Harris & J. Chamberlain (Eds.), *Studies in Thai Linguistics*. Bangkok: Central Institute of English Language, 1-16.

Chao, Y.R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chuang, C.K., Hiki, S., Sone, T. and Nimura, T. (1971). The acoustical features and perceptual cues of the four tones of standard colloquial Chinese. In *Proceedings of the 7th International Congress on Acoustics* (Vol. 3). Budapest: Akademia Kiado, 297-300.

Chuang, C.K. and Wang, W. S-Y. (1978). Psychophysical pitch biases related to vowel quality, intensity, and sequential order. *Journal of the Acoustical Society of America*, 64, 1004-1014.

Gandour, J. (1978). The perception of tone. In V. Fromkin (Ed.), *Tone: A Linguistic Survey*. New York: Academic Press, 41-76.

Gandour, J. (1981). Perceptual dimensions of tone: Evidence from Cantonese. *Journal of Chinese Linguistics*, 9.

Gandour, J. and Harshman, R. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech*, 21, 1-33.

Hombert, J.M. (1978). Consonant types, vowel quality, and tone. In V. Fromkin (Ed.), *Tone: A Linguistic Survey*. New York: Academic Press, 77-111.

Howie, J.M. (1972). Some experiments on the perception of Mandarin tones. In A. Rigault & R. Charbonneau (Eds.), *Proceedings of the 7th International Congress of Phonetic Sciences*. The Hague: Mouton, 900-904.

Howie, J.M. (1974). On the domain of tone in Mandarin. *Phonetica*, 30, 129-48.

Howie, J.M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.

Kim, C.W. (1968). Review of Lieberman 1967. *Language*, 44, 830-842.

Klatt, D.H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, 3, 971-995.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.

Li, K.P. (1966). Machine recognition of Mandarin monosyllables. In *Project on Linguistic Analysis*, Report No. 9. Columbus, Ohio: The Ohio State University.

McCawley, J.D. (1978). What is a tone language? In V. Fromkin, (Ed.), *Tone: A Linguistic Survey*. New York: Academic Press, 113-131.

Petersen, N.R. (1978). Intrinsic fundamental frequency of Danish vowels. *Annual Report of the Institute of Phonetics*. University of Copenhagen, 1-27.

Peterson, G.E. and Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175-184.

Tseng, C. (1981). An acoustic phonetic study on tones in Mandarin Chinese. Ph.D. dissertation, Brown University, Providence, R.I.

Wang, W. S.-Y. (1977). Review of Howie 1976. *Modern Language Journal*, 301.