Perception of acoustic source characteristics: Walking sounds[a)]

Xiaofeng Li, Robert J. Logan, and Richard E. Pastore
Department of Psychology, Center for Cognitive and Psycholinguistic Sciences, State University of New York at Binghamton, Binghamton, New York 13902-6000

(Received 14 January 1991; accepted for publication 15 August 1991)

The current study investigated the ability of subjects to perceive source characteristics (the gender of a human walker) of a naturally occurring auditory event (human walking). A number of acoustic properties were measured and subjected to statistical analyses in order to identify those properties that differentiate male and female walking footsteps. A principal component analysis on the statistical properties of the spectral energy distributions then identified two classes of information that were important in determining subject perception of the gender of a walker: (1) the spectral peak which integrates the information about the spectral central tendency of frequency and shape of the spectral peak; and (2) the contribution of high-frequency spectral components. A follow-up experiment then manipulated these spectral properties to verify their contributions in the perceptual classification of walker gender. Additionally, the effect of shoe on the gender judgement in walking sequences was assessed by having both male and female walkers wear male's shoes.

PACS numbers: 43.66.Lj [WAY]

INTRODUCTION

People sometimes exhibit excellent abilities to use sounds to identify the objects or events that produced those sounds. It is common for listeners to judge the sex and approximate age of an unfamiliar person speaking on the telephone, and to even identify a voice of a familiar person. This simple anecdotal example illustrates that humans use properties of sounds to perceive ecologically important characteristics of environmental acoustic sources. The identification of source characteristics of sounds probably represents an important and often used skill that may be critical to certain tasks (e.g., operators of passive sonar identifying underwater objects). The need to understand the actual processes and potential in the identification of sound source characteristics is one major goal for auditory research (CHABA, 1989). Unfortunately, surprisingly little research has directly evaluated the nature and extent of this ability; such evaluations are necessary precursors to any investigation of how this important skill might be improved (CHABA, 1989).

The current study investigated the perception of distal source characteristics of human walking on a hard surface as mediated by proximal acoustic properties. The decision to utilize walking stimuli was motivated by reasons of ecological validity and subjective experience. Few previous studies on source perception utilized stimuli that were readily identified as originating from the appropriate source event. For example, the separate clapping stimuli employed by Repp (1987) and Tousman and Pastore (1989), and the synthetic propeller cavitation sounds employed by Howard (1983) were not readily and consistently identified in terms of the class of appropriate source events. However, we have observed an excellent and consistent ability to identify gender, and even the identity of specific colleagues, based solely upon the sound of their walking in the hallway outside our offices.[1] Although the study of walking itself may not be the primary interest for research, it does represent a class of ecologically valid stimuli which impose a nonarbitrary structure on the acoustical waveform. The methodology used in, and conclusions derived from, this investigation may provide insights for understanding the perception of other complex auditory events.

I. PREVIOUS RESEARCH

Few studies have directly evaluated the ability of humans to recognize source characteristics of environmental sounds. Warren and Verbrugge (1984) investigated the dynamic physical characteristics of stimuli that govern the perception of bouncing versus breaking events. This study evaluated the relationship between acoustic structure and perception in terms of temporal stimulus properties. Listeners were presented with the edited and unedited sounds of objects bouncing and breaking, with the stimuli varying primarily in temporal attributes. Warren and Verbrugge (1984) found that there was a high correlation between the temporal properties of the stimuli and perceptual judgements of bouncing versus breaking. They concluded that dynamic stimulus cues carry the information that determines bouncing versus breaking events.

Other studies focused on spectral parameters in investigating the relationship between the acoustic waveform and perception. The methodologies employed in these studies fall into two categories based upon whether the techniques were used to identify hypothetical dimensions or gross statistical properties of the waveform. We will use the study by Repp (1987) to illustrate the first methodology (the identification of hypothetical dimensions), although Repp used

[a)] Requests for reprints should be sent to X. Li or R. E. Pastore, Department of Psychology, State University of New York at Binghamton, NY 13902-6000, or R. J. Logan, IBM Corporation, P.O. Box 950, D59A/701, Poughkeepsie, NY 12602.

this method for data reduction rather than for hypothetical dimension extraction. Repp found a moderate ability of subjects to identify positions of hands as articulators in producing claps, then analyzed the correlation between physical parameters and perceptual classification of hand positions. Conducting a principal component factor analysis on various clapping spectra, he found four factors that could account for most spectral shapes which characterized four prototypical hand configurations. Repp interpreted the derived dimensions as those which define the prototypical classes of claps. Unlike the current investigation, Repp could not identify the physical parameters that differentiate clapper gender. A different version of the hypothetical dimension methodology is exemplified by the use of multidimensional scaling techniques to construct quantitative dimensions to describe various auditory events (e.g., Grey, 1977; Grey and Gordon, 1978; Pols *et al.*, 1969; Wedin and Goude, 1972).

The second category of methodological techniques (the identification of gross statistical properties of spectra) uses various statistical measures of spectral distributions as regressors for perceptual judgements. This general type of methodological approach has been found to be effective in the investigation of timbre perception, such as spectral "sharpness" (von Bismarck, 1974), and spectral "brightness" (Grey and Gordon, 1978), as well as the classification of vowels (Potter and Steinberg, 1950) and fricatives (Jassem, 1979). A recent nonspeech study by Freed (1990) examined the correlation of perceptual ratings of mallet hardness for percussive sound events with two statistical measures of static and dynamic spectra. These two statistical measures were overall spectral levels and spectral centroid. The latter is equivalent to the first statistical moment of the spectral distribution. In the speech domain, Forrest *et al.* (1988) evaluated the relationship between three moments (mean, skewness, and kurtosis) and perception. These three statistical moments were found to accurately classify word-initial voiceless obstruents regardless of speaker gender.

These various studies on the perception of complex acoustical events suggest that the two methodological approaches can be used effectively to explore the correlation between physical parameters of a sound and subjects' perceptual judgements.

II. CURRENT RESEARCH

The current study used human walking as the stimulus events. The subjects' task was to identify the gender of a walker. Before presenting the current study, we will analyze the biomechanical aspect of human walking with the intent to identify the possible skeletal differences between male and female walkers, which may be significant for the gender judgement. We then discuss a general methodology that guides the current study.

A. Biomechanical analysis of human walking

A series of visual studies have demonstrated that humans are able to identify their biological motion (e.g., human walking) through viewing a constellation of lights mounted on the skeletal joints even with all other cues of actual body movements removed (Johansson, 1973). Kozlowski and Cutting (1977; also Barclay *et al.*, 1978; Cutting, 1978) showed that subjects are able to identify gender of a walker by viewing such an array of lights. These authors argued that the percept of an object in the environment is determined by the distal structure that remains invariant during certain transformations. Aspects of this ecological research evoked our interest to investigate walking in this auditory domain.

Cutting *et al.* (1978) provide a thorough analysis of the relationship between human skeletal structure and walking. Walking, like other locomotion behaviors, involves the movement of the entire body from the point of contact the supporting surface. Accomplishing this movement requires that the feet apply force to a surface which must resist to the force, which, in turn, results in sound. The surface reaction moves the body in accordance with Newton's third law of motion. According to biomechanical analysis of human walking, a prototype of human walking consists of two phases: stance and swing. During the stance phase, a heel strike is initiated and then followed sequentially by foot flat, heel off, knee flexion, and toe off. The stance phase is followed by a swing phase during which toe clearance and leg swing occur to prepare for the next cycle of walking (Cooper *et al.*, 1982). These two phases are repeated periodically in the continuation of walking. It was found that the stance phase consumes most of energy (Cooper *et al.*, 1982; Zuniga and Leavitt, 1974). Based upon Newton's third law of motion, this finding suggests that the supporting surface would resist greater impact from a foot during the stance phase which is characterized by a heel strike than during the swing phase. Accordingly, we anticipate that a louder acoustic event (possibly the major acoustic component in a walking cycle) would occur during the stance phase. The current study thus focused attention to the analysis of the mapping between perception of a walker gender and the major acoustic component of walking—the heel strike.

Zuniga and Leavitt (1974) found several gender differences during these two phases. For example, female footsteps usually exhibit an absence of foot flat during the stance phase. This difference was attributed to the typical shorter foot length of female walkers than male walkers. Several of such biological differences can be anticipated to contribute to the physical properties of walking sounds. Relative to the average male, the average female is shorter, lighter, and has a lower center of gravity and a pelvic structure compatible with childbearing (also see the analysis of anthropomorphic data below).

Consistent with the findings in Cutting *et al.*'s study (1978), Zuniga and Leavitt (1974) suggest that gender differences in human walking may inherit from biological differences between male and female. Cutting *et al.* indicated that the *visual* perception of gender of a walker could be determined by different patterns of upper-body movements. This suggests that gender differences in viewing a human walking may be due to the skeletal differences between male and female walkers. Therefore, in the current study, we will predict that the perception of a walker gender by listening to the footsteps is a function of certain biological differences.

Various anthropomorphic data were collected for all the walkers to indicate the biological differences among walkers. We anticipated that the perception of walker gender will be correlated with some of these anthropomorphic measures. The degree of correlation, however, strongly depends upon the choice of these anthropomorphic measures.

B. Methodology of the current study

This investigation was guided by three types of pairwise relationships illustrated in Fig. 1. In order to identify properties of acoustic source events, those events must produce sounds with a nonarbitrary acoustic structure which, in turn, must be recovered by the listener. Moreover, in order to perceive source events, the listener must be able to map the relevant acoustic structure to auditory source categories defined in terms of the source attributes. This simple sequential model is illustrated with the arrow lines connecting the three boxes in Fig. 1.

Also depicted in Fig. 1 is an effective and efficient strategy for investigating the perception of source events; the evaluation of the pairwise relationships between the three stages. These pairwise relationships will be referred to as analysis 1, 2, and 3 as indicated in Fig. 1 by the solid brackets. First, categorization tasks should be used to map the relationship between perception and the categories of ecologically valid events. Second, the acoustic waveforms should be analyzed to identify acoustic properties which reliably differentiate source event categories. Finally, a correlation analysis between the identified acoustic properties and listener's perception should be used to evaluate which of these acoustic properties may have mediated perception. The validity of this final analysis then should be confirmed by altering these identified acoustic properties and observing the resulting effects on perception. This approach to investigating the perception of source characteristics is developed on the basis of early, highly successful efforts to understand the perception of speech categories or phonemes (Forrest *et al.*, 1988).

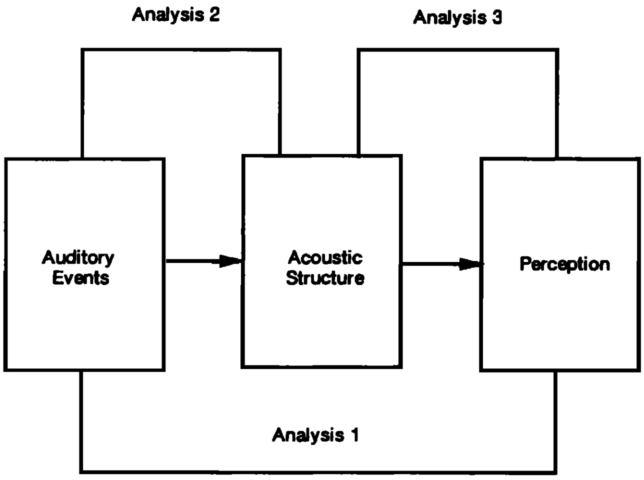


FIG. 1. The illustration of the general methodology which guides the current research. An auditory event (the left box) maps to a percept of the auditory event (the right box) mediated through the acoustic representation of the auditory event (the middle box). Three types of analyses can be conducted to understand how the auditory event is recovered based upon the acoustic representation of the auditory event designated by three brackets (for the description of the three analyses, see text).

C. Overview of the experiments

The current study consists of three behavioral experiments and an experiment of the physical analysis on the stimuli. Experiment 1 was designed to examine the relationship between auditory events and subject categorization of the stimuli (analysis 1). It addressed the issue of whether subjects were able to identify the gender of a walker. Experiment 2, which was not behavioral, used seven statistics of the spectral distributions to evaluate the mapping between auditory events and the acoustic structure (analysis 2), and the mapping between the acoustic structure and perception (analysis 3). The gross spectral descriptions were derived based upon these seven statistics using a principal component analysis. Experiment 3 was conducted to evaluate the effects on the perception of altering the three statistics which were identified in experiment 2 as important for subject judgements of walker gender. Finally, experiment 4 attempted to resolve the issue of whether the observed perceptual gender difference (from experiment 1) was due to different types of shoes worn by the male and female.

III. EXPERIMENT 1

This experiment used sequences of humans walking on a hard surface as stimuli to investigate whether subjects were able to judge the gender of the walker. This experiment was designed to provide analysis 1 in Fig. 1—specification of the relationship between perception and auditory events.

A. Method

1. Stimuli

The stimuli were recordings of human walking on a hardwood stage in a modern, 1200 seat performing arts theater (Anderson Center Concert Hall at SUNY—Binghamton). Nine graduate, six undergraduate students, and a faculty member (8 male and 8 female) voluntarily participated as walkers in the recording phase. Critical anthropomorphic characteristics of the 16 walkers are summarized in Table I and will be discussed below.

Fourteen walkers (M1, M2, M3, M5, M6, M7, and all female walkers) wore their own leather sole shoes with low, solid, synthetic heels typical of modern shoes. M4 and M8, who did not own such shoes, were recorded using the shoes of M3.

Starting with the left foot (and always beginning from a fixed location approximately 24 feet from the microphone and ending at similar locations), each walker took eight natural steps directly toward the microphone. This procedure was repeated at least three times for each walker, since stimuli with unusual sounds (e.g., off-stage sound or sound of loose board) were discarded. The microphone (B&K model 4135) was placed on a solid box approximately 3 ft above the floor. The stimuli were recorded on half-track tape (Tandberg TD 20A operating at 15 ips). The recorded stimuli were digitized (10-kHz sample rate, 5-kHz low-pass filter).

TABLE I. Anthropomorphic characteristics of the 16 walkers. The numbers in the bottom row are the correlation coefficients between the anthropomorphic measures and maleness judgments (Table II). The numbers in parentheses are the level of the statistical significance for the corresponding correlation coefficients.

Walker	M1	M2	M3	M4	M5	M6	M7	M8	Mean	Std
Weight (lbs)	168	170	170	130	155	155	170	150	158	13.8
Height (in.)	70	71	72	72	68	69	64	68	69	2.7
Ground-hip	41	41	41.5	42	39.5	41	36	42	40.5	2.0
Ankle-knee	18	18	18	19.5	17	17	16.6	18.5	17.6	1.4
Knee-hip	21	21	21	20.5	22.5	23	18.5	21.5	20.8	1.2
Foot length	11.3	11	11	11	9.8	8.5	7.5	11	10.5	1.3
Foot width	3.8	4	3.75	3.75	3.5	3.5	4	3.75	3.8	0.2
Walker	**F8**	**F7**	**F6**	**F5**	**F4**	**F3**	**F2**	**F1**	**Mean**	**Std**
Weight (lbs)	125	140	140	115	110	135	120	155	130	15.1
Height (in.)	66	69	67	64	62	65	64	66	65	2.1
Ground-hip	38	43.5	41	38.5	34.5	38	39	39	38.9	2.6
Ankle-knee	16	17	17.5	15.5	15	16	15	17.5	16.5	1.0
Knee-hip	21.5	23	21.5	19	19	19.5	21.5	19	20.5	1.6
Foot length	8	9.8	8	8.5	9	6.3	8	10	8.5	1.2
Foot width	3.25	3.5	3.25	3.3	3	3	3.5	3.5	3.3	0.2

Weight	Height	Ground-hip	Ankle-hip	Knee-hip	Foot length	Foot width
0.57 (0.02)	0.84 (00)	0.49 (0.05)	0.60 (0.01)	0.25 (0.35)	0.63 (0.01)	0.65 (0.01)

To avoid possible irregular or uncontrolled properties associated with the initiation and termination of walking, only the middle four steps were used in the experiments. Thus, this research used three 4-step walking sequences for each walker. The physical analyses (described in experiment 2) are based upon all 48 stimuli (16 walkers $\times$ 3 walking sequences). The perceptual experiments (1 and 3) used only one (arbitrarily selected) walking sequence for each walker. The stimuli used in experiment 4 are described below. All stimuli were diotic and presented to listeners over TDH-49P headphones through a 12-bit digital-to-analog converter at approximately 75 dB(A) on average [range: 70–78 dB(A)].

2. Subjects

Thirteen SUNY—Binghamton undergraduate students participated in experiment 1 as a course requirement. This experiment used the stimuli produced by the eight walkers (i.e., M1, M3, M5, M6, F7, F5, F3, and F1). In a later phase of this experiment, another group of 13 SUNY—Binghamton undergraduate students were paid to participate in the experiment using the stimuli produced by the other eight walkers (M2, M4, M7, M8, F8, F6, F4, and F2). Subjects were seated in a standard acoustic chamber. All subjects reported normal hearing.

3. Procedure

On each trial, a randomly selected four-footstep walking sequence was presented. Subjects were told that they would listen to a person walking on a hardwood surface and the stimulus consisted of four footsteps in a sequence. Subjects were instructed to identify whether this walking sequence was produced by a male or female by pressing a corresponding response button on the response box. Each group of the subjects finished 10 blocks of 48 trials (8 walkers $\times$ 6 repetitions) with a short break after every second block.

B. Results and discussion

Table II summarizes the average probability that each walker was judged as male. The continuum of probabilities from most probable male judgement to most probable female walker (therefore, least probable male walker) represents a continuum of maleness. The probabilities in Table II provided the basis for evaluating the pairwise relationship between complex auditory events and the perception of these events. We have used these ordered probabilities to assign labels to the walkers. The initial letter indicates the actual gender of the walker (M or F). The number then indicates the rank ordering of the degree of the perceived gender.

TABLE II. Composition of a step and the probability of being judged as a male walker for eight walkers (analysis 1 in Fig. 1). The numbers in the bottom rows are the means of maleness judgments for the male and female walkers. The numbers in parentheses are the standard deviation.

Walker	Heel strike	Toe strike	% "male"
M1	yes	yes	94.1
M2	yes	no	92.4
M3	yes	no	89.1
M4	yes	no	85.4
M5	yes	yes	51.4
M6	yes	yes	50.9
M7	yes	yes	47.5
M8	yes	yes	41.1
F8	yes	no	54.0
F7	yes	yes	43.8
F6	yes	no	30.2
F5	yes	no	25.4
F4	yes	yes	16.3
F3	yes	yes	15.9
F2	yes	no	12.1
F1	yes	no	6.1
Male			69.0 (23.1)
Female			25.5 (16.5)

Overall, the male walkers were judged as male most of the time (69%), while the probability that female walkers were judged as male (25.5%) was well below chance, indicating the female walkers were most of times judged as female (see Table II).

Within each actual gender category, the probability that the walking sequence was judged as male was not uniform, indicating that there is a continuum of maleness, rather than discrete perceptual categories. However, within this continuum, there was reasonable judgment of source gender with six male walkers judged as male above 50% of the time and the remaining two male walkers (M7 and M8) only slightly below chance (47% and 41%). All but one of the female walkers were judged as male less than 50% of the time with F8 only slightly above chance (54%). Therefore, the subjects were able to identify the gender of the walkers on the basis of acoustic information present in the waveforms. Walkers M1 and F1 were judged respectively as the most male and female (i.e., with the highest respective probabilities). Walkers M5, M6, F8, and F7 produced more ambiguous stimuli in terms of source gender. There was a slight tendency for walkers (M7, M8, and F8) to be judged as the gender opposite to their actual gender, indicating perceptual confusability of source gender on the basis of acoustic information. Finally, although M4 and M8 wore shoes of M3, judgement of maleness greatly differed among these three walkers, indicating that shoe characteristics did not have a dominant effect on judgement of walker gender. However, a more careful analysis of the effect of shoe will be performed in experiment 4.

1. Anthropomorphic analysis

Table I summarizes critical anthropomorphic information. Seven anthropomorphic measures were collected under the assumption that they may be correlated with the biological differences between male and female walkers.

Two-tail independent sample t tests were performed between the male and female walkers on the seven anthropomorphic data. Between the male and female walkers, statistically significant differences were found in both weight and height [for the weight, $t(14) = 3.88, p < 0.05$; for the height, $t(14) = 3.21$, $p < 0.05$], indicating that the male walkers were both heavier and taller than the female walkers. The male walkers also wore shoes that were significantly longer and wider than the female walkers [for the length, $t(14) = 3.58$, $p < 0.05$; for the width, $t(14) = 5.52$, $p < 0.05$]. For the length measures for lower extremities, generally, the male walkers did not differ significantly from the female [for ground to hip, $t(14) = 1.36$, $p > 0.05$; for knee to hip, $t(14) = 0.36, p > 0.05$]; the one possible exception was ankle to knee, which was marginally significant [$t(14) = 2.20, p < 0.0454$]. On the basis of these statistical tests, we concluded that, for the male and female walkers in our study, the significant anthropomorphic differences exist only in weight, height, and shoe size. These gross measures do not evaluate difference in pelvic structure or weight distribution whose quantification is beyond the scope of current research.

To examine whether these anthropomorphic characteristics are inherent sources that produce perceptual gender differences in this experiment, we calculated pairwise Pearson product correlation coefficients between the maleness (see Table II) and each of the seven anthropomorphic data, which are shown in the bottom of Table I. Table I shows that significant correlations were found for the weight, height, foot length, foot width, and the length between the ankle and knee. The strongest correlation was found between height and maleness. According to the research on visual perception of human gait (Cooper *et al.*, 1982), the center of gravity is always considered as the most important factor influencing human locomotion behaviors (e.g., walking). Therefore, the high correlation of maleness judgment with the height for acoustic stimuli is consistent with this finding for the visual stimuli. We believe that, rather than the height on its own, the center of gravity, which is the combined effect of weight distribution and height, determines the gender of the walker.

The significant correlation between the maleness judgment and the length and width of shoes suggests that size of the shoes may influence the maleness judgment. However, the size of the shoes apparently was not the primary factor that determined the maleness judgement, since M3, M4, and M8 (all wearing M3's shoes) were not judged equally in the maleness. These two incongruent findings about shoe effect motivated experiment 4 which was designed to investigate this issue by having a subset of male and female walkers wear two pairs of new shoes.

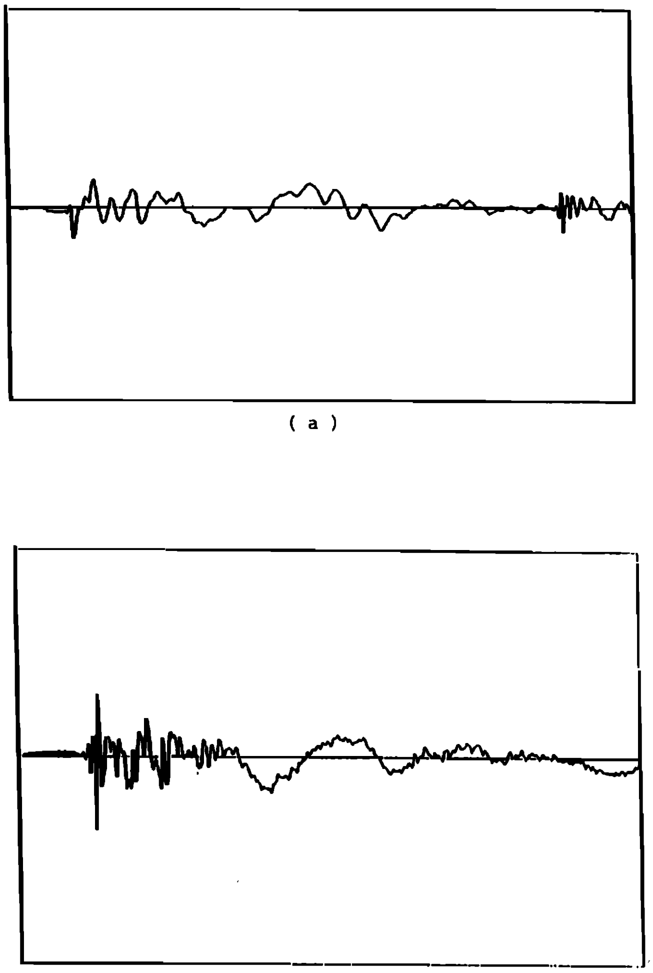
IV. EXPERIMENT 2

Experiment 2 provided a series of physical analyses that identified critical statistical properties of walking spectra (analysis 2 in Fig. 1). In the physical analyses, the gross statistical properties first were derived on the basis of fast Fourier transforms of stimulus waveforms. Hypothetical dimensions were extracted from the gross statistical properties, and these extracted hypothetical dimensions then were interpreted in terms of important classes of information that differentiated walker genders. These properties were correlated with the perceptual data from experiment 1 to identify the physical attributes used by the listeners (analysis 3).

A. Acoustic measurements

1. Spectral analyses

Figure 2 displays the waveforms of the digitized M1 and F1 walking sequences. Panel (a) in Fig. 2 represents a typical walking sequence with both heel and toe strikes, while panel (b) shows a typical walking sequence only with a heel strike. For every walking sequence, a digital waveform editor was used to isolate each single footstep. The beginning of a single footstep was identified by observing a sudden amplitude rise followed by rapid amplitude oscillations. The end of a footstep was marked by the appearance of slow, low-amplitude oscillations representing the noise present between two adjacent footsteps. The beginning and end of each footstep were identified at the nearest zero crossing. The isolation of steps based upon visual inspection of the



(a)

(b)

FIG. 2. The waveform of a single footstep from (a) a male (M1) and (b) female (F1) walker. For the male walker [panel (a)], the moderately rapid oscillation on the left represents the heel strike, and the low, fast oscillation on the right is the toe strike. For the female walker [panel (b)], only heel strike is observed on the left.

waveform was always confirmed by listening to each step.

For some walkers, a measurable toe strike was present. A chi-square test on the contingency of the presence of toe strikes on the gender category indicated that the presence of this physical component seemed to be independent of the actual gender of the walker [chi-square $(1) = 1.0, p > 0.05$]. The presence of toe strikes merely reflected the individual walker's style, and thus, may relate to walker identification rather than gender identification. Therefore (except as noted below for the duration analysis), only heel strikes were used in the spectral analyses. For each walker, 12 individual footsteps were analyzed (three sequences of four steps). A total of 192 waveform tokens (12 steps $\times$ 16 walkers) were used to compute 196 spectra using fast Fourier transform (FFT) with the 1024-point Blackman window.[2] The resulting spectrum consisted of 513 pairs of real and imaginary parts. The following formula was used to calculate the spectral power $P(k)$:

$$P(k) = X_{\text{re}}(k)^2 + X_{\text{im}}(k)^2, \tag{1}$$

where X_{re} and X_{im} are real and imaginary components of a Fourier spectrum, and k denotes the Fourier frequency sample in the range $0 \leqslant k \leqslant 512$. In the following statistical analyses, only 409 Fourier spectral points were used with zero frequency and any frequencies above 4000 Hz discarded.

2. Duration analysis

For each walker, the walking pace was represented as the mean number of footsteps per second. In addition, a cycle of a walking sequence was divided into two types of temporal phases (see the Introduction): the stance phase in which a footstep is accomplished, and the swing phase in which legs alternate. We treated the duration of a footstep as the approximate measure for the stance phase, while the noise duration between two adjacent footsteps as the measure for the duration of the swing phase. These durations were computed by converting the number of computer words for a footstep into milliseconds based upon the 10-kHz sample rate. Due to the different walking pace among the walkers, each of the two durations then was converted into a percentage representing the proportion of the time that each phase occupied in a cycle of a walking sequence.

B. Statistical measures

A mean spectrum was obtained for each walker by averaging the spectra for all 12 steps. The statistical measures were based upon these mean spectra. To describe the shape of spectral distribution, seven statistical measures were calculated by treating each mean spectrum as a probability density function (pdf) of the frequency. The pdf was derived according to the following formula:

$$p(k) = \frac{P(k)}{\Sigma_{i=1}^{409} P(k)}, \tag{2}$$

where $p(k)$ is the probability of the kth Fourier frequency sample. The seven statistical measures were three spectral moments (first, third, and fourth), mode, low- and high-frequency slopes, and average spectrum level. Two principles guided the choice of these statistical measures. First, a set of statistical measures was sought to completely describe the distribution of a spectrum. Second, the statistical measures have been used successfully in previous research on the perception of speech and nonspeech complex auditory events (e.g., Forrest *et al.*, 1988; Freed, 1990).

1. Spectral moments

Each mean spectrum is represented by two equal-length vectors. One vector $f(i)$ is an array of Fourier frequency samples in Hertz, where $1 \leqslant i \leqslant 409$. The other $p(i)$ vector is an array of the proportion of spectral power associated with each frequency sample, relative to the total spectral power [see Eq. (2)]. The linear spectral moments were derived as follows:

$$M_1 = \sum_{i=1}^{409} f(i) p(i), \tag{3}$$

$$M_2 = \sum_{i=1}^{409} [f(i) - M_1]^2 p(i), \tag{4}$$

$$M_3 = \sum_{i=1}^{409} [f(i) - M_1]^3 p(i), \quad (5)$$

$$M_4 = \sum_{i=1}^{409} [f(i) - M_1]^4 p(i), \quad (6)$$

where M_1, M_2, M_3, and M_4 are the first, second, third, and fourth moments for a probability distribution. The first moment is the expected mean of a distribution. The square root of the second moment is the standard deviation of a distribution. The skewness and kurtosis of a distribution are computed as follows:

$$\text{skewness coefficient} = M_3/(M_2)^{3/2}, \quad (7)$$

$$\text{kurtosis coefficient} = [M_4/(M_2)^2] - 3. \quad (8)$$

These dimensionless versions of the third and fourth moments were normalized according to the variance to make the values from different subjects comparable. In this study, we followed Forrest *et al.* (1988) in the use of mean, skewness, and kurtosis as statistical measures.

2. The spectral mode

The mode of a spectral distribution was obtained by identifying the Fourier frequency sample with the maximum probability. Specifically,

$$\text{the mode} = f(i), \quad (9)$$

$$f(i) = \text{MAX } [p(i)], \quad 1 \leqslant i \leqslant 409. \quad (10)$$

The spectral mode provides information about the frequency which carries the most spectral power. In speech literature, formants correspond to such modes and are the bands of energy corresponding to vocal tract resonances. Formants are indispensable cues to distinguish various vowel categories. Since all the spectra had a single major peak, only a single mode was reported for each spectrum.

3. Low- and high-frequency slopes

In his systematic study of the timbre of static spectra, von Bismarck (1974) found that low- and high-frequency slopes substantially influenced the perception of timbre. In the research of human speech perception, Stevens and Blumstein (1981) claimed three types of spectral templates represent three phonetic categories for place of articulation. These three templates correspond respectively to the falling, rising, and compact spectra. This claim, together with empirical evidence, suggests that the slope of a spectrum may be used to define a category of auditory events. In this study, the peak of a spectrum (covering 90% of spectral power in the entire spectrum starting with the second Fourier frequency sample) was treated conceptually as a triangle. The vertex of the triangle corresponds to the peak of a spectral distribution. The two sides of the triangle, which are the low- and high-frequency slopes, were estimated using least-squares linear regression.

4. Average spectral level (ASL)

Freed (1990) found the average spectral level was important to determine the perception of the mallet hardness of a percussive event. In this study, a logarithmic ASL for each mean spectrum was computed by taking the logarithm of the averaging spectral power over 409 frequency samples.

To evaluate the relationship between auditory events and acoustic structure, these seven statistical measures were treated as predictors to discriminate male and female categories (analysis 2 in Fig. 1). To examine the mapping between acoustic structure and perception, the seven statistical measures were used to regress the probability (from experiment 1) that each stimulus was judged as a male walker (analysis 3 in Fig. 1).

C. Results and discussion

1. Duration analysis

For each walker, the walking pace and temporal proportions for the stance and swing phases are displayed in Table III. A two-tail t test was conducted on the walking speeds between the male and female walkers. The absence of the statistical significance failed to confirm the apparent difference in the walking speed between the male and female walkers as noted in Table III [$t(14) = 1.73, p < 0.10$]. The statistical measures for two temporal proportions for stance and swing phase are shown in the second and third columns of Table III. For the stance phase, a two-tail t test indicates that the female walkers had shorter stance phase than the male walkers [$t(14) = 2.31, p < 0.05$]. As expected proportion measures, the female walkers had longer swing phase than the male walkers [$t(14) = 2.31, p < 0.05$]. This finding is consistent with the biomechanical study by Zugina and Leavitt (1974).

To examine whether the overall walking pace and the relative temporal difference in the stance and swing phases between the male and female walkers may affect the gender

TABLE III. Mean temporal measures of the 16 walkers. The numbers in the bottom rows are the means of the corresponding temporal measures and the numbers in parentheses are the standard deviation.

Walker	Stance(%)	Swing(%)	Pace (# of steps per s)
M1	31.2	68.8	1.799
M2	10.8	89.2	1.809
M3	11.9	88.1	1.891
M4	12.2	87.8	1.808
M5	12.6	87.4	1.870
M6	21.3	78.7	1.994
M7	31.3	68.7	1.873
M8	24.3	75.7	1.883
F8	12.1	87.9	1.808
F7	18.6	81.4	1.947
F6	11.2	88.8	1.891
F5	9.5	90.5	1.830
F4	10.9	89.1	1.974
F3	17.2	82.8	2.018
F2	5.4	94.6	1.928
F1	5.1	94.9	2.143
Correlation	0.35	−0.35	−0.68
Male	19.44 (8.75)	80.56 (8.75)	1.87 (0.06)
Female	11.25 (4.87)	88.75 (4.87)	1.94 (0.11)

judgement, we computed pairwise Pearson product correlation coefficients between the maleness (from experiment 1) in Table II and the temporal measures in Table III. These coefficients are reported in the bottom of Table III. The significant correlation ($r = -0.68$, $p < 0.05$) was found between the maleness and the walking pace, indicating that a faster walker is more likely judged as a female, although the walking pace, on average, did not significantly differ between male and female categories.

2. Spectral analyses

Figure 3 displays the mean single-step spectra produced by each of the 16 walkers ordered in terms of perceived maleness (M1 to F1). The spectra were based upon spectral power, rather than on the decibel scale.[3] These spectra clearly differentiate between male and female walkers. By visual inspection, four aspects of spectral differences are noted for male (relative to female) walkers: (1) spectral mode is at a lower frequency; (2) spectral mode is narrower; (3) spectrum is more skewed toward low frequencies; and (4) spectrum has a more rapid (low frequency) rise and (high frequency) fall. These spectral differences are not necessarily independent (as will be confirmed below). The spectral properties represented by these four differences are quantified in terms of the seven statistical measures of the acoustic spectra summarized in Table IV. If these spectral properties (alone or in combination) are used by listeners to judge gender, these properties should be highly correlated with the perceptual continuum of maleness (Table II).

Moderate intercorrelations were found among these statistical measures with the mean absolute correlation coefficient of 0.57 (std = 0.24). This suggests that it may be possible to derive a few descriptive dimensions which will adequately summarize these statistical measures. Therefore, a principal component analysis was conducted on these seven statistical measures. By convention adopted in a principal component analysis, the principal components with the eigenvalues greater than 1 would be retained as interpretable factors. Table V lists the loadings of the two principal components with eigenvalues greater than 1.

The five statistics of mean, mode, skewness, kurtosis, and the low-frequency slope are correlated with the first principal component. These five statistical measures appear to be the description of the shape of the spectral peak, and thus define the nature of the first principal component. Based upon the opposite signs, the first principal component divided the correlations with these five statistical measures into two groups. The value of the first principal component is directly correlated with three statistics (skewness, kurtosis, and low-frequency slope) and is inversely correlated with the other two (spectral mean and mode). The second principal component is correlated with the high-frequency slope and overall spectral level. Since the value of the second principal component increases with the increase of the high-frequency slope and therefore energy in the high-frequency region, we suspect this second component may be responsible for describing the characteristics presented in the high-frequency portion of the spectrum.

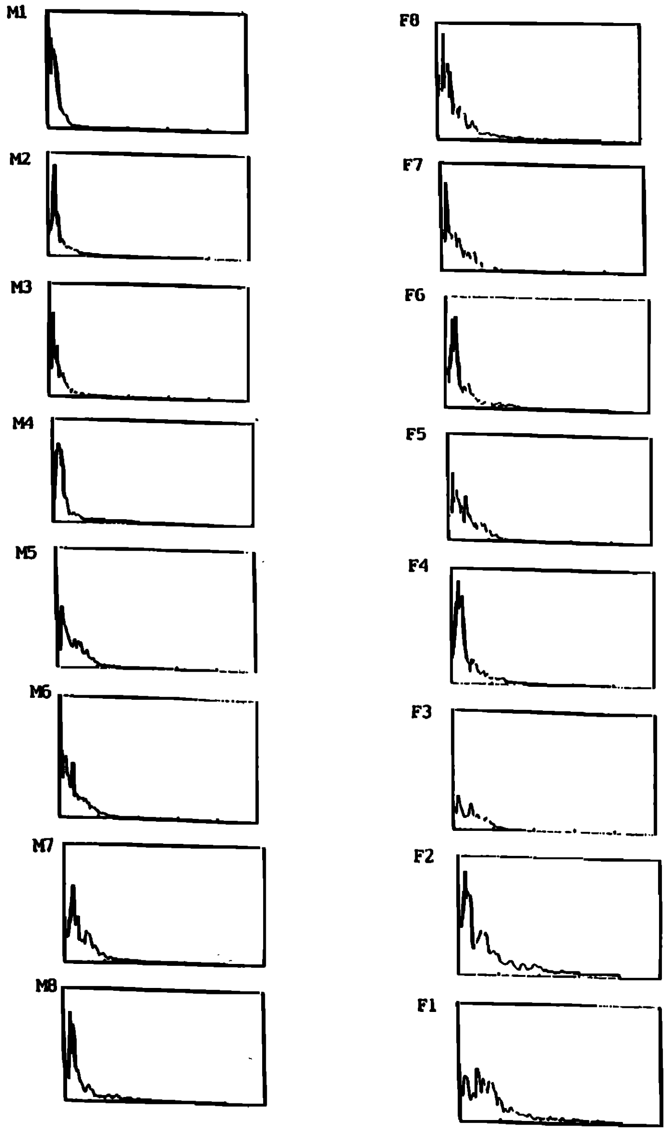


FIG. 3. The mean spectra for the 16 walkers. Letters M and F represent male and female walkers. The number for each spectrum represents the rank order of the perceived gender. The abscissa represents a 5-kHz linear frequency range. The ordinate represents a linear intensity range.

3. Relationship between the auditory events and acoustic structure

The two derived principal components were used in a discriminant analysis to predict the gender categories of the walkers. This analysis provides an additional indication of the relationship between the auditory events and acoustic structure (analysis 2 in Fig. 1). Table VI shows the coefficients for the discriminant function and the results of the classification based on the linear discriminant function. In this linear discriminant function, the male category is coded as 1, whereas the female category is coded as 0. The coefficients indicate that principal components 1 and 2, respectively, were positively and negatively correlated with male category. These correlations suggest that stimuli with (1) low spectral mean and mode, (2) high values of skewness, kurtosis, and low-frequency slope, and (3) low to small high-fre-

TABLE IV. Statistical measures for the 16 walkers. Mode = spectral mode, skew = skewness, kurt = kurtosis, L SLP = low-frequency slope, H SLP = high-frequency slope. For the male and female walkers, the upper numbers in the bottom rows are means and lower numbers are their standard deviation.

Walker	Mean	Mode	Skew	Kurt	L SLP	H SLP	log(ASL)
M1	273.84	29.30	4.68	27.03	372.02	− 13.16	2.55
M2	488.07	36.02	2.92	11.72	157.91	− 18.69	3.38
M3	368.08	29.30	3.39	15.48	237.92	− 4.43	2.45
M4	539.94	36.02	2.42	8.62	133.24	− 10.64	3.39
M5	544.31	136.72	2.68	11.47	2.13	− 2.05	2.49
M6	432.07	39.06	3.00	14.66	149.23	− 3.46	2.61
M7	604.90	234.38	2.51	10.02	39.10	− 10.26	3.31
M8	674.45	175.78	2.05	6.74	45.99	− 5.83	3.28
F8	683.76	175.78	2.12	7.18	30.37	− 8.72	3.43
F7	501.82	136.72	2.82	13.10	14.78	− 2.91	2.60
F6	743.13	244.14	1.88	6.00	54.42	− 5.52	3.41
F5	552.84	126.95	2.68	12.79	14.33	− 2.28	2.59
F4	517.97	175.78	2.70	11.39	9.77	− 15.92	3.36
F3	581.06	136.72	2.74	13.08	3.78	− 0.88	3.32
F2	898.30	185.55	1.49	4.48	11.81	− 1.96	3.09
F1	956.95	468.75	1.55	4.92	0.23	− 1.07	2.77
Male	490.71	89.57	2.96	13.22	142.19	− 8.57	2.93
	129.62	81.22	0.81	6.29	120.63	5.66	0.44
Female	679.48	206.30	2.25	9.12	17.42	− 4.91	2.94
	174.16	112.49	0.56	3.83	17.42	5.16	0.43

quency energy favor male category. Conversely, stimuli with (1) high spectral mean and mode, and significant high-frequency energy tend to be judged as female walking. Moreover, the classification results based upon this linear discriminant function (Table VI) demonstrated an order of maleness for the walkers similar to the perceptual results summarized in Table II (which defined the ordinal relationship for walker labels). A high correlation between the perceptual classification of gender (from experiment 1) and that predicted by the linear discriminant function ($r = 0.82$, $p < 0.05$) qualifies these two derived principal components to effectively delimit the gender categories.

4. Relationship between the acoustic structure and perception

A multiple regression analysis was conducted to regress the maleness continuum (Table II) on the two derived principal components. This analysis provides an evaluation of the relationship between acoustic structure and perception (analysis 3 in Fig. 1). Table VII summarizes the standardized and unstandardized coefficients for the resulting regression equation. These coefficients indicate that subjects judged walker maleness to be higher when the compromised values describing the shape of the spectral peak was increased. The reverse relationship was found for increasing values describing the contribution of high-frequency com-

TABLE V. Results of the principal component analysis. PC1 = the first principal component, PC2 = the second principal component.

Factors	Eigenvalue	Variances accounted for
1	4.44	63%
2	1.66	24%
	Eigenvectors	
Variables	PC 1	PC 2
Mean	− 0.45	0.02
Mode	− 0.38	0.10
Skewness	0.46	0.06
Kurtosis	0.45	0.15
L slope	0.40	− 0.12
H slope	− 0.19	0.68
log(ASL)	− 0.19	0.46

TABLE VI. Results of the discriminant analysis (analysis 2 in Fig. 1). The first part represents the discriminant function of the derived components and the second part represents the classification results based on discriminant function.

Variables	Coefficients		
intercept	− 0.24		
PC 1	0.36		
PC 2	− 0.23		
Walkers	Original gender category	Classified gender category	Prob(male| walker)
M1	male	male	0.98
M2	male	male	0.88
M3	male	male	0.80
M4	male	male	0.71
M5	male	female	0.34
M6	male	male	0.68
M7	male	female	0.48
M8	male	female	0.33
F8	female	female	0.39
F7	female	female	0.43
F6	female	female	0.25
F5	female	female	0.38
F4	female	male	0.69
F3	female	female	0.31
F2	female	female	0.12
F1	female	female	0.04

TABLE VII. Results of the multiple regression analysis for the two principal components (analysis 3 in Fig. 1).

Variables	Unstandardized coefficients	Standardized coefficients	Significance levels
Intercept	0.47	0	0.0001
PC 1	0.10	0.75	0.0004
PC 2	− 0.08	− 0.35	0.0460
Model	R-square = 68%	Adj. R-square = 63%	0.0006

ponents. These findings confirmed our preliminary visual inspection of the walking spectra (see Fig. 2) which seemed to indicate that male walking spectra can be characterized by low central tendency with high positive skewness and kurtosis, and rapid spectral rising and falling. Conversely, the female walking spectra can be characterized by high central tendency with low skewness and kurtosis, and slow spectral rising and falling.

V. EXPERIMENT 3

Experiment 1, together with the various statistical analyses of the spectral statistical properties (experiment 2), suggests that subjects may use two classes of information to identify a walker's gender: the information about the shape of the spectral peak and contribution of high-frequency components. These two classes of information also have been found to be important in the classification of vowels (Chistovich, 1985). The conclusions about the importance of these two complex factors for the perception of walker gender are based upon the evaluation of the pairwise relationships summarized in Fig. 1. If these two principal components are actually used by subjects in classifying walker gender, then it should be possible to alter the classification of the gender of a walker in a predicted manner by the manipulation of these two factors in an otherwise constant stimulus waveform.

Experiment 2 provides a crude, preliminary test of the validity of these conclusions by systematically altering the spectral mode and slopes for two walking stimuli previously judged to be relatively neutral in terms of perceived gender. It is predicted that shifting the spectral mode to lower frequencies and increasing the spectral slope will enhance the perception of maleness. Conversely, a shift to higher frequencies and decreasing the spectral slope should increase the perception of femaleness.

A. Method

1. Stimuli

The M6 and F7 walking sequences used in experiment 1 served as the base stimuli for this experiment. Stimuli from M6 and F7 were selected because each was perceived as relatively neutral for each gender. These base stimuli were passed through one channel of a frequency equalizer (Realistic model 31-2020A) with a 24-dB dynamic range. The equalizer settings represent the experimental manipulations in this experiment. There were five settings for spectral mode (65, 125, 250, 500, and 1000 Hz) which were factorially crossed with each of two levels of spectral slopes (6 and 24 dB/oct) to create 20 experimental conditions for each walker. The octave band representing the mode was set to maximum amplification, with the settings of the lower and higher octave bands reduced by 6 or 24 dB per octave to a maximum attenuation of 24 dB relative to the setting for the spectral mode.

The stimuli were 4-kHz low-pass (antialiasing) filtered, then passed through the frequency equalizer, and recorded on tape cassette (TEAC W-350). The 40 stimuli (20 conditions $\times$ 2 walkers) then were digitized (10-kHz sample rate, 4-kHz low pass). The stimuli were presented to listeners over TDH-49P headphones through a 12-bit digital-to-analog converter at approximately 75 dB(A).

Although the stimuli are described in terms of the settings of the frequency equalizer, these stimulus labels do not necessarily represent the actual spectral characteristics of the stimulus waveforms. The equalizer served to modify the spectral properties of the original waveforms and thus provides only a rough manipulation of spectral mode (in octave band steps) and a crude, very limited manipulation of spectral slopes for attack and decay. For example, because the original spectrum for walker M6 has a spectral mode at 39 Hz and a steep high-frequency slope, the stimuli based upon a modification of the walker M6 stimulus never will have a shallow high frequency slope. Furthermore, all high spectral modes for stimuli based upon walker M6 will be an octave band in width. These manipulations of spectral parameters are adequate for a simple verification of stimulus parameter importance, and are not intended to quantify these parameters. A more precise specification of parameter values critical for perception would require the development of a means to synthesize walking stimuli which vary systematically for specific parameters (with other stimulus parameters held constant).

2. Subjects

Fifteen SUNY—Binghamton undergraduate students participated in the experiment for course credit. Subjects were seated in a commercial acoustic chamber. All subjects reported normal hearing. The data for five out of these 15 subjects were eliminated because they failed to respond (at least during the computer-monitored interval) on more than 70% of the trials; the remaining subjects responded on almost every trial. Therefore, the results reported in this experiment were based upon responses from ten subjects.

3. Procedure

This experiment consisted of nine blocks of 40 trials (20 conditions $\times$ 2 walkers). The instructions and method of the stimulus presentation were the same as those used in experiment 1.

B. Results and discussion

For each subject, the average percentage of a stimulus being judged as male over nine blocks of trials was calculated for each experimental condition. A $5\times2\times2\times2$ (5 levels of mode $\times$ 2 levels of gender of the base stimuli $\times$ 2 levels of slope of attack $\times$ 2 levels of slope of decay) within-subject analysis of variance was conducted on these average percentages. A main effect of spectral mode indicated that judgments of maleness were altered significantly by the spectral mode [$F(4,36) = 62.17$, $p < 0.01$]. The main effect of original stimulus source gender was not significant [$F(1,9) = 0.00$, $p > 0.05$].

Figure 4 shows the maleness judgment as a function of spectral mode. The solid line in Fig. 4 is collapsed across both the original stimuli (M6 and F7) and spectral slopes. The other two functions are for the M6 and F7 stimuli with 24-dB low- and high-frequency spectral slopes (discussed below). These data demonstrate that judgment of maleness increased from an original 50% to 90% or more as the spectral mode was shifted to lower frequencies (at least to 125 Hz). Perception of these stimuli also decreased to 10% male (therefore, 90% female) with the mode at 1000 Hz. Therefore, the crude manipulation of spectral mode was successful in altering the perception of gender.

The main effect for low-frequency slope was also significant [$F(1,9) = 54.43$, $p < 0.01$], but opposite to expected, with more male judgments at shallow (6 dB per octave) than steep (24 dB) low-frequency slope. This finding (not illustrated in Fig. 4) can probably be attributed to the crude nature of the experimental manipulations. The original M6 stimulus has a sharp mode at 39 Hz. With the equalizer set for a 62-Hz mode, the original peak is reduced by 6 dB, or 1/2 the original voltage. Based upon examination of the original spectral distribution, the mode for the M6 stimulus at this setting still will be in the very low-frequency region, and not in the 62-Hz octave band. With the 24 dB per octave setting, we always will effectively shift the spectral mode to the equalizer setting. Therefore, our experimental manipulation does not allow us evaluate spectral slope independent of the spectral mode, and it is not surprising that there was a significant interaction between mode and spectral slopes for both low frequency [$F(4,36) = 10.38$, $p < 0.01$] and high frequency [$F(4,36) = 19.41$, $p < 0.01$]. However, since the basic findings for spectral mode should be reasonably valid for the stimuli with the 24-dB high- and low-frequency slopes (see functions for M6 and F7 in Fig. 4), this confounding of spectral slope and mode should not alter our validation of the importance of spectral mode using the 24-dB slope results.

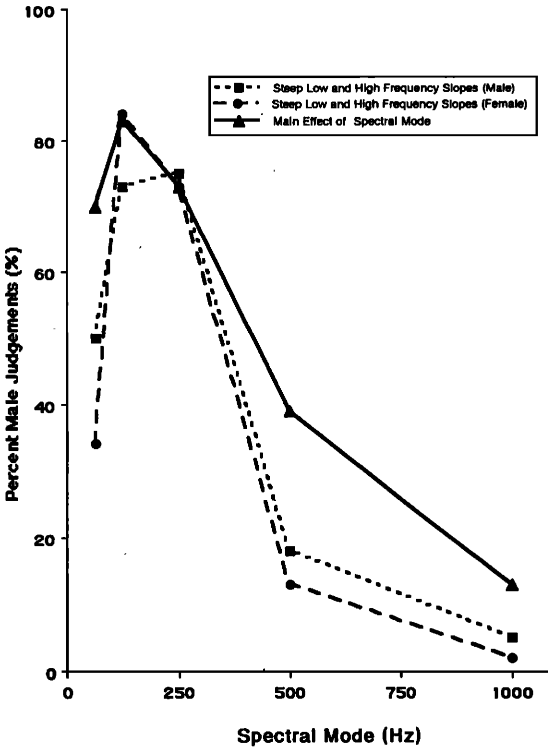


FIG. 4. The probability of maleness judgment as a function of spectral mode.

The M6 and F7 functions for the 24 dB per octave stimuli are shown in Fig. 4. There was a significant interaction between mode and original gender for the base stimuli [$F(4,36) = 9.70$, $p < 0.01$]. This significant interaction and the absence of a main effect of the original gender indicate that the energy distributions over frequencies for the two walkers (M6 and F7) chosen for the base stimuli were not the same. M6 contained more significant low-frequency energy than F7, with the original mode for F7 located near 125 Hz. This indication is consistent with the findings in the physical analyses (see Table IV).

Three conclusions can be drawn from this experiment. First, manipulating spectra can alter perceived judgment of walker maleness judgments. Second, maleness judgments can be based upon spectral differences between male and female walkers independent of the original gender prior to the spectra manipulations. Third, spectral mode plays a major role in classification of walker gender.

VI. EXPERIMENT 4

Experiments 1 and 2 discovered two incongruent findings about the effect of shoe. First, the maleness judgment about M8 differed from M3 and M4 even though all wore the same shoes (i.e., M3's shoes). This finding seemed to argue against the effect of shoe upon the maleness judgment. Second, a significant correlation between size of shoes and maleness judgment seemed to indicate the shoes might still have effects on the maleness judgment, although correlation cannot be used to infer causation between pairs of variables. Experiment 4 was designed specifically to investigate the effect of shoe on the maleness judgment by asking three female and a male walker to wear two pairs of new male shoes (single size) from different manufacturers. If types of shoes do affect the maleness judgment, these female walkers should be judged as male walkers most of the time, and for the male walker, the magnitude of maleness judgment should depend upon the type of shoes.

A. Method

1. Stimuli

The three female walkers (F2, F6, and F8) and one male walker (M7) were asked to wear two pairs of M7's shoes. Other aspects of the stimulus production were the same as those in experiment 1. For each walker in each pair of shoes, and for the three female walkers also in their own shoes, one of the three four-footstep walking sequences was randomly selected as the experimental stimulus. Thus, the experimental stimuli consisted of 11 walking sequences (three female walkers in three pairs of shoes and a male walker in two pairs of shoes).

2. Subjects

Eight SUNY—Binghamton undergraduate students were paid for their participation in this experiment. All subjects reported normal hearing.

3. Procedure

Each subject was instructed to indicate the gender of the walker by listening to a four-footstep walking sequence. The subjects finished 11 blocks of 33 trials (11 walking sequences $\times$ 3 repetitions) with the first block treated as practice trials.

B. Results and discussion

Table VIII shows the probability of the four walkers judged as a male walker when they wore the two pairs of the male shoes (i.e., M7's) and their own shoes. Table VIII shows that, for the male walker, the magnitude of maleness judgment differs between the two pairs of shoes. For the female walkers, the similar dependency of maleness judgment upon the shoes is also manifested.

Because shoes A and B are M7's shoes, M7 did not have data under "own shoes" in Table VIII. Therefore, a chi-square (contingency table) test was conducted on the frequencies of reporting a male walker contingent upon the female walkers and type of the shoes. The main effect of types of shoes indicated that maleness judgment for a walker can be changed when the person wears shoes of the opposite sex [chi-square(2) = 150.87, $p<0.05$]. Table VIII shows that the greatest increase in reporting a male walker occurred when the walker wore shoes A. However, the magnitude of the change in the maleness judgment depends upon the walker, which was indicated by the presence of the interaction between the walker and types of shoes [chi-square(4) = 43.29, $p<0.05$]. Additionally, the maleness judgment differed among the walkers themselves regardless of types of shoes [chi-square(2) = 69.39, $p<0.05$], which was consistent with the findings about the actual perception of gender reported in experiment 1.

This simple experiment demonstrated that type of shoe does influence the maleness judgment. The perception of a walker gender can sometimes be reversed by wearing shoes of the opposite sex. However, both experiments 1 and 4 reported the different magnitude of maleness judgment among walkers, suggesting that the type of shoes is not a primary factor responsible for the walker gender identification. Conversely, this experiment suggests that the walker gender identification is more likely determined by the integration of multiple factors, including those identified in the first three experiments.

TABLE VIII. The probability of three female and a male walkers in male shoes judged as a male walker.

Walker	Shoes A	Shoes B	Own Shoes	Mean
F8	0.92	0.69	0.39	0.67
F6	0.71	0.32	0.30	0.44
F2	0.70	0.44	0.06	0.40
M7	0.83	0.40	···	0.62
Mean	0.79	0.46	0.25	

VII. GENERAL DISCUSSION

Using the methodological paradigm illustrated in Fig. 1, the current study investigated the perception of source gender for acoustic stimuli produced by humans walking on a hard surface. The perception of walker gender is important primarily (1) in demonstrating that listeners can identify some source characteristics from properties of acoustic signals; and (2) in demonstrating the effectiveness of the current approach in identifying the properties of, and the relationship between, the stages summarized in Fig. 1. Specifically, this study suggested that people may judge the gender of a walker by assessing two classes of information presented in walking spectra: (1) the spectral peak including shape and location of the spectral peak; and (2) the contribution of the distribution of high-frequency energy. Additionally, this study also demonstrated that there were true differences in anthropomorphic measures between male and female walkers. These differences were highly correlated with the maleness judgment. In the following discussion, we will propose a preliminary conjecture that may link sound articulators with their correlates of the spectral characteristics, and therefore, to account for the mechanism with which the gender identification is undertaken.

Various theories have been proposed to understand how people perceive objects in the environment (e.g., see review of theories of perception by Gordon, 1989). Among various theories, Gibson's ecological (direct) perception theory is related to the current study (Gibson, 1979). Rather than using acoustic stimulation of footsteps, the ecological perception theory asserts that people directly perceive the informational invariance which exists in natural events. The different patterns of invariance between male and female footsteps would be recovered by the organism to perceive the male and female walkers. However, this theory leaves unspecified many details about the process, and thus, does not lead directly to specific empirical predictions.

In contrast with the direct perception theory, a theory based upon detailed inference computations exists in the literature. Richards and Bobick (1988; also, Richards, 1988), recently quantified by Bennett *et al.* (1989), argued that the perceptual system analyzes various proximal data to recover the natural mode of the object. The perceptual system has the built-in inference machinery to conduct inductive (rather than deductive) inference tests. The natural mode of an object category is claimed to exist because structure in the environment is not arbitrarily organized and the object properties are characterized as modes along various dimensions. For example, the sound that the object emits reflects information about the structure of the source. If a sound is narrow band, then the source must have a system of tuned resonators (e.g., human speech), which is unlikely possessed by an inanimate object. Thus, all objects belonging to inanimate categories should be rejected as the source. Additionally, the fundamental frequency is usually determined by the size of the source (for example, male speakers with the large vocal cavity usually have a low fundamental frequency). Thus the fundamental frequency of the sound may indicate the size of the source. These examples suggest that various properties of a sound may provide indications for the presence of certain characteristics of the source.

By analyzing the process of a sound production, we can identify three components which are involved in the sound production: (1) power; (2) oscillators; and (3) resonators (Fant, 1960; Richards, 1988). These three sound production components together determine the characteristics of a sound. Humans may classify a sound by calculating the contributions of these three components on the basis of spectral information. The current study varied two of these three components: power and oscillators varied, with the resonators roughly fixed (since the walkers were walking on the same floor). Within this context, we discovered the true anthropomorphic differences between male and female walkers. For example, typical female walkers usually have lighter weight and are shorter than male walkers, and typical source properties determine typical power differences between the male and female walkers. Furthermore, we found in experiment 4 that the gender perception of a female walker could be altered when she wore male shoes. This shoe effect was also supported by the presence of a high correlation between the maleness judgment and size of shoes. Shoe characteristics contribute to the oscillator differences between the male and female walkers. As an acoustic consequence, the differences in these two sound production components may result in high pitch and some high-frequency components in female walking spectra. Humans may map these acoustic consequences to sound classes stored in memory with the physical parameters used to produce these acoustic consequences. Therefore, the perception of source characteristics might be accomplished by integrating acoustic consequences, and then establishing the association between the acoustic consequences and the sound classes.

We should emphasize that this conjecture differs from the claim that the perception of a walker gender is based upon a stereotypical association in which a single stimulus property (e.g., high pitch) is always associated with the feminine source. On the contrary, we argue that this stereotypical coincidence is the by-product of the true differences in the source.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Grant BNS8911456 to the third author. Any opinions, findings and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. Further support was provided by the Center for Cognitive and Psycholinguistic Sciences at the State University of New York at Binghamton. We thank Michael Hall, Bruno Repp, William Yost, and an anonymous reviewer for their valuable comments on earlier drafts of this paper, and Jennifer Cho, Renee Gauci, and Paul Thompson for their assistance in conducting this research. The cooperation of the Anderson Center for the Performing Arts at the State University of New York at Binghamton is greatly acknowledged. Robert Logan currently is affiliated with International Business Machine (IBM) Corporation, Poughkeepsie, New York.

[1] This ability has certain benefits for survival in graduate school, but discussion of such Darwinian aspects of this ability is not consistent with survival of the current acoustic research study.

[2] A 1024-point Blackman window is large relative to the temporal resolving power of the human auditory system. This window was employed in this initial study because it allowed each step to be represented in terms of a static spectrum, and thus, facilitated the derivation and interpretation of gross statistical properties of the stimuli. Our findings indicate that this relatively gross analysis does provide an adequate first approximation of the stimulus properties which reflect subject perception of gender. A finer, dynamic analysis which may (or may not) provide a more adequate description of stimulus properties, is beyond the scope of this initial investigation of source perception.

[3] All walking stimuli were recorded in a relatively quiet, natural environment, but still were embedded in a low, relatively constant background noise. Therefore, signal-to-noise (S/N) ratio was determined primarily by stimulus amplitude. Spectra based upon relative amplitude would have tended to enhance the contribution of the background noise for low S/N and thus make cross-spectra comparisons more difficult. For this reason, spectra were based upon spectral power.

Barclay, C. D., Cutting, J. E., and Kozlowski, D. M. (**1978**). "Temporal and spatial factors in gait perception that influence gender recognition," Percept. Psychophys. **23**, 145–152.

Bennett, B. M., Hoffman, D. D., and Prakash, C. (**1989**). *Observer Mechanics* (Academic, New York).

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics, National Research Council). (**1989**). *Classification of Complex Nonspeech Sounds* (National Academy, Washington, DC).

Chistovich, L. A. (**1985**). "Central auditory processing of peripheral vowel spectra," J. Acoust. Soc. Am. **77**, 789–805.

Cooper, J. M., Adrian, M., and Glassow, R. B. (**1982**). *Kinesiology*, (Mosby, St. Louis), 5th Ed.

Cutting, J. E. (**1978**). "Generation of synthetic male and female walkers through manipulation of a biomechanical invariant," Perception **7**, 393–405.

Cutting, J. E., Proffitt, D. R., and Kozlowski, L. T. (**1978**). "A biomechanical invariant for gait perception," J. Exp. Psychol.: Hum. Percept. Perform. **4**, 357–372.

Fant, G. (**1960**). *Acoustic Theory of Speech Production* ('s-Gravenhage, Mouton).

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (**1988**). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," J. Acoust. Soc. Am. **84**, 115–123.

Freed, D. J. (**1990**). "Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events," J. Acoust. Soc. Am. **87**, 311–322.

Gibson, J. J. (**1979**). *The Ecological Approach to Visual Perception* (Houghton-Mifflin, Boston).

Gordon, I. E. (**1989**). *Theories of Visual Perception* (Wiley, New York).

Grey, J. M. (**1977**). "Multidimensional perceptual scaling of musical timbres," J. Acoust. Soc. Am. **61**, 1270–1277.

Grey, J. M., and Gordon, J. W. (**1978**). "Perceptual effects of spectral modifications on musical timbres," J. Acoust. Soc. Am. **63**, 1493–1500.

Howard, J. H. (**1983**). "Perception of simulated propeller cavitation," Human Factors **25**, 643–656.

Jassem, W. (**1979**). "Classification of fricative spectra using statistical discriminant function," in *Frontiers of Speech Communication Research*, edited by B. Lindblom and S. Ohman (Academic, London), pp. 71–91.

Johansson, G. (**1973**). "Visual perception of biological motion and a model for its analysis," Percept. Psychophys. **14**, 201–211.

Kozlowski, L. T., and Cutting, J. E. (**1977**). "Recognizing the sex of a walker from a dynamic point-light display," Percept. Psychophys. **21**, 575–580.

Pols, L. C. W., van der Kamp, L. J. Th., and Plomp, R. (**1969**). "Perceptual and physical space of vowel sounds," J. Acoust. Soc. Am. **46**, 458–467.

Potter, R. K., and Steinberg, J. C. (**1950**). "Toward the specification of speech," J. Acoust. Soc. Am. **22**, 807–820.

Repp, B. H. (**1987**). "The sound of two hands clapping: An exploratory study," J. Acoust. Soc. Am. **81**, 1100–1110.

Richards, W. (**1988**). *Natural Computation* (MIT, Cambridge, MA).

Richards, W., and Bobick, A. (**1988**). "Playing twenty questions with nature," in *Computational Processes in Human Vision: An Interdisciplinary Perspective*, edited by Z. Pylyshyn (Ablex, Norwood, NJ).

Stevens, K. N., and Blumstein, S. E. (**1981**). "The search for invariant acoustic correlates of phonetic features," in *Perspectives in the Study of Speech*, edited by P. D. Eimas and J. L. Miller (Erlbaum, Hillsdale, NJ).

Tousman, S. A., and Pastore, R. E. (**1989**). "Source characteristics: A study of hand clapping," J. Acoust. Soc. Am. Suppl. 1 **85**, S53.

Warren, H., and Verbrugge, R. R. (**1984**). "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," J. Exp. Psychol.: Hum. Percept. Perform. **10**, 704–712.

Wedin, L., and Goude, G. (**1972**). "Dimension analysis and the perception of instrumental timbre," Scand. J. Psychol. **13**(3), 228–240.

von Bismarck, G. (**1974**). "Sharpness as attribute of the timbre of steady sounds," Acustica **30**, 159–172.

Zuniga, E. N., and Leavitt, L. A. (**1974**). "Analysis of gait: A method of measurement," in *Biomechanics IV*, edited by R. C. Nelson and C. A. Morehouse (University Park, Baltimore, MD).