Stewart, J. 1964. 'The typology of the Twi tone system, with comments by Paul Schachter and William Welmers' Legon University of Ghana, Institute of African Studies.

Surintramont, A. 1973. 'Some aspects of underlying syllable structure structure in Thai: evidence from Khamphuan - a Thai word game. *Studies in the Linguistic Sciences* 3,1:121-142.

Welmers, W. 1962. 'The Phonology of Kpelle' *Journal of African Languages* 1:69-93.

Woo, N. 1969. 'Prosody and phonology' MIT Doctoral dissertation. Reproduced by the *Indiana University Linguistics Club* (1972).

Contextual Factors Influencing Tone Discrimination[1]

J. M. Hombert and S. Greenberg

[Department of Linguistics

University of California, Los Angeles]

Introduction

Tone languages do not exploit the various possibilities of tonal contrasts with equal frequency. For example, falling tones have phonemic status in many tone languages while rising contours are relatively rare. Such disparities in the distribution of tonal patterns may be caused in part by articulatory constraints imposed by the mechanics of laryngeal displacement (Ohala and Ewan, 1973; Sundberg, 1973).

Additional constraints may be imposed by limitations on the auditory system's ability to encode frequency information of a waveform whose periodicity changes constantly.

Detection and recognition of a pure tone is adversely affected upon the introduction of a second tone in temporal proximity to the first. When the interfering tone *precedes* the target, the decrease in performance is significant only if the interval separating the tones is less than 60 msec (Samoilova, 1960; Elliot, 1962). However, the presence of a tone *following* the target can interfere with tonal recognition for inter-stimulus intervals up to 250 msec (Massaro, 1972). A tone *preceding* the target exerts maximal interference when its frequency is slightly below the target's. The situation is a bit more complex when the interfering tone *follows* the target. Identification of temporal order is most severely affected by a tone lying one-sixth to one-third of an octave above the pattern frequency range (Divenyi and Hirsh, 1975). In a tone recognition paradigm, however, the degree of interference appears independent of the direction of frequency change (Massaro, 1972).

In an earlier study Hombert (1975) found that the onset of a falling tone was more accurately perceived than the onset of a rising tone in synthesized vowel stimuli. The data indicated that 1) low tones are more effective at masking higher tones than the converse; and 2) forward masking is more significant than backward masking in the perception of fundamental frequency contours. In the case of a rising tone, each frequency would be masked by the immediately preceding lower frequency. This would not occur in the case of a falling tone since the fundamental frequency descends.

The experiment reported here was designed in order to clarify the role of backward versus forward masking on the one hand, and the role of relative frequency height on the other.

More generally, the purpose of this paper is to investigate whether interference caused by a neighboring tone (a phonomenon generally referred to as "masking" in the psychoacoustic literature) is relevant for the perception of pitch corresponding to the frequency range used by the human voice in speech and to use this knowledge to acquire new insights into the evolution of tone systems in the languages of the world.

Experimental Paradigm

Eight subjects, all native speakers of American English, participated in an experiment in which they were asked to compare the fundamental frequency of two synthesized vowels in the presence of an adjacent syllable. Each of the subjects had achieved a score of 95% or better on a pretest involving the comparison of the F_o of two neighboring vowels.

As can be seen in Figure 1, the set of stimuli[2] was composed of 1) a synthesized VCV sequence ([ini]) on which two fundamental frequency patterns were superimposed (either a low steady-state F_o of 119 Hz on the first vowel followed by a high steady-state F_o of 150 Hz or a high steady-state F_o followed by a low steady-state F_o), and 2) two sets of comparison vowels: the high set (140 Hz, 145 Hz, 148 Hz, 152 Hz, 155 Hz, 160 Hz) which were compared with the high reference vowel (F_o = 150 Hz); the low set (109 Hz, 113 Hz, 117 Hz, 121 Hz, 127 Hz, 130 Hz) which were compared with the low reference vowel (F_o = 119 Hz).

Subjects were asked to judge whether the comparison vowel's fundamental frequency was higher or lower than that of the reference vowel. If the comparison vowel was higher (lower) they were instructed to circle the letter "H" ("L") on their answer sheet.

These stimuli were grouped into eight conditions as presented in Figure 2: 1) The comparison vowel either followed the reference vowel (conditions 1, 2, 3, 4) or preceded it (conditions 5, 6, 7, 8); 2) The comparison vowel was either adjacent to the reference vowel (conditions 1, 2, 5, 6) or non-adjacent (conditions 3, 4, 7, 8); 3) The comparison and reference vowels were either high (conditions 1, 3, 5, 7) or low (conditions 2, 4, 6, 8).

The interval separating the comparison vowel and the disyllable was 350 msec. The subject had two seconds in which to make a response.

Each condition was composed of two identical subparts. Each subpart consisted of 66 trials (10 repetitions for each of six comparison vowels + six trials at the beginning of each subpart which were excluded from the data analysis).

The experiment was run in two sessions (Session A and Session B) separated by one or two days. Subjects were divided into four groups. The order in which the conditions were presented was counterbalanced across these four groups.

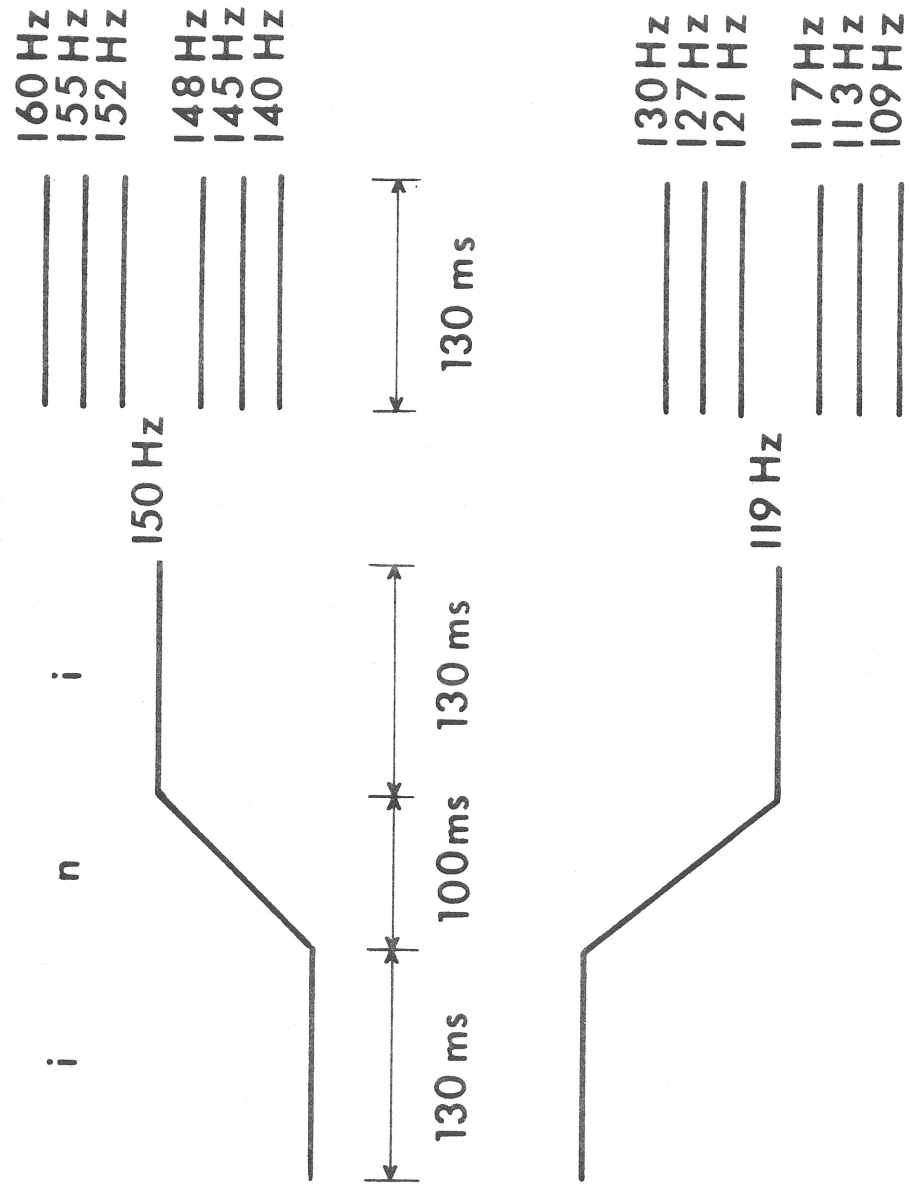

Figure 1: Stimulus configuration

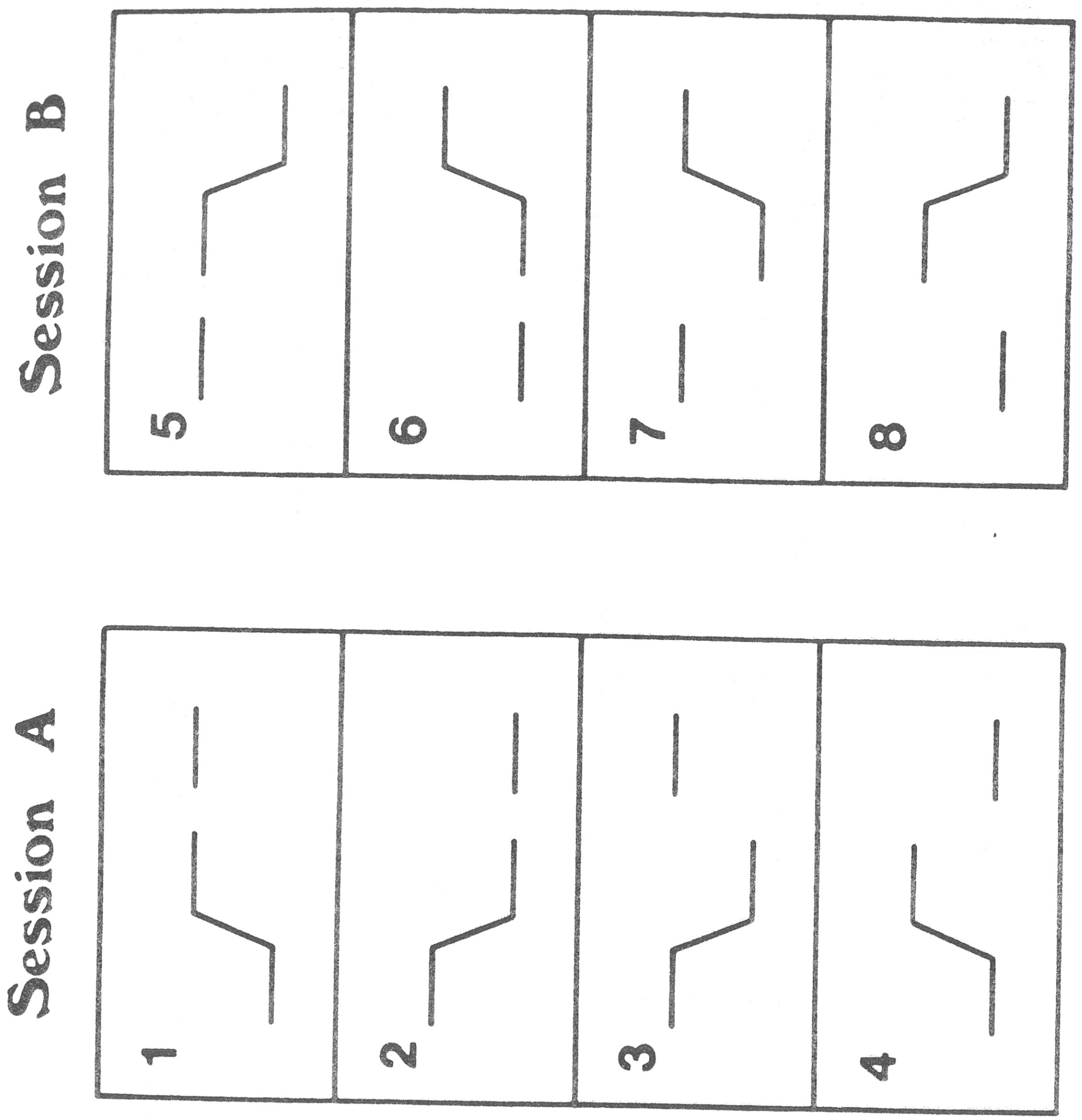

Figure 2: Experimental paradigm

Results

The results of this experiment are presented in Table 1. Our data were subjected to a four-way analysis of variance. Three parameters were within subjects; 1) Position of the comparison vowel; 2) Contiguity of the comparison and reference vowels; 3) Frequency height of the comparison and reference vowels. The fourth parameter involved the distribution of the results by group.

The d' measurements refer to a measure of the subjects' ability to judge the relative pitch of the vowels, which compensates for response bias (Green and Swets, 1966). As can be seen from the table, the position of the comparison vowel relative to the reference vowel does not affect the subjects' performance.

As expected, these results show that the comparison of two tones is easier when the comparison and reference vowels are contiguous. A more surprising result can be seen in the third row of the table. Comparison of low tones was better than comparison of high. This suggests that for our experimental paradigm, at least, low tones are perceptually more salient.

On Figure 3 performance level (the ordinate) is plotted as a function of the experimental condition (the abscissa). Subjects have been divided into two groups according to their performance. Performance values corresponding to the four best subjects (Group A) are indicated by a triangle. The performance values for the four subjects with relatively lower scores (Group B) are represented by circles. Sets of performance values (in d' units) averaged across each group of subjects are connected by a solid line.

Discussion

Analysis of the frequency distribution of errors makes it unlikely that the results are due to factors unrelated to fine frequency discrimination. For five of our eight subjects, ninety-four percent or more of the errors occurred when the fundamental frequency of the comparison vowel was either one step above or one step below that of the reference vowel.

The results are not consistent with our original hypothesis concerning the effects of backward and forward masking. A consistent pattern of interference effects is absent, both when the data is averaged and when it is analyzed by subject.

The most significant factor contributing to the observed interference involves the contiguity relationship of the comparison and reference vowels. Not surprisingly, discrimination is affected least when the compared vowels are adjacent to each other.

A second, less significant factor involves the frequency range of the compared stimuli. When the comparison involved vowels with fundamental frequencies in the lower range, discrimination is superior for all subjects in Session A and for the lower performance group in Session B.

Table 1: Experimental results

PARAMETERS		PERFORMANCE (in d' units)	
Position of c.v.	preceding	2.79	N.S.
	following	2.87	
Contiguity of c.v. and t.v.	adjacent	2.95	P < 0.02
	non-adjacent	2.71	
Frequency height of c.v. and t.v.	high	2.73	P < 0.05
	low	2.92	

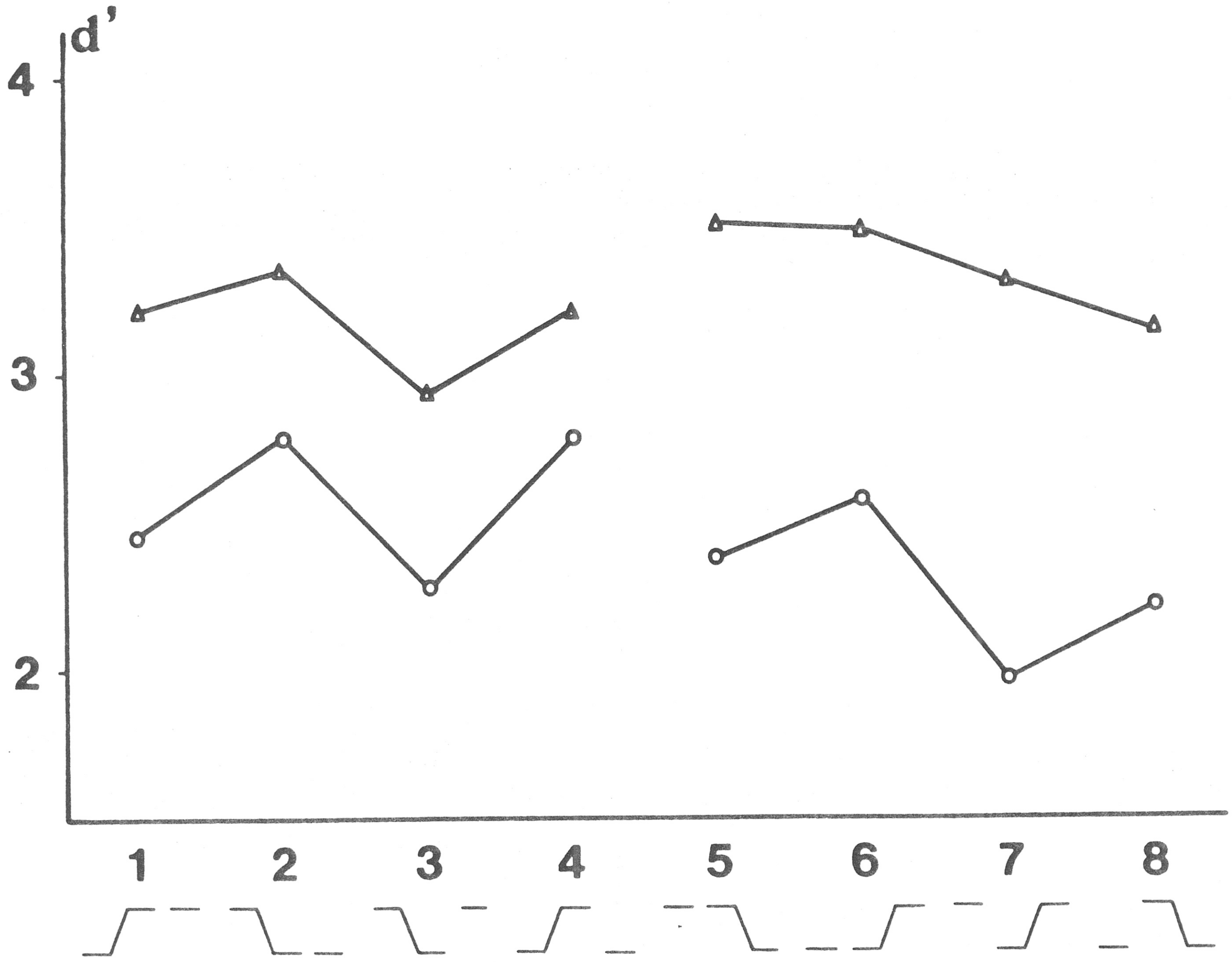

Figure 3: Performance level as a function of experimental condition

This result is somewhat surprising in view of the frequent observation of high tone prominence in many tone languages. One possible explanation could be that the violation of the normal correlation of lower intensity with low tone provides the low tone with an unaccustomed perceptual salience not found in the normal speech context. An investigation of this hypothesis is planned for the near future.

It is entirely possible that certain peculiarities of the experimental paradigm contributed to the observed effects. For instance, several subjects reported using a discrimination strategy which ignored the reference vowel in an attempt to base judgements solely upon the absolute frequency of the comparison vowel. This sort of strategy could account for the uniformly high performance level achieved by the superior performance group throughout the conditions of Session B. Randomly varying the absolute frequency of the target vowel among several possibilities within a restricted range should render this strategy ineffective. In addition, it may be advisable to increase the interval separating the comparison and reference stimuli in order to minimize possible secondary interference effects.

Conclusion

This study did not confirm our hypothesis concerning the relative importance of backward and forward interference of linguistic tones. In fact, our data appear to indicate that the findings concerning backward and forward masking are not directly relevant for stimuli having a fundamental frequency corresponding to the human voice range.

As expected, it was found that the comparison of fundamental frequency is more accurate when the two vowels are contiguous. As well, our data suggest that the perception of low tones is more accurate than the perception of relatively higher tones. This finding points indirectly to the role of amplitude information involved in the perception of tones in tone languages.

Footnotes

1. This paper was presented at the 90th Meeting of the Acoustical Society of America, San Francisco, California, November 3-7, 1975.

2. The fundamental frequencies of the comparison vowels were not in equal linear intervals because of the manner in which the OVE III synthesizer sets F_0. The OVE III generates a limited number of F_0 values. Specification of a specific frequency value provides only the synthesizer's closest approximation. We were aiming at equal spacing between our comparison vowels since pitch perception is approximately linear below 1000 Hz (Stevens et al., 1937).

Acknowledgements

The research reported in this paper was supported through NSF and USPHS grants to the UCLA Phonetics Laboratory. The stimuli were generated on the Haskins Laboratories OVE III Synthesizer. We would like to thank Terry Halwes and Cecilia Dewey for their assistance on the synthesizer. The stimuli were subsequently digitized on the UCLA Phonetics Laboratory PDP-12 Computer.

The authors thank the members of the UCLA Tone Seminar for their valuable comments. Special thanks are owed to Peter Ladefoged and Louis Goldstein for their criticism and discussion.

References

Divenyi, P. and Hirsh, I. 1975. 'The effect of blanking on the identification of temporal order in three-tone sequences' *Perception and Psychophysics* 17:246-252.

Elliot, L. L. 1962. 'Backward and forward masking of probe tones of different frequencies' *Journal of the Acoustical Society of America* 34:1116-1117.

Green, D. M. and Swets, J. A. 1966. *'Signal Detection Theory and Psychophysics'* John Wiley and Sons: New York.

Hombert, J. M. 1975. 'Perception of contour tones: an experimental investigation' *Procedings of the First Annual Meeting of the Berkeley Linguistics Society* pp. 221-232.

Massaro, D. W. 1972. 'Preperceptual images, processing time, and perceptual units in auditory perception' *Psychological Review* 79:124-145.

Ohala, J. and Ewan, W. 1973. 'Speed of pitch change' *Journal of the Acoustical Society of America* 53:345(A).

Samoilova, I. K. 1960. 'Masking of short tone signals as a function of the time interval between masked and masking sounds. *Biophysics* 4(5):44-52.

Stevens S. S., Volkmann J. and Newman, E. B. 1937. 'A scale for the measurement of the psychological magnitude pitch' *Journal of the Acoustical Society of America* 8:185-190.

Sundberg, J. 1973. 'Data on maximum speed of pitch changes' *Speech Transmission Laboratory Quarterly Progress and Status Report* (Stockholm) 4:39-47.