A study of sentence stress production in Mandarin speakers of American English

Yang Chen[a)]
Division of Speech-Language Pathology and Audiology, University of Wyoming, Laramie, Wyoming 82071

Michael P. Robb, Harvey R. Gilbert, and Jay W. Lerman
Department of Communication Sciences, University of Connecticut, Storrs, Connecticut 06269

(Received 7 August 2000; accepted for publication 12 January 2001)

Acoustic characteristics of American English sentence stress produced by native Mandarin speakers are reported. Fundamental frequency (F_0), vowel duration, and vowel intensity in the sentence-level stress produced by 40 Mandarin speakers were compared to those of 40 American English speakers. Results obtained from two methods of stress calculation indicated that Mandarin speakers of American English are able to differentiate stressed and unstressed words according to features of F_0, duration, and intensity. Although the group of Mandarin speakers were able to signal stress in their sentence productions, the acoustic characteristics of stress were not identical to the American speakers. Mandarin speakers were found to produce stressed words with a significantly higher F_0 and shorter duration compared to the American speakers. The groups also differed in production of unstressed words with Mandarin speakers using a higher F_0 and greater intensity compared to American speakers. Although the acoustic differences observed may reflect an interference of L1 Mandarin in the production of L2 American English, the outcome of this study suggests no critical divergence between these speakers in the way they implement American English sentence stress. [DOI: 10.1121/1.1356023]

PACS numbers: 43.70.Fq [AL]

I. INTRODUCTION

Individuals who speak English as a second language (L2) vary in their ability to produce phonetic features of English precisely. Phonetic characteristics of the native or first language (L1) are thought to interfere with the production of L2 (Cheng, 1987). The L1 interference with L2 production can occur at both the segmental and suprasegmental level (Ingram and Pittam, 1997; Shen, 1990). The more interference which exists between L1 and L2, the less likely phonetic features of the L2 will be produced accurately. There have been many acoustic and perceptual studies examining the effects of various Asian languages on the phonetic features of English spoken as L2 (Flege, 1989; Flege and Davidian, 1985; Hutchinson, 1973; Ingram and Park, 1997; Ingram and Pittam, 1987; Kim, 1972; Magnuson and Akahane-Yamada, 1996; Pittam and Ingram, 1992; Robson, 1982; Tarone, 1980). The magnitude of phonetic inaccuracies appears to be correlated to the amount of experience (or length of time) speaking English, or the age period during which L2 was acquired (Flege, 1995; Guion *et al.*, 2000; Johnson and Newport, 1991). It is believed that after a ''critical period,'' L2 learners have difficulty not only accurately articulating L2 segments (Flege, 1987; Lennenberg, 1967; Magnuson and Akahane-Yamada, 1996; Pittam and Ingram, 1992), but also in acquiring the suprasegmental features of L2 (Chun, 1982; Guion *et al.*, 2000; Guzma, 1973; Scuffil, 1982).

One aspect of American English noted to be of difficulty for native speakers of tonal languages, specifically East Asian languages, is the production of stress placed on syllables or words (Cheng, 1968, 1987; Chun, 1982). Wijk (1966) noted over 30 years ago that correct stressing of words presents a major difficulty in the pronunciation of English for individuals who learn English as L2. There are two forms of stress which occur in the production of American English: lexical stress and sentence stress. Lexical stress is concerned with the emphasis of individual syllables comprising a polysyllabic word. Sentence stress is concerned with the stress placed on words in order to indicate (or contrast) their importance/prominence in relation to other words in a sentence. Varying degrees of syllable/word stress are indicated by changes in vocal fundamental frequency (F_0), vowel duration, and vowel intensity (Berinstein, 1979; Bolinger, 1958; Crystal, 1969; Fry, 1958; Potisuk *et al.*, 1996).

The focus of the present study was on the production of American English sentence stress by native speakers of Mandarin.[1] Clear differences exist in the segmental and suprasegmental characteristics of Mandarin and English (Chao, 1948, 1968; Cooper and Sorensen, 1981; Ho, 1976; Kratochvil, 1962; Lehiste, 1970; Pike, 1945; Tseng, 1981). Such being the case, we wished to determine whether L1 Mandarin interferes with the production of sentence stress in L2 English. Prior to undertaking this research, literature related to sentence stress production in American English and Mandarin was reviewed. Results of this review were used to formulate specific research predictions. The literature review and predictions follow.

A. Sentence stress in English

English is a polysyllabic language with diverse syllable structure. English is often described as a stress-timed language, whereby the speech rhythm of English involves an

a)Author to whom correspondence should be addressed. Electronic mail: yc@uwyo.edu

interplay of prominent (or long duration) syllables and less prominent (or short duration) syllables (Chun, 1982; van Santen and Shih, 2000).[2] English is also a nontone language whereby the meaning of a word is derived from the phonetic composition of the word. In the case of English, tones are part of what is usually called intonation, and can be spread across any number of syllables (Chun, 1982; Clark and Yallop, 1995). The primary acoustic correlate of a tone is the F_0 contour and in English the fundamental tonal choice is between rise and fall. The tone is normally placed on the last word of an utterance to convey either a statement (''She gave him the *KEYS*.'') or a question (''She gave him the *KEYS*?''). However, tones can also be placed on virtually any syllable within a sentence (i.e., sentence stress), in which case the tone has contrastive value (''He doesn't live *IN* Auckland'' [but nearby]) (Clark and Yallop, 1995).

There have been several examinations of the sentence stress of English (Brown and McGone, 1974; Crystal, 1969; Fromkin and Ohala, 1968; Ohala and Hirano, 1967; Lieberman, 1960, 1967; Pike, 1945). Notable are the studies reported by Cooper and his colleagues (Cooper *et al.*, 1985; Cooper and Sorensen, 1981; Eady and Cooper, 1986; Eady *et al.*, 1986). Results from a majority of studies examining sentence stress indicate that the most consistent acoustic correlates underlying stress are an increase in magnitude of F_0, duration, and intensity of stressed words in relation to unstressed words. However, an increase in magnitude of each acoustic correlate is not a necessary requirement for stress to occur. Liberman (1967) noted that typically two of these three correlates show an increase in magnitude during production of sentence stress.

B. Sentence stress in Mandarin

A syllable in Mandarin consists of segmental and suprasegmental features (Shih, 1986). Segmental features consist of a vocalic nucleus which may or may not be accompanied by pre-vocalic and post-vocalic consonants (Chao, 1968). The syllable structure of Mandarin, unlike English, is monosyllabic with primarily a basic CV word shape (Guo, 1992). Suprasegmental features include four ''basic'' tones[3] (Cheng, 1968; Gandour, 1978; Howie, 1976; Leben, 1978). In a tone language such as Mandarin, a change in tone of a syllable (i.e., a word) leads to a change of lexical meaning (Chao, 1948; Cheng, 1968; Chun, 1982; Dreher and Lee, 1966; Tseng, 1981; Pike, 1948). Mandarin is also a syllable-timed language and is generally thought of as showing no strong pattern of stress, with syllable duration remaining relatively constant across a sentence (Clark and Yallop, 1995; van Santen and Shih, 2000).

There are perceptual and acoustic studies which demonstrate that sentence stress occurs in the production of Mandarin. For example, Chao (1932) and Chun (1982) examined sentence stress as perceived in the production of Mandarin phrases (i.e., a sequential CV production). Results from both studies indicated that sentence stress in Mandarin can be achieved through differential use of F_0 or by lengthening of syllable duration, depending on the position of the stressed word in the sentence. The acoustic studies performed by Lin and Yan (1980) and Shen (1993) found Mandarin speakers to produce stressed words with higher intensity compared to unstressed words. Cao (1986) and Shen (1993) reported lengthening of stressed words compared to unstressed words, and Xu (1999) found Mandarin speakers to vary F_0 during the production of Mandarin phrases to differentiate stressed from unstressed words.

C. Sentence stress in Mandarin speakers of American English

When Mandarin speakers produce English, they may encounter difficulty determining the appropriate location for stress within a sentence (Cheng, 1987; Chun, 1982). For example, Chun (1982) reported that L1 Mandarin speakers learning English were perceived to either misplace sentence stress or produce stressed words with abnormally short durations. Chun (1982) attributed the findings to the lack of a fixed pitch pattern in Mandarin sentences and to the syllable-timed nature of Mandarin, which appeared to interfere with the production of English. To date, information related to English sentence stress production by Mandarin speakers is limited to perceptual judgments. An acoustic examination of the production of American English sentence stress by native speakers of Mandarin remains to be performed.

Based on what is known about the characteristics of sentence stress in the English and Mandarin languages, three predictions were made regarding the production of English sentence stress by native speakers of Mandarin. The predictions were developed with reference to the specific influence of F_0, duration, and intensity on sentence stress. The first prediction was that Mandarin speakers will not significantly differ from American English speakers in their use of F_0 to signal stress in the production of English sentences. Although tonal languages such as Mandarin are assumed to alter F_0 exclusively to signal a change in word meaning rather than intent (Chao, 1948; Howie, 1976), the recent report by Xu (1999) would suggest otherwise. The second prediction was that Mandarin speakers will show no significant differences from American English speakers in their use of duration to signal English sentence stress. Being a syllable-timed language, Mandarin might be expected to show minimal variation in syllable duration across a phrase; however, Cao (1986) and Shen (1993) have found contrary evidence. The third prediction was that Mandarin speakers will show no significant differences from American English speakers in their use of intensity to signal sentence stress. Mandarin speakers have been shown to alter intensity to signify stress in phrases (Lin and Yan, 1980; Shen, 1993). It remains to be determined whether native Mandarin speakers are capable of using intensity to signify stress in English sentences.

II. METHOD

A. Participants

Two groups of subjects were recruited from within the University of Connecticut community (Storrs, CT). The first group consisted of 40 adults (20 males, 20 females) who spoke Mandarin as L1 and American English as L2. All Mandarin subjects were born in mainland China. The average age of the Mandarin males was 33 years

(range=30–46 years). The average age of the Mandarin females was 28 years (range=21–42 years). Selection criteria for inclusion in the Mandarin group consisted of: (1) a university education; (2) formal instruction in English; (3) the ability to speak standard Mandarin as judged by the first author (who is a native speaker of Mandarin); (4) speaking English for a minimum of 30% of their daily conversation; (5) the ability to orally read English fluently; and (6) residence in the U.S. for a minimum of 2 years. The average length of U.S. residence was 3 years, 9 months for Mandarin female speakers (range=2–12 years), and 4 years, 7 months for Mandarin male speakers (range=2–17 years). The second group consisted of 40 adults (20 males, 20 females) who spoke American English as L1. The average ages of the American male and female speakers were 33 years (range =22–46 years) and 27 years (range=23–41 years), respectively. All subjects (native Mandarin and English) were judged to be free of speech, language, or hearing disorders.[4]

B. Speech materials

Sentence stress was evaluated during production of the sentence, ''I bought a cat there.'' The sentence was produced four ways with primary stress placed on one of four different words. The four sentences with varying stress (noted in uppercase italics) were: ''*I* bought a cat there,'' ''I *BOUGHT* a cat there,'' ''I bought a *CAT* there,'' and ''I bought a cat *THERE*.'' A similar procedure was used by Copper and Sorenson (1981). Each of the 4 sentences was produced 3 times by each participant for a total of 12 sentences per participant. Individuals were allowed to practice the speech materials as often as they wished prior to the actual audio recordings. A sentence was deemed acceptable if it was perceptually fluent, contained no misarticulations, conformed to the prescribed placement of primary stress, and was judged by one of the researchers to be produced at a conversational pitch and loudness. For the present study, sentence stress was defined as the stressed placed on the vocalic nucleus of each target word.

C. Audio recordings

All recordings took place in a sound-attenuated booth. Order of sentence presentation was randomized across participants. Audio recordings were made using a cassette recorder (Marantz PMD-360) and a unidirectional dynamic microphone (Shure, 515SD). The microphone was placed approximately 20 cm from the speaker's mouth. Three pure tones of varying intensities were recorded on audiotape for later intensity calibration. A sound level meter was placed alongside the microphone for recording the actual sound intensity level. Gain control settings remained in fixed positions during the recording of the calibration and speech signals.

D. Acoustic analysis

Each of the four sentences, along with the corresponding intensity calibration signals, were introduced via a cassette recorder (Nakamichi, MR-2) into a computer (CPU586). All acoustic signals were digitized at 10 000 Hz using a speech analysis system (CSL 4300B). Each sentence produced by each subject was displayed as an amplitude-by-time waveform. The four target words in each sentence recitation were measured for F_0, intensity, and vowel duration. Measurements of F_0, intensity, and duration were made only for the vowel nucleus of the four target words. The measurements performed were as follows.

1. Fundamental frequency (F_0)

Three portions of the vocalic nucleus were measured for F_0. A 50-ms windowed cursor was placed at beginning, midpoint, and end locations of the vowel segment, respectively. To minimize possible coarticulatory influences of pre- and post-vocalic consonants on the vowel of interest, the beginning of the vowel was defined as the first 50 ms of periodic activity following the third glottal pulse. The end of the vowel was defined as the last 50 ms of periodic activity preceding the last three glottal pulses. Vowel midpoint was the mathematical middle point between the beginning and end points. Once the 50-ms window was fixed on the waveform, the signal was transformed into an amplitude-by-frequency ''power'' spectrum using a Hamming window weighting. The center frequency of the first harmonic peak was extracted to represent the F_0. Instances when the resulting F_0 value was thought be questionable were handled by re-examining the vowel nucleus using a narrow band (45-Hz) spectrogram. The F_0 determined for each vocalic nucleus was based on the median of the F_0 values identified at the three measurement locations. Group data were calculated as an average of the median values.

2. Vowel intensity

Five data points (i.e., dB values) were extracted from each of the three calibration tones, and a calibration equation was then generated by using a linear regression method (Ng, 1996). The calibration equation was used to calculate the intensity level in dB SPL during the speech sample recordings. The onset and offset of each vowel sample was demarcated using the same beginning and end locations used to measure the F_0 for each target word. The demarcated vowel was then transformed into an energy-by-time plot. The same three vowel portions used for F_0 measurement were extracted from the energy-by-time plot and the intensity levels were obtained. These intensity values were then ''corrected'' on the basis of the calibration equation to determine the true intensity values. The intensity value for each vocalic nucleus was based on the median of the intensity values identified at the three measurement locations. Group data were calculated as an average of the median values.

3. Vowel duration

Measurement of vowel duration for each target word was based on the amplitude-by-time waveform display. Vertical cursors were manually placed at the first and last glottal pulses of the vowel to demarcate the onset and offset points. The time interval between the two cursors was taken to be the vowel duration.

TABLE I. Mean (M), standard deviation (s.d.), and range values for F_0 (Hz), intensity (dB), and duration (ms) for each stressed word produced by the group of American (A) female and Mandarin (M) female speakers.

Group	Stressed word	F_0 (Hz)			Intensity (dB)			Duration (ms)		
		M	s.d.	Range	M	s.d.	Range	M	s.d.	Range
A-Female	*I*	219	28	166–303	61	7	43–75	277	89	103–508
	BOUGHT	225	34	166–332	63	7	44–75	209	44	125–305
	CAT	232	52	168–491	58	8	38–76	160	29	113–247
	THERE	201	27	146–285	56	9	39–73	335	48	238–454
M-Female	*I*	257	28	215–329	63	13	37–83	265	75	70–437
	BOUGHT	258	29	204–336	65	11	42–83	207	48	111–370
	CAT	270	31	222–336	60	13	41–80	160	43	71–290
	THERE	237	27	182–290	58	14	29–82	344	68	211–503

E. Sentence stress calculation

The acoustic measurements of each speaker's production of sentence stress were analyzed using two different calculations. The first calculation involved computing mean values of F_0, intensity, and duration for stressed words only. This was referred to as *average* sentence stress. Analysis of average sentence stress provided absolute values of the acoustic features for stressed words. This analysis alone might not indicate whether Mandarin speakers produced sentence stress in a manner similar to American speakers. Laver (1994) indicated that a stressed syllable/word is one that is produced more prominently than another syllable/word. Therefore, comparison of the acoustic characteristics of stressed and unstressed words would be necessary to comprehensively evaluate the contrastive (or differentiated) production of sentence stress.

The second calculation of sentence stress was designed to evaluate differentiated stress. This calculation involved comparing the acoustic characteristics of a stressed word (e.g., stressed ''*BOUGHT*'') to the same but unstressed words (e.g., the remaining unstressed ''bought'' productions) produced across the four different sentences. This was referred to as *across-sentence* stress. A ''difference'' value of an acoustic parameter (F_0, intensity, or duration) was obtained which indicated the change associated with the stressed word compared to the average of the remaining unstressed words. For example, to calculate the across-sentence $F_0(\Delta F_0)$ for the word ''bought'' across the four sentences:''*I* bought a cat there,'' ''I *BOUGHT* a cat there,'' ''I bought a *CAT* there,'' and ''I bought a cat *THERE*,'' the formula is:

$$\Delta F_0 = F_{0\ BOUGHT} - (F_{0\ \text{bought}} + F_{0\ \text{bought}} + F_{0\ \text{bought}})/3.$$

F. Measurement reliability

Ten percent of the sentence tokens (96 sentences) were randomly selected across Mandarin and American speakers (i.e., four Mandarin and four American speakers) for assessment of intra-judge and inter-judge measurement reliability. The first author remeasured the 96 sentences for intra-judge reliability assessment, while inter-judge reliability assessment was performed by another individual who was experienced in acoustic measurements. Average absolute errors for F_0, intensity, and vowel duration for intra-judge measurement were 4.10 Hz, 0.55 dB, and 1.82 ms, respectively. Pearson correlation coefficients for F_0, intensity, and vowel duration between the first and second measurements were 0.95, 0.79, and 0.97 ($p<0.01$), respectively. Average absolute errors for F_0, intensity, and vowel duration for inter-judge measurements were 1.48 Hz, 2.01 dB, and 26.88 ms, respectively. Pearson correlation coefficients for measurement of F_0, intensity, and vowel duration between the two judges were 0.79, 0.79, and 0.91 ($p<0.01$), respectively.

III. RESULTS

A. Average sentence stress

The F_0, duration, and intensity values for average sentence stress are shown in Table I for American and Mandarin

TABLE II. Mean (M), standard deviation (s.d.), and range values for F_0 (Hz), intensity (dB), and duration (ms) for each stressed word produced by the group of American (A) male and Mandarin (M) male speakers.

Group	Stressed word	F_0 (Hz)			Intensity (dB)			Duration (ms)		
		M	s.d.	Range	M	s.d.	Range	M	s.d.	Range
A-Male	*I*	140	29	87–222	61	7	46–77	299	82	163–471
	BOUGHT	144	25	106–203	67	6	54–79	208	44	118–327
	CAT	154	32	106–242	62	6	43–74	170	39	103–297
	THERE	127	26	121–203	61	6	41–74	350	66	218–488
M-Male	*I*	160	23	104–215	63	12	35–82	258	92	78–530
	BOUGHT	163	22	126–215	68	11	39–84	191	60	69–392
	CAT	167	33	107–257	65	11	40–82	140	31	71–209
	THERE	145	19	121–182	62	12	39–82	314	57	202–467

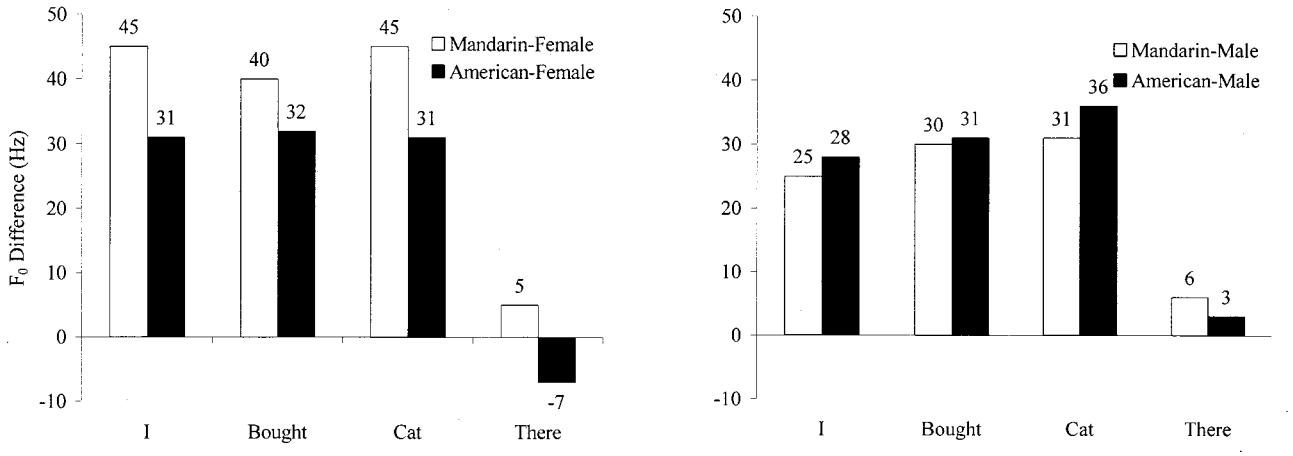

FIG. 1. Comparison of across-sentence F_0 between American and Mandarin speakers. The values represent the F_0 difference between stressed and unstressed words across sentences. Negative values indicate higher F_0 values in the nonstressed words compared to stressed words.

female speakers, and Table II for American and Mandarin males. Two-way repeated analysis of variance (ANOVA) tests were performed separately for each acoustic feature according to gender group. Native language (American English vs Mandarin) was the between groups factor and stressed words (*I, BOUGHT, CAT, THERE*) was the within groups factor.[5]

1. F_0

Results of the two-way ANOVA for the F_0 of female speakers revealed a significant main effect for language group [$F(1,479)=145.1$, $p<0.01$; $\eta^2=0.235$], with Mandarin females producing stressed words with a higher F_0 compared to American females. There was also a significant main effect for stressed word [$F(3,479)=19.8$, $p<0.01$; $\eta^2=0.112$]. Post hoc *t*-tests found *THERE* to be produced with a significantly lower F_0 compared to the remaining three stressed words ($p<0.01$). There was no significant interaction between female language groups and the F_0 of stressed words [$F(3,479)=0.14$, $p>0.01$; $\eta^2=0.001$]. Results for the male speakers revealed a significant main effect for language group [$F(1,479)=51.8$, $p<0.01$; $\eta^2=0.099$], with Mandarin males showing a higher F_0 for stressed words compared to American males. There was a significant main effect for stressed word [$F(3,479)=18.1$, $p<0.01$; $\eta^2=0.103$]. Post hoc *t*-tests found the F_0 of *THERE* to be significantly lower compared to the remaining three stressed words ($p<0.01$). No significant interaction was found between language group and stressed word [$F(3,479)=0.35$, $p>0.01$; $\eta^2=0.002$].

2. Duration

A two-way ANOVA for vowel durations produced by female speakers was nonsignificant for language group [$F(1,479)=0.06$, $p>0.01$; $\eta^2=0.000$]. A significant main effect occurred for stressed words [$F(3,479)=212.6$, $p<0.01$; $\eta^2=0.575$]. Post hoc *t*-tests found each of the four stressed words to be significantly different in duration compared to the other ($p<0.01$). There was no significant interaction between female language groups and the duration of stressed words [$F(3,479)=0.59$, $p>0.01$; $\eta^2=0.004$]. The male speakers demonstrated a significant main effect for language group [$F(1,479)=29.7$, $p<0.01$; $\eta^2=0.059$], with Mandarin males producing stressed words with shorter vowel durations compared to American males. There was also a significant effect for stressed words [$F(3,479)=196.2$, $p<0.01$; $\eta^2=0.555$] with each of the stressed words significantly different in duration compared to the other ($p<0.01$). There was no significant interaction between language group and stressed word [$F(3,479)=0.77$, $p>0.01$; $\eta^2=0.005$].

3. Intensity

Results of a two-way ANOVA for vowel intensity revealed a nonsignificant main effect for the female language groups [$F(1,479)=5.0$, $p>0.01$; $\eta^2=0.010$]. There was a significant main effect for the intensity of stressed words [$F(3,479)=11.5$, $p<0.01$; $\eta^2=0.068$], with *BOUGHT* being produced with greater intensity than *THERE* ($p<0.01$). There was no significant interaction between female language group and intensity of stressed words [$F(3,479)=0.04$, $p>0.01$; $\eta^2=0.000$]. The ANOVA results for male speakers were nonsignificant for language group [$F(1,479)=4.2$, $p>0.01$; $\eta^2=0.009$]. There was a significant main effect for stressed word [$F(3,479)=8.5$, $p<0.01$; $\eta^2=0.051$], with *BOUGHT* being produced with greater intensity than the three remaining stressed words ($p<0.01$). No significant interaction was found between language group and stressed word [$F(3,479)=0.25$, $p>0.01$; $\eta^2=0.002$].

B. Across-sentence stress

1. F_0

Results of the analysis for F_0 are shown in Fig. 1. Both American and Mandarin speakers produced stressed words with a higher F_0 compared to unstressed words. Results of a two-way ANOVA for the F_0 of female speakers revealed a significant main effect for language group [$F(1,479)=22.7$, $p<0.01$; $\eta^2=0.046$]. Mandarin females differentiated stressed from unstressed words with a higher F_0 compared to American females. There was a significant effect for

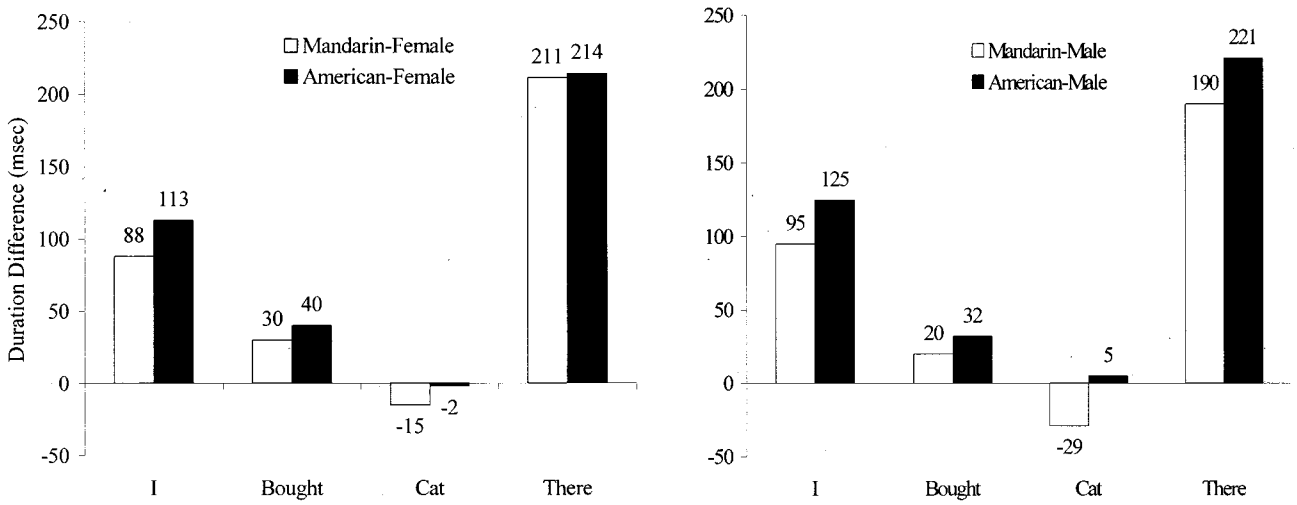

FIG. 2. Comparison of across-sentence duration between American and Mandarin speakers. The values represent the duration difference between stressed and unstressed words across sentences. Negative values indicate longer duration values in the nonstressed words compared to stressed words.

the F_0 of stressed words [$F(3,479)=15.4$, $p<0.01$; $\eta^2=0.089$]. Post hoc testing found the word *CAT* to have a higher differentiated F_0 compared to the remaining three words ($p<0.01$). There was no significant interaction between female language groups and stressed words [$F(3,479)=1.14$, $p>0.01$; $\eta^2=0.007$]. The ANOVA results for male speakers showed a nonsignificant effect for language group [$F(1,479)=0.89$, $p>0.01$; $\eta^2=0.002$], and a significant effect for the F_0 of stressed words [$F(3,479)=17.3$, $p<0.01$; $\eta^2=0.099$]. The word *CAT* had a higher differentiated F_0 compared to the remaining three words ($p<0.01$). There was no significant interaction between male language groups and stressed words [$F(3,479)=2.08$, $p>0.01$; $\eta^2=0.013$].

2. Duration

Results of the duration analysis are displayed in Fig. 2. Both American and Mandarin speakers produced stressed words with a longer duration compared to unstressed words. A two-way ANOVA for vowel durations produced by the female speakers showed a significant effect for language group [$F(1,479)=7.2$, $p<0.01$; $\eta^2=0.015$], with Mandarin females producing differentiated stress with shorter vowel durations. There was a significant effect for the duration of stressed words [$F(3,479)=123.4$, $p<0.01$; $\eta^2=0.44$]. The word *I* was produced with longer differentiated duration compared to the remaining three words ($p<0.01$). There was no significant interaction between female language groups and production of stressed words [$F(3,479)=2.53$, $p>0.01$; $\eta^2=0.016$]. A two-way ANOVA for male speakers was significant for language group [$F(1,479)=28.2$, $p<0.01$; $\eta^2=0.056$], with Mandarin males producing differentiated stress with shorter vowel durations. There was a significant effect for stressed word [$F(3,479)=126.3$, $p<0.01$; $\eta^2=0.445$], with *I* being produced with longer differentiated duration compared to the remaining three words ($p<0.01$). There was a significant interaction between male language groups and stressed word [$F(3,479)=9.35$, $p<0.01$; $\eta^2=0.056$]. Post hoc testing found Mandarin males to produce *I* and *THERE* with significantly shorter differentiated duration compared to American males ($p<0.01$).

3. Intensity

Results of the intensity analysis are shown in Fig. 3. American and Mandarin speakers produced stressed words with greater intensity compared to unstressed words. Results of the two-way ANOVA for female speakers revealed a nonsignificant effect for language group [$F(1,479)=1.9$, $p>0.01$; $\eta^2=0.004$]. There was a significant effect for stressed words [$F(3,479)=22.2$, $p<0.01$; $\eta^2=0.124$], with the differentiated intensity of *THERE* being higher than the remaining three words ($p<0.01$). There was no significant interaction between female language groups and stressed word [$F(3,479)=0.67$, $p>0.01$; $\eta^2=0.004$]. Results of the two-way ANOVA for male speakers was nonsignificant for language group [$F(1,479)=8.6$, $p>0.01$; $\eta^2=0.018$]. There was a significant effect for the intensity of stressed words [$F(3,479)=23.5$, $p<0.01$; $\eta^2=0.130$], with the differentiated intensity of *THERE* being higher than the remaining three words ($p<0.01$). There was no significant interaction between language group and stressed word [$F(3,479)=0.60$, $p>0.01$; $\eta^2=0.004$].

IV. DISCUSSION

A. Sentence stress according to F_0

It was predicted that Mandarin speakers would be able to alter F_0 in L2 English to differentiate stress from unstressed words. The present results indicate that Mandarin speakers were able to use F_0 to differentiate stressed from unstressed words in the production of American English sentences. Similar to American speakers, stressed words were consistently produced with the highest F_0. Although results obtained from the across-sentence analyses indicate Mandarin speakers' appropriate use of F_0 to signal stress, it is interesting to consider the results of the significance testing regarding the characteristics of F_0 for stressed and unstressed words. The ANOVA results obtained from the aver-

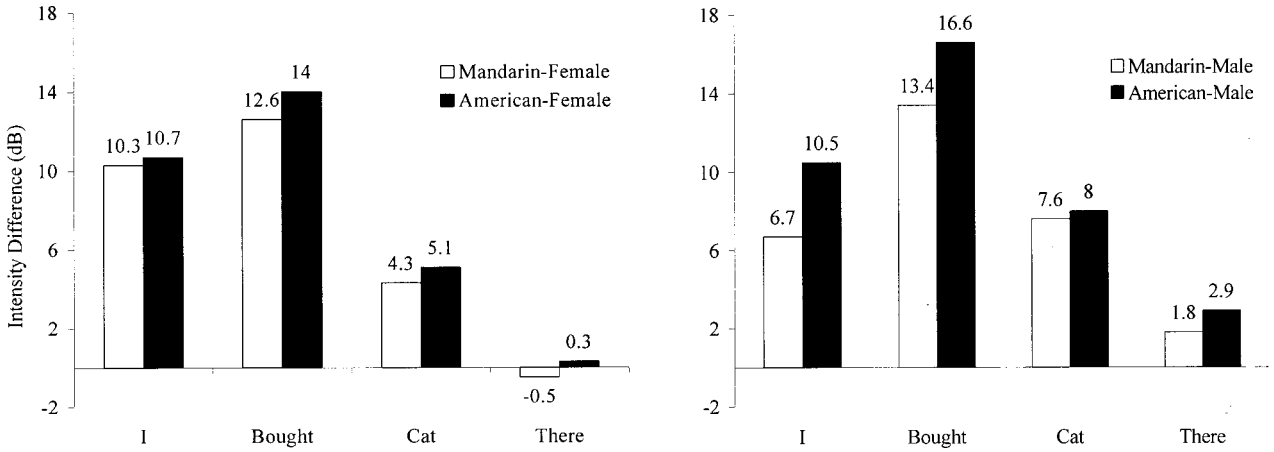

FIG. 3. Comparison of across-sentence intensity between American and Mandarin speakers. The values represent the intensity difference between stressed and unstressed words across sentences. Negative values indicate higher intensity values in the nonstressed words compared to stressed words.

age stress analysis found male and female Mandarin speakers to produce stressed words with a significantly higher F_0 than their American counterparts. A significantly higher F_0 was also found for the female Mandarin speakers in the across-sentence analysis.

Two possible explanations may account for the use of a higher F_0 for sentence stress among male and female Mandarin speakers. The first explanation is that the higher F_0 among Mandarin speakers may simply reflect basic anthropometric differences between Asian and Caucasian speakers. Yang (1996) examined the acoustic characteristics of Korean speakers compared to Caucasian speakers and found that the vocal F_0 and formant frequencies were higher in the Korean speakers. He attributed the higher acoustic values in Korean speakers to an overall smaller laryngeal framework compared to Caucasian speakers. However, Ng (1996) found no differences in vocal F_0 between Cantonese-speaking males compared to American-English speaking males. In the absence of anatomic measures of the larynx, it is unknown whether the observed differences in F_0 resulted from discrepancies in laryngeal anatomy.

Another reason for the higher F_0 among Mandarin speakers may relate to the influence of Mandarin tones on the production of American English. Eady (1982) found that the continuous speech of Mandarin contains significantly greater F_0 fluctuations at the syllable-level compared to the continuous speech of English. The greater F_0 fluctuations in Mandarin sentences were attributed to the fact that every syllable spoken in Mandarin has its own tone contour. This finding was examined in greater detail by Xu (1999) who found that the lexical tone determines the F_0 contour for a syllable, as well as the shape of F_0 contours for surrounding syllables. In addition, the location of stress within a sentence serves to vary the overall (global) F_0 of the entire sentence. Shen (1990) found that when native speakers of Mandarin produce French as L2 (a nontonal language), greater pitch fluctuations were found in the L2. Shen believed that the interplay between sentence intonation and lexical tones served to increase overall pitch level. The suggestion of an interplay between intonation and lexical tones was also espoused by Chao (1968) who suggested that the interplay was similar to small ripples riding on large waves which results in an algebraic sum of two kinds of waves. Considering the results from these various studies, it is possible that the higher average F_0 shown by Mandarin speakers in the present study was indicative of interference from L1, whereby Mandarin speakers produced words with greater pitch fluctuations compared to American speakers.

To evaluate whether the Mandarin groups produced a higher F_0 regardless of sentence stress, a series of post hoc analyses were performed. The F_0 values comprising the entire sample of unstressed vowels were summed and compared between Mandarin and American English groups. Results of the analysis found Mandarin females to produce unstressed words with a significantly higher F_0 compared to American females $[F(1,1439)=402.71,\ p<0.01;\ \eta^2=0.22]$. Results for the Mandarin males also indicated a significantly higher F_0 for unstressed words compared to American males $[F(1,1439)=253.9,\ p<0.01;\ \eta^2=0.15]$. Thus while Mandarin speakers were able to use F_0 to differentiate stressed from unstressed words, there was a tendency for these speakers to use a high F_0 across stressed and unstressed words in their production of L2 English. This finding lends further support to Shen's (1990) contention of an increased overall pitch level during L2 production because of an interplay between sentence intonation and lexical tones in L1 Mandarin. Future acoustic comparisons of Mandarin and American English speakers should consider the possible influence of anthropometric differences, as well as the interplay between sentence intonation and lexical tones to ascertain why Mandarin speakers of English use a significantly higher F_0.

B. Sentence stress according to duration

It was predicted that Mandarin speakers would show no significant differences from American speakers in their use of duration to signal American English sentence stress. Results of the analysis showed that Mandarin speakers produced stressed words with a longer vowel duration compared to unstressed words. The pattern of stress duration shown across the words was similar to that found for American

speakers (Fig. 2). However, female and male Mandarin speakers produced stressed words with significantly shorter vowel durations compared to the same productions by American speakers. This behavior was particularly apparent in the across-sentence analysis which examined differentiated vowel duration.

To evaluate whether Mandarin speakers differed from American speakers in vowel duration regardless of sentence stress, a series of post hoc analyses were performed. The vowel duration values comprising the entire sample of unstressed words were summed and compared between Mandarin and American English groups. Results of the analysis found Mandarin females to produce unstressed words with a significantly longer duration compared to American females [$F(1,1439)=27.3$, $p<0.01$; $\eta^2=0.02$], while there was no significant difference between Mandarin and American males in the vowel duration of unstressed words [$F(1,1439)=2.29$, $p>0.13$; $\eta^2=0.002$]. The behavior noted for production of unstressed words would suggest that Mandarin speakers' application of duration in L2 English is somewhat variable. These results are in agreement with a recent report by Guion *et al.* (2000) who found that temporal variability is a common feature of L2 English. Because Mandarin is considered a syllable-timed language, whereby syllable duration is relatively constant across a sentence, it is not surprising that Mandarin speakers would have difficulty regulating vowel duration in multisyllabic American English sentences (Chun, 1982).

C. Sentence stress according to intensity

The third prediction was that Mandarin speakers would show no significant differences from American English speakers in their use of intensity to signal sentence stress. Results from the across-sentence and average stress analysis found Mandarin and American speakers to produce stressed words with comparable intensity. The finding of no differences in intensity between the two language groups agrees with previous studies which indicated that intensity serves as a cue for stress in both Mandarin and American English languages (Fry, 1955; Lieberman, 1967; Lin and Yan, 1980; Shen, 1993).

There are reports of intensity being the least salient and less consistent acoustic parameter to differentiate stress from unstressed words (Clark and Yallop, 1995). Yet, this feature of sentence stress was perhaps the most comparable across Mandarin and American speakers. Additional post hoc analyses examined whether the intensity of unstressed words was also comparable across Mandarin and American speakers. The analysis revealed that Mandarin females produced unstressed words with a significantly higher intensity compared to American females [$F(1,1439)=23.2$, $p<0.01$; $\eta^2=0.02$]. Mandarin males also produced unstressed words with a significantly higher intensity than the American males [$F(1,1439)=40.2$, $p<0.01$; $\eta^2=0.16$]. Therefore, while it is likely that Mandarin and American groups used intensity to signify sentence stress, Mandarin subjects appeared to use intensity to a greater extent across both stressed and unstressed words.

V. SUMMARY AND CONCLUSIONS

This study confirmed the predictions that Mandarin speakers of English would be capable of signifying sentence stress using F_0, duration, and intensity. Although the present group of Mandarin speakers were able to signal sentence stress, the acoustic characteristics of stress were not identical to American speakers. Mandarin speakers produced stressed words with a significantly higher F_0 and shorter vowel duration compared to American speakers. Acoustic examination of unstressed words revealed additional differences between language groups. Mandarin speakers produced unstressed words with significantly higher F_0 and greater intensity than American speakers, and the vowel duration of unstressed words was either similar or longer than that of American speakers. The most reasonable explanation for these differences is an interference of L1 Mandarin in the production of L2 American English. The outcome of this study suggests that, although there are acoustic differences between Mandarin and American speakers in the production of sentence stress, there is no critical divergence between these speakers in the way they implement sentence stress in English. The work here adds to the limited body of data examining the influence of Mandarin on the accurate production of L2 American English.

[1]Putonghua (i.e., the general/standard Chinese language) has been promoted as the ''official language'' by the government throughout China since the 1950s (Chao, 1968; Cheng, 1987). It is based on the phonological and grammatical system known as Beifanghua (i.e., the northern dialect). In English, Beifanghua is also often referred to as Mandarin (Hua and Dodd, 2000).

[2]The interplay of stress and unstressed syllables/words within a sentence is thought by some researchers to reflect an isochronal (i.e., equal timing) pattern of speech rhythm (Chun, 1982; Crystal, 1969; Lehiste, 1980); however, this remains a matter of debate (cf. Cruttenden, 1986). The intent of the present study was to simply note the general timing pattern of English as contrasted to Mandarin.

[3]Besides the four ''basic'' tones in Mandarin Chinese, there is also a fifth tone, which is usually referred to as a neutral tone. According to Chao (1968), almost any morpheme normally having one of the regular four tones can be produced in a neutral tone under certain conditions. There are also a very small number of morphemes, such as suffixes, which tend to be produced in a neutral tone and do not possesses one of the four basic tones.

[4]Equal numbers of male and female subjects were selected for two reasons. First, there are limited data concerning the production of American English sentence stress by native speakers of Mandarin and several acoustic studies examining Mandarin speakers are based on subject samples of eight or less (Eady, 1982; Shen, 1990; Xu, 1997, 1999). The small sample of male and female speakers has led to the collapsing of data across genders, thereby requiring that absolute values for features such as F_0 be converted to logarithms (e.g., Xu, 1999). The relatively large number of male and female subjects sampled in the present study allowed for grouping subjects according to gender and to provide representative estimates of central tendency. Second, because a portion of the acoustic analyses involved examining absolute measures of F_0, it was deemed inappropriate to collapse the results across males and females because of the known influence of laryngeal and vocal tract anatomy on the subsequent speech product. Although it was not the intent of the present study to explore whether gender differences were apparent, we believed that reporting data according to gender served to provide additional descriptive data regarding English spoken as L2.

[5]The phonetic composition of the four words was not identical which might naturally contribute to observed acoustic differences between words. However, the ANOVA test applied to the present database was thought to provide the most comprehensive analysis within and between native Mandarin and English groups. Any significant differences which may have occurred

between groups in the production of specific word types would be identified by using word as a within-groups factor.

Berinstein, A. (**1979**). ''A cross-linguistic study on the perception and production of stress,'' UCLA Working Papers of Phonetics, **47** (UCLA Phonetics Laboratory, Los Angeles).

Bolinger, D. (**1958**). ''A theory of pitch accent in English,'' Word **14**, 109–149.

Brown, W., and McGlone, R. (**1974**). ''Aerodynamic and acoustic study of stress in sentence productions,'' J. Acoust. Soc. Am. **56**, 971–974.

Cao, J. (**1986**). ''Pu Tong Hua Qing Sheng Yin Jie Te Xin Fen Xi'' [Analysis of characteristics of unstressed syllables in Mandarin], Ying Yong Sheng Xue **5**, 23–37.

Chao, Y. (**1932**). ''A preliminary study of English intonation (with American variants and its Chinese equivalents),'' The Tsai Yuan Pei Anniversary Volume (Suppl. of the Bulletin of the Institute of History and Philosophy), Vol. 1, pp. 105–156.

Chao, Y. (**1948**). *Mandarin Primer* (Harvard University Press, Cambridge, MA).

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

Cheng, C. (**1968**). ''English stress and Chinese tones in Chinese sentences'' Physica C **18**, 77–88.

Cheng, L. (**1987**). *Assessing Asian Language Performance: Guidelines for Evaluating Limited-English-Proficient Students* (Aspen, Rockville, MD).

Chun, D. (**1982**). ''A contrastive study of the suprasegmental pitch in modern German, American English and Mandarin Chinese,'' unpublished doctoral dissertation, University of California, Berkeley.

Clark, J., and Yallop, C. (**1995**). *An Introduction to Phonetics and Phonology*, 2nd ed. (Blackwell, Cambridge, MA).

Cooper, W., Eady, S., and Mueller, P. (**1985**). ''Acoustical aspects of contrastive stress in question-answer contexts,'' J. Acoust. Soc. Am. **77**, 2142–2156.

Cooper, W., and Sorensen, J. (**1981**). *Fundamental Frequency in Sentence Production* (Springer-Verlag, New York).

Cruttenden, A. (**1986**). *Intonation* (Cambridge University Press, Cambridge).

Crystal, D. (**1969**). *Prosodic Systems and Intonation in English* (Cambridge University Press, Cambridge).

Dreher, F., and Lee, P. (**1966**). ''Instrumental investigation of single and paired Mandarin tonemes'' (Douglas Advanced Research Laboratory).

Eady, S., (**1982**). ''Differences in the *F*0 patterns of speech: Tone language versus stress language,'' Lang. Speech **25**, 29–42.

Eady, S., and Cooper, W. (**1986**). ''Speech intonation and focus location in matched statements and questions,'' J. Acoust. Soc. Am. **80**, 402–415.

Eady, S., Cooper, W., Klouda, P., Mueller, D., and Lotts, P. (**1986**). ''Acoustic characteristics of sentential focus: Narrow vs broad and single vs dual focus environments,'' Lang. Speech **29**, 233–251.

Flege, J. (**1987**). ''The instrumental study of L2 speech production: Some methodological considerations,'' Lang. Learn. **37**, 285–296.

Flege, J. (**1989**). ''Chinese subjects' perception of the word-final English /t/–/d/ contrast: Performance before and after training,'' J. Acoust. Soc. Am. **86**, 1684–1697.

Flege, J. (**1995**). ''Second language speech learning: Theory, findings, and problems,'' in *Speech Perception and Linguistic Experience*, edited by W. Strange (York Press, Baltimore).

Flege, J., and Davidian, R. (**1985**). ''Transfer and developmental processes in adult foreign language speech production,'' J. Appl. Psycholing. Res. **5**, 323–347.

Fromkin, V., and Ohala, J. (**1968**). ''Laryngeal control and a model of speech production,'' UCLA Working Papers Phonet. **10**, 98–110.

Fry, D. (**1955**). ''Duration and intensity as physical correlates of linguistic stress,'' J. Acoust. Soc. Am. **27**, 765–768.

Fry, D. (**1958**). ''Experiments in the perception of stress,'' Lang. Speech. **1**, 126–152.

Gandour, J. (**1978**). ''The perception of tone,'' in *Tone: A Linguistic Survey*, edited by V. Fromkin (Academic, New York).

Guion, S., Flege, J., Liu, S., and Yeni-Komshian, G. (**2000**). ''Age of learning effects on the duration of sentences produced in a second language,'' Appl. Psycholing. **21**, 205–228.

Guo, J. (**1992**). *Zhong He Yu Yin Xue* [Comprehensive Phonetics] (Fujian Ren Min Chu Ban She, Fujian, China).

Guzma, S. (**1973**). ''*Les intérferences de l'intonation interrogative dans le français parlé par les Chiliens*,'' Thèse de 3ème cycle, Paris III.

Ho, A. (**1976**). ''Mandarin tones in relation to sentence intonation and grammatical structure,'' J. Chinese Ling. **4**, 1–13.

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

Hua, Z., and Dodd, B. (**2000**). ''The phonological acquisition of Putonghua (Modern Standard Chinese),'' J. Child Lang. **27**, 3–42.

Hutchinson, S. (**1973**). ''An objective index of the English-Spanish pronunciation dimension,'' unpublished masters thesis, University of Texas, Austin, TX.

Ingram, J., and Pittam, J. (**1987**). ''Auditory and acoustic correlates of perceived accent change: Vietnamese school children acquiring Australian English,'' J. Phonetics **15**, 127–143.

Ingram, J., and Park, S. (**1997**). ''Cross-language vowel perception and production by Japanese and Korean learners of English,'' J. Phonetics **25**, 343–370.

Johnson, J., and Newport, E. (**1991**). ''Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language,'' Cognition **39**, 215–258.

Kim, D. (**1972**). ''A contrastive study of English and Korean phonology (on the basis of aural perception),'' Lang. Teach. **5**, 1–36.

Kratochvil, P. (**1962**). ''The role of stress in the syntactical analysis of modern Peking dialect,'' Arch. Orient. **30**, 147–49.

Laver, J. (**1994**). *Principles of Phonetics* (Cambridge University Press, Cambridge).

Leben, W. (**1978**). ''The representation of tone,'' in *Tone: A Linguistic Survey*, edited by V. Fromkin (Academic, New York).

Lehiste, I. (**1970**). *Suprasegmentals* (Cambridge University Press, Cambridge).

Lehiste, I. (**1980**). ''Phonetic Manifestation of Syntactic Structure in English,'' Annual Bulletin Research Institute of Logopedics and Phoniatrics, No. 14, pp. 1–27, Tokyo.

Lenneberg, E. (**1967**). *Biological Foundations of Language* (Wiley, New York).

Lieberman, P. (**1960**). ''Some acoustic correlates of word stress in American English,'' J. Acoust. Soc. Am. **32**, 451–454.

Lieberman, P. (**1967**). *Intonation, Perception and Language* (Research Monograph, 38) (M.I.T. Press, Cambridge, MA).

Lin, M., and Yan, J. (**1980**). ''Beijing Hua Qin Sheng De Shen Xue Xin Zhi [Acoustic characteristics of unstressed syllables in Beijing Dialect],'' Fang Yan **3**, 16–24.

Magnuson, J., and Akahane-Yamada, R. (**1996**). ''Acoustic correlates to the effects of talker variability on the perception of English /r/ and /l/ by Japanese listeners,'' in *Proceedings of the Fourth International Congress on Spoken Language Processing* (A. duPont Institute, France).

Ng, L. (**1996**). ''A perceptual and acoustic study of alaryngeal speech in adult Cantonese-speaking males,'' unpublished doctoral dissertation, University of Connecticut, Storrs.

Ohala, J., and Hirano, M. (**1967**). ''Studies of pitch change in speech,'' UCLA Working Papers Phonet. **7**, 80–84.

Pike, K. (**1945**). *The Intonation of American English* (University of Michigan Press, Ann Arbor, MI).

Pike, K. (**1948**). *Tone Languages* (University of Michigan Press, Ann Arbor, MI).

Pittam, J., and Ingram, J. (**1992**). ''Accuracy of perception and production of compound and phrasal stress by Vietnamese Australians,'' Appl. Psycholing. **13**, 1–12.

Potisuk, S., Gandour, J., and Harper, M. (**1996**). ''Acoustic Correlates of Stress in Thai,'' Phonetica **53**, 200–220.

Robson, B. (**1982**). ''Teaching English pronunciation to speakers of Korean,'' Focus **11**, 1–8.

Scuffil, M. (**1982**). *Experiments in Comparative Intonation—A Case Study of English and German* (Max Niemeyer, Tübingen).

Shen, X. (**1990**). ''Ability of learning the prosody of an intonational language by speakers of a tonal language: Chinese speakers learning French prosody,'' Int. Rev. Appl. Ling. Lang. Teach. **28**, 119–134.

Shen, X. (**1993**). ''Relative duration as a perceptual cue to stress in Mandarin,'' Lang. Speech **36**, 415–433.

Shih, C. (**1986**). *The Phonetics of the Chinese Tonal System: Technical Memorandum* (AT&T, Bell Laboratories).

Tarone, E. (**1980**). ''Some influences on the syllable structure of interlanguage phonology,'' Int. Rev. Appl. Ling. Lang. Teach. **18**, 139–152.

Tseng, C. (**1981**). ''An acoustic phonetic study on tones in Mandarin Chinese,'' unpublished doctoral dissertation, Brown University, Providence, RI.

van Santen, J., and Shih, C. (**2000**). ''Suprasegmental and segmental timing models in Mandarin Chinese and American English,'' J. Acoust. Soc. Am. **107**, 1012–1026.

Wijk, A. (**1966**). *Rules of Pronunciation for the English Language: An Account of the Relationship Between English Spelling and Pronunciation* (Oxford University Press, London).

Xu, Y. (**1997**). ''Contextual tonal variations in Mandarin,'' J. Phonetics **25**, 61–83.

Xu, Y. (**1999**). ''Effects of tone and focus on the formation and alignment of F_0 contours,'' J. Phonetics **27**, 55–105.

Yang, B. (**1996**). ''A comparative study of American English and Korean vowels produced by male and female speakers,'' J. Phonetics **24**, 245–261.