

# 21

## *THE PERCEPTION OF SECOND LANGUAGE PROSODY\**

Ellen Broselow

Richard R. Hurtig

Catherine Ringen

### INTRODUCTION

The question of the extent to which transfer from a first language affects the learning of subsequent languages continues to be a perplexing one. In recent years, researchers have questioned the importance of transfer, suggesting that if transfer operates at all in second language acquisition, it is a factor only in the acquisition of phonology, not of syntax (Richards, 1974; Dulay and Burt, 1974), and that even in the acquisition of second language phonology, transfer may play a role in the production of second language strings but not in the

\*An earlier version of this paper was presented at the Tenth International Congress of Phonetic Sciences at Utrecht, The Netherlands, in August 1983. We wish to express our thanks to Carol Binder, Jack Gandour, David Jung, Kee Ho Kim, Jocelyn Liu, James Tai, Ingo Titze, Andrea Tyler, Robert Wachal, Yuen-Mei Yin, and Hsio-Ching Yuan for their assistance and suggestions. We would also like to thank Pat Colsher and William Cooper for providing the fundamental frequency plots in Figure 21.2. That work was performed in Cooper's lab and was supported by NIH NS20071.

perception of the second language (Neufeld, 1980). If this is in fact the case, it suggests that transfer is a rather limited phenomenon, involving the transfer of motor skills rather than of linguistic competence.

It is the contention of this paper that transfer does play a role—in fact, a significant role—in the perception of a second language, and that what we might call “perceptual transfer” is an extremely interesting phenomenon, both because it makes certain aspects of the learning of a second language comprehensible, and because it reveals something of the speaker’s knowledge of his or her native language grammar. The phenomenon investigated in this paper is the perception of tone in sequences of Mandarin Chinese syllables by native speakers of English. The paper discusses the pattern of responses of English speakers presented with the task of identifying Mandarin tone sequences. We argue that the initially mysterious response pattern of English subjects becomes comprehensible when the Mandarin tonal system is compared with the corresponding system in the subjects’ native language—that is, when Mandarin tone is compared with English sentence intonation. While pitch functions very differently in Mandarin and in English, English speakers’ perception of Mandarin tone appears to be strongly influenced by their knowledge of the function of pitch in the intonation system of English.

The paper discusses two types of evidence that transfer plays a role in the perception of Chinese tone by English speakers. The first involves an examination of the relative ease with which each of the four Mandarin tones is perceived in different positions in strings of two and three syllables. Only one of the Mandarin tones exhibits any significant positional effect (aside from effects that are attributable to acoustic variations in the tones themselves), and this tone is the only one that is markedly similar in its acoustic properties to a common English intonation contour. This tone is perceived significantly better when it occurs in the position in which it is normally found in English sentences (namely, in final position). Thus this tone is more easily perceived when it occurs in the position in which subjects are accustomed to hearing it in their native language, and it is more difficult to perceive when it occurs in an unfamiliar context. The second type of evidence of transfer involves the sorts of errors that English speakers make when they misidentify the Mandarin tone that is phonetically equivalent to an English intonation contour. Again, this misidentification appears to be affected by the position in the string of the contour, and the pattern of misidentification suggests that English speakers tend to analyze this tonal contour in terms of English intonation. This analysis leads to a particular error pattern that can be understood by assuming that speakers analyze only part of the tonal contour as having lexical significance; the rest of the contour is presumably dismissed as part of the sentence intonation pattern. In both these cases, we argue, the performance of the English speakers is what we would expect if we assume that they are perceiving the second language strings in terms of their native language phonological system.

## METHOD

To examine the ability of English speakers to identify Mandarin tones, we presented a tape made by a female Mandarin speaker from Taiwan to 50 students in an introductory linguistics course at the University of Iowa. All these students were native speakers of English and none had studied a tone language. The tape consisted of both training and identification tasks. The four lexical tones of Mandarin Chinese were first introduced with the syllable *ti*, and this syllable was produced three times using each tone. Following this very brief training, the subjects were directed to identify the tones of individually presented syllables. The order of the tones was randomized, with each tone appearing three times. The procedure was then repeated with a different syllable (*ma*). The subjects were then asked to identify tones presented in series of two and three syllables (a total of 56 syllables in doublets and 72 syllables in triplets). In the doublets and triplets, each tone occurred the same number of times in each position and with each of the other tones. In the doublets, each tone was also paired with itself. A single example was presented prior to each set of identification trials (6 for singlets and triplets, 7 for doublets), and subjects were given feedback after the presentation of singlets and doublets. No other training was given. Two native speakers of Mandarin were asked to identify the tone on the tape as a control. Their judgements coincided exactly with those of the speaker who had made the tape.

### Mandarin Tones

The tape provided no acoustic description of the four tones, identifying them only as first, second, third, or fourth tone. The traditional representation of these tones is presented in Figure 21.1.

The numerical descriptions presented in Figure 21.1 assume a fundamental frequency range extending from a high point of five to a low point of one. The first one, the only level tone of the series, is spoken on the highest pitch and neither rises nor falls. The second tone begins in the middle range and rises. The third and fourth tones both involve falling contours. The third tone begins quite low and falls to the lowest point of the register, where it is often associated with laryngealization. In absolute final position it also contains a rise, indicated by the parenthesized four. The fourth tone also falls, but it begins at a much higher point and falls much more sharply than the third tone.

Pitch in Mandarin is phonemic; changing the pitch of a syllable can change its meaning. Thus the tone of a Mandarin word must be part of its lexical entry. Pitch is used quite differently in English, however, where the pitch pattern of a string may be used to convey such things as the emotional state of the speaker or whether the string is intended to be a question or a statement. However, the Mandarin fourth tone is acoustically quite similar to a common

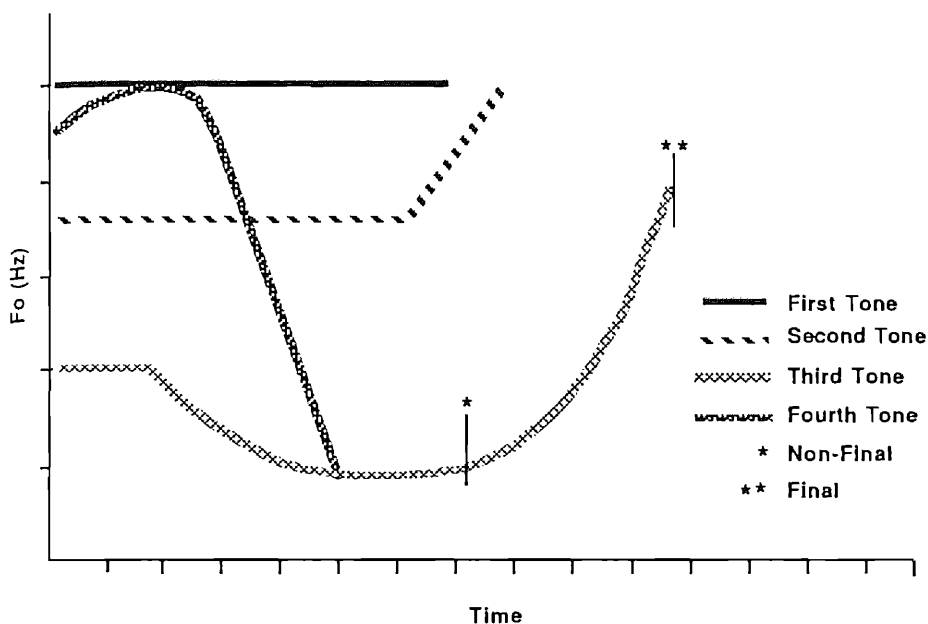


Figure 21.1

First Tone:	High level	55	MA	'mother'
Second Tone:	Rising	35	MA	'hemp'
Third Tone:	Falling-rising	21(4)	MA	'horse'
Fourth Tone:	Falling	51	MA	'scold'

English intonational contour. The unmarked pattern for declaratives in English, traditionally called “rising-falling,” involves a rising pitch on the so-called “tonic” syllable—the rightmost pitch-accented syllable—followed by a fall on all material following the tonic syllable. This contour is represented below:

My name is Mary.

When the tonic syllable of a sentence is also the last syllable of that sentence, then both the high pitch generally associated with the tonic syllable and the fall associated with material following the tonic syllable are realized on that single syllable, as below:

My name is John.

This rising-falling contour is also typical of single-word utterances spoken as declaratives, as in *a*, and of the final item in a series, as in *b*:

- a.* What's your name? John.
- b.* I invited Susan, Mary, and John.

Tracings of the fundamental frequency showing this rising-falling contour are given in Figure 21.2.

The phonetic similarity between the pitch contour of a Mandarin syllable spoken on a fourth tone and a sentence-final tonic syllable in an English declarative accounts for the impression of many English speakers that the fourth tone is the only Mandarin tone that sounds “normal.” This phonetic similarity also accounts for the production error mentioned by Chiang (1979), who reports that English learners of Chinese tend to incorrectly use falling (that is, fourth) tone on syllables that occur at the end of Mandarin sentences. Chiang attributes this tendency to interference from English intonation patterns. We argue that the similarity between Mandarin fourth tone and English final declarative intonation also accounts for the error pattern revealed in our study.

## RESULTS

### Data Analysis

The subjects’ response to each tone were cast into standard confusion matrices. Thus for each of the four tones in each of the three conditions (singlets, doublets and triplets) and for each of the serial positions in the doublets and triplets we obtained the percentage of correct identification as well as the percentage of

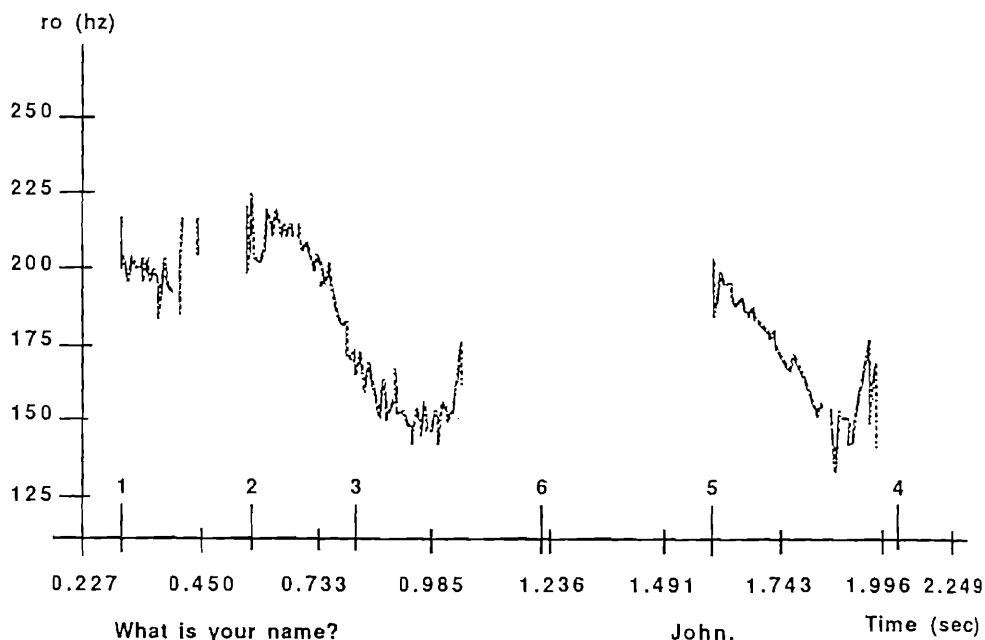


Figure 21.2

misidentification. From these data we computed the index of detectability ( $d'$ ) and response bias ( $\beta$ ):

$$d' = z(X|\bar{X}) - z(X|X)$$

$$\beta = f(X|X)/f(X|\bar{X})$$

We utilized the statistic proposed by Gourevitch and Galanter (1967) to assess the significance of the differences that we obtained in the detectability of the tones.

### Relative Detectability of Mandarin Tones

If fourth tone is the only Mandarin tone readily perceived in terms of English intonation patterns, we might expect it to be perceived differently from the other three tones of Mandarin. An examination of the relative detectability of the four tones in different positions does reveal some interesting differences between the perception of fourth tone and the perception of the other tones. As Table 21.1 shows, in a task involving the perception of single tones, fourth tone is the most easily perceived tone. Its index of detectability ( $d' = 3.44$ ) is significantly different from that of all the other tones, ( $p < .02$ ). This is shown graphically in Figure 21.3.

Thus the fourth tone appears to be the easiest tone for English speakers to identify—a property we might ascribe to its combination of a very high beginning point and very steep fall. This result is consistent with the results of the various studies of perception of Mandarin tone reviewed in Gandour (1978). These studies confirm that both native speakers and learners of Mandarin find tones one and four easiest to identify, while tones two and three, which are acoustically most alike, are most often confused. All these studies investigated the perception of tones in single syllables, and their findings suggest that the perceptibility of a tone is a direct consequence of its acoustic properties. However, when we examine the percent of correct responses for longer strings, a simple account of detectability only in terms of acoustic properties is not maintainable.

As Table 21.2 shows, detection of the fourth tone declines dramatically in nonfinal position. In fact, fourth tone in the middle position of a triplet

TABLE 21.1. Single Tones

	% Correct	$d'$
Tone 1:	81%	2.44
Tone 2:	67%	1.72
Tone 3:	78%	2.18
Tone 4:	94%	3.44

TABLE 21.2. Percent Correct Responses: Strings of Two and Three Syllables

	Position				
	1	2	1	2	3
Tone 1:	70%	69%	65%	52%	61%
Tone 2:	55%	47%	47%	48%	40%
Tone 3:	49%	71%	41%	47%	61%
Tone 4:	31%	79%	49%	28%	77%

actually has, at 28 percent, the lowest percent correct score of any tone. Note that while the percent of correct responses declines for all tones in longer strings, the pattern of correct responses for tone four is very different from that of tones one and two. First and second tones show no dramatic positional difference in the percent correct responses in two- and three-syllable strings; as Table 21.3 shows, the indices of detectability of these tones in different positions are fairly close, and the differences between them are not statistically significant. Only tones three and four show a significant difference in detectability between different positions ( $p < .01$ ). For both these tones, detectability in final position—that is, the second position in two-syllable strings and the third position in three-syllable strings—is significantly greater than detectability in nonfinal positions. This is illustrated graphically in Figure 21.3.

The greater detectability of third tone in final vs. non-final position has a straightforward explanation in the facts of Mandarin. The third tone is really produced differently in final position; an examination of narrow band sonagrams and Fo tracings of our tape confirmed that the third tone has a marked rise in final position that is generally absent in non-final positions, as indicated in Figure 21.1. Speakers of Mandarin have of course learned to ignore this allophonic variation in the third tone, but our subjects, whose training exposed them only to the final variant of third tone, predictably have difficulty in recognizing the non-final variant.

No such explanation is available, however, for the positional differences in the detection of the fourth tone, which has essentially the same acoustic properties in all positions, yet is identified correctly far more often in final than

TABLE 21.3. D': Strings of Two and Three Syllables

	Position				
	1	2	1	2	3
Tone 1:	1.71	1.91	1.47	1.13	1.56
Tone 2:	.81	.92	.60	.45	.70
Tone 3:	.85	1.63	.85	.76	1.36
Tone 4:	.98	2.46	1.31	.83	2.02

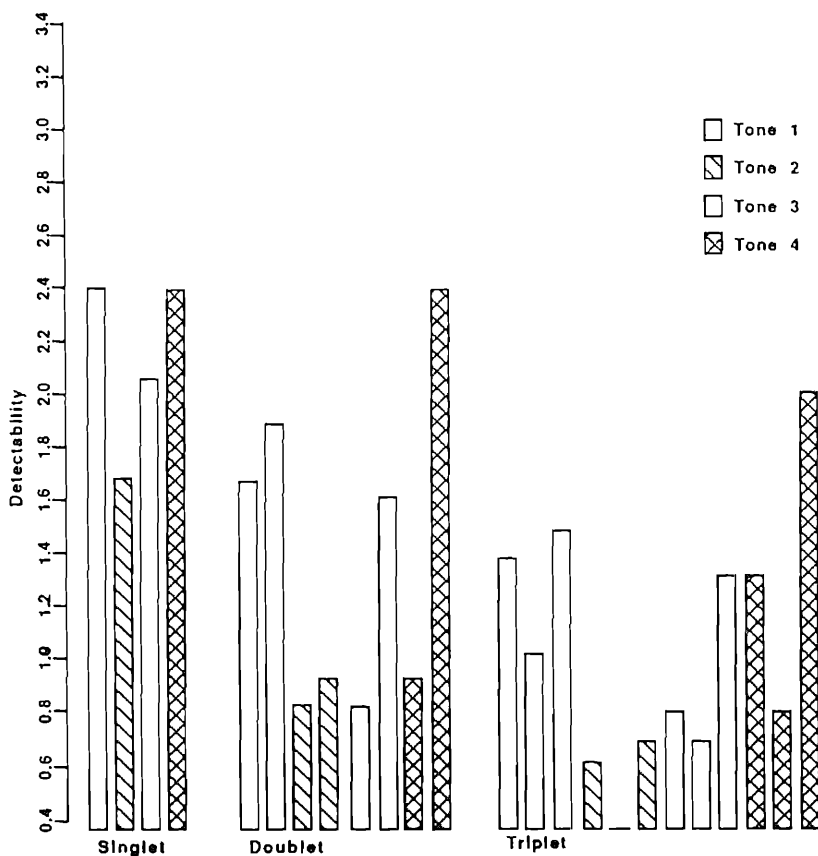


Figure 21.3

in non-final position. This effect, we argue, is a result of interference from English intonation. The contour of the fourth tone is a familiar one, since it corresponds to the fall at the end of declaratives in English. When the fall occurs in final position, then, it is easily recognized as something the listener is accustomed to hearing. In non-final position, however, fourth tone sounds as odd to an English speaker as would an English phoneme occurring in a position in which it does not normally occur in English, such as word-initial velar nasal or word-initial *t* followed by *l*. The relative ease of detectability of the fourth tone in final and non-final positions, then, appears to be a clear result of transfer from the native language. The relative ease of detectability of the fourth tone in string-final position is a result of positive transfer: a familiar item occurring in a familiar position is easy to hear. The difficulty of hearing the fourth tone in non-final position, on the other hand, is a result of negative transfer; the high falling pitch contour normally does not occur in non-final position in English strings.



## The Misidentification of the Fourth Tone in Final Position

In the preceding section we argued that English speakers are better at identifying the fourth tone in final than in non-final position because in this position the fourth tone corresponds to something that is familiar to them from their native language. We now turn our attention to the pattern of errors that appears when the fourth tone is misidentified. A wealth of anecdotal evidence testifies to a rather surprising pattern of errors: In strings of Mandarin syllables, students of Mandarin tend to identify the fourth (high falling) tone as the first (high level) tone,<sup>1</sup> although the fourth tone and the first tone are otherwise the most distinguished tones.

The answer to the question of why fourth tone should be misidentified as first tone more often in final than in non-final position is fairly obvious once we recall that the contour of the fourth tone is virtually identical to the pitch contour found at the end of an English declarative sentence with a sentence-final tonic syllable. In English, the high pitch on the tonic syllable is linguistically significant; the location of this high pitch indicates which element of the sentence is focussed. Thus the English speaker is accustomed to listening for the high portion of the rising-falling contour, since its location carries semantic information. The fall that follows the high pitch, however, is simply what comes after the tonic syllable. English speakers who identify a final fourth tone as a first tone are apparently dismissing the fall at the end of the string as redundant—part of the sentence contour rather than associated with any particular syllable.

Thus the identification of falling tone as high level tone at the end of a string is interpretable as a consequence of transfer from the native language: English speakers may ignore the final fall because they analyze it as the fall normally associated with the end of a declarative. They identify the starting point of the fall as the portion of the tone contour that is lexically significant—that is, associated with the fourth-tone syllable itself rather than with the sentence contour. This error pattern bears in an interesting way on a controversy in linguistic theory: the question of whether pitch contours are to be represented on the underlying level in terms of contours, such as rising and falling, or whether these surface contours are actually, at a deeper level of analysis, a series of level pitches, such as high and low. A number of convincing analyses of contour tones in terms of level pitches have been offered (for example, Goldsmith, 1976). Analyses of English intonation contours in terms of level pitches have been proposed by Pike (1945), Trager and Smith (1951), among others, and most recently by Liberman (1978) and Pierrehumbert (1980). In all these analyses, the phonetic pitch contours of English intonation represent the surface realization of an underlying pattern of level pitches, with pitch changes merely the necessary transitions from one level pitch to the next.

Given this analysis of English intonation, along with the assumption that English speakers tend to analyze the strings of Mandarin tones in terms of their

native language system, the identification of a final falling tone as a high level tone by English speakers makes perfect sense. The fourth tone begins at the highest point of the register but falls immediately to the bottom of the pitch range, while the first tone is a sustained high tone. Thus what the fourth and first tones have in common is their starting points. If the fall characteristic of the ends of English declaratives is, at a deeper level of analysis, composed of two level pitches, a high pitch followed by a low pitch, it is not at all surprising that English speakers should exhibit a tendency to identify the high falling tone as a high level tone in final position. English speakers hearing the Mandarin fourth tone in final position disregard the fall, analyzing it as the surface reflex of a low tone associated with the sentence boundary—that is, an effect of the intonation contour of the sentence—and identify what remains when this sentence contour is stripped away as a high tone. The final fall is apparently decomposed into a high tone, associated with the final syllable, and a low tone, associated with the sentence boundary, and therefore irrelevant to the pitch of the syllable itself. Thus an analysis of the pitch contours of English intonation as the surface realization of an underlying series of level pitches is entirely consistent with the error of identifying final fourth tone as first tone. Confirmation of the assumption that English speakers tend to analyze foreign language pitch phenomena in terms of level pitches is provided by a recent study (Gandour, 1983), which showed that English speakers given the task of rating the dissimilarity of pairs of tones relied far more on the dimension of tone height than on a dimension of direction of change. It should be noted that the alternative theory of English intonation, which posits pitch contours such as rise and fall as underlying units of representation, contributes nothing to our understanding of the identification of final fourth tone as first tone, since this theory offers no explanation for the reason it should be the starting point of the falling contour that is identified as the portion associated with the final syllable, or why a contour tone should be misidentified as a level tone.

Although our subjects made few errors on fourth tone in final position, an examination of the pattern of the response bias (see Figure 21.4) reveals a pattern consistent with the anecdotal evidence that English speakers learning Mandarin Chinese tend to identify fourth tone as the first tone in final position. As predicted, the difference in the response bias between final and nonfinal position as shown in Table 21.4 was greatest for first tone; note also the complementary decrease in  $\beta$  for fourth tone in final position.

## CONCLUSION

We have argued that transfer from the intonational system of English accounts for two aspects of English speakers' perception of Mandarin Chinese tones. First, we found that subjects in this study found the fourth tone significantly easier to perceive when it occurred in a position in which the phonetically

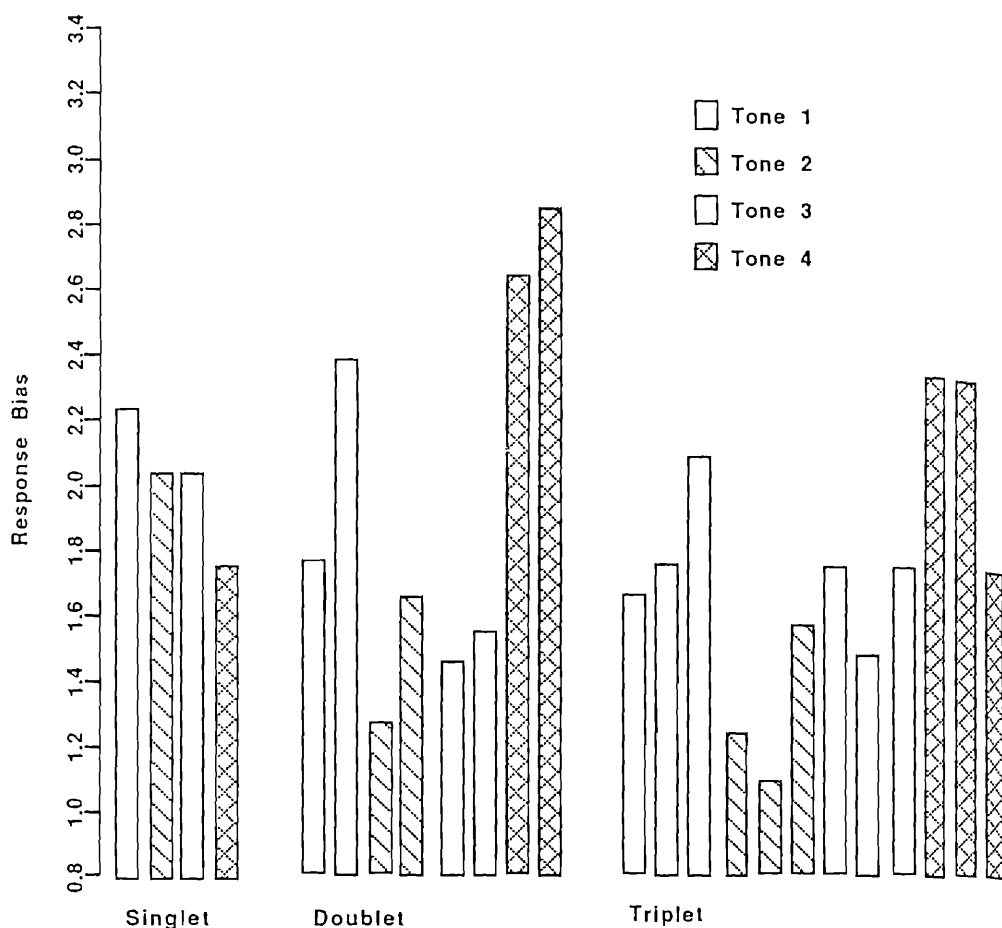


Figure 21.4

similar English pitch contour appears, and that it is significantly more difficult to perceive when it occurred in a position in which it would be anomalous in English. The fourth tone is both the only tone that is markedly similar to an English intonational contour and the only tone that shows a significant posi-

TABLE 21.4. Response Bias  $\beta$

	Position					
	1	2		1	2	3
Tone 1:	1.74	2.39		1.66	1.79	2.18
Tone 2:	1.25	1.64		1.26	1.13	1.53
Tone 3:	1.47	1.54		1.74	1.42	1.72
Tone 4:	2.64	2.81		2.27	2.28	1.73

tional effect. We argued that these two facts can be related once the effect of language transfer is taken into account. Second, we have shown that the response bias is consistent with the anecdotal evidence that a final fourth tone is likely to be heard as a first tone. We can understand the identification of a final fourth tone as a first tone in terms of transfer by assuming that the English speaker analyzes the fourth tone as composed, like the English rising-falling intonation contour, of a redundant fall associated with the sentence boundary preceded by a high tone associated with the syllable itself. Thus the concept of transfer provides an explanation of an otherwise puzzling array of facts (i.e. the positional variation in the detectability of the fourth tone) and it suggests a means of testing the predictions of competing linguistic theories.

## NOTE

1. This phenomenon was noticed by one of the authors in several phonetics classes in which students were taught to identify Mandarin tones. The persistence of this error was confirmed by teachers and by English-speaking learners of Mandarin. We are indebted to Susan Schmerling for the report that the entire phonetics class in which she was a student made this error on a final exam.

## REFERENCES

- Chiang, T. 1979. "Some Interferences of English Intonations with Chinese Tones," *IRAL*, 17, 245-250.
- Dulay, H., and M. Burt. 1974. "Errors and Strategies in Child Second Language Acquisition," *TESOL Quarterly*, 8, 129-136.
- Gandour, J. 1977. "Counterfeit Tones in the Speech of Southern Thai Bidialectals," *Lingua*, 41, 125-143.
- Gandour, J. 1978. "The Perception of Tone," in V. Fromkin (Ed.), *Tone: A Linguistic Survey*. New York: Academic Press.
- Gandour, J. 1983. "Tone Perception in Far Eastern Languages," *Journal of Phonetics*, 11, 149-175.
- Gandour, J., and R. Harshman. 1978. "Crosslanguage Differences in Tone Perception: A Multidimensional Scaling Investigation," *Language and Speech*, 21, 1-33.
- Goldsmith, J. 1976. "An Overview of Autosegmental Phonology," *Linguistic Analysis*, 2, 23-68.
- Gourevitch, V., and E. Galanter. 1967. "A Significance Test for One Parameter Isosensitivity Functions," *Psychometrika*, 32, 1, 25-33.
- Kirilloff, C. 1969. "On the Auditory Perception of Tones in Mandarin," *Phonetica*, 20, 63-67.
- Neufeld, G. 1980. "On the Adult's Ability to Acquire Phonology," *TESOL Quarterly*, 14, 285-298.
- Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. Doctoral dissertation, MIT.
- Pike, K. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Richards, J. 1974. "A Non-Contrastive Approach to Error Analysis," in J. Richards (Ed.), *Error Analysis*. London: Longman.
- Selkirk, E. 1984. *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, MA: MIT Press.
- Trager, G., and H. Smith. 1951. *Outline of English Structure*. Norman, OK: Battenburg Press. (Revised edition 1957, Washington: American Council of Learned Societies.)
- Tseng, Chiu-Yu. 1981. *An Acoustic-Phonetic Study of Tones in Mandarin Chinese*. Doctoral dissertation, Brown University.