THE SOUND OF TWO HANDS CLAPPING: AN EXPLORATORY STUDY*

Bruno H. Repp

Abstract. Clapping is a little-studied human activity that may be viewed either as a form of communicative group behavior (applause) or as an individual sound-generating activity involving two "articulators"--the hands. The latter aspect was explored in this pilot study by means of acoustical analyses and perceptual experiments. Principal components analysis of 20 subjects' average clap spectra yielded several dimensions of interindividual variation that were related to observed hand configuration. This relationship emerged even more clearly in a similar analysis of a single clapper's deliberately varied productions. In perception experiments, subjects proved sensitive to spectral properties of claps: For a single clapper, at least, listeners were able to judge hand configuration with good accuracy. Besides providing some general information on individual variations in clapping, the present results support the general hypothesis that sound emanating from a natural source informs listeners about the changing states of the source mechanism.

Introduction

Clapping, the production of sound by striking the hands together, is perhaps the most common audible activity of humans that is (a) intended to be heard by others and (b) does not involve either the vocal tract or a musical instrument. It is practiced by virtually all individuals from an early age and, probably, in all cultures. Its most frequent function, at least in Western society, is to signal approval, in which case it is a rhythmic, repetitive activity maintained for at least several seconds, often collectively in a group. Given the widespread occurrence and the communicative function of clapping, it is surprising that scientific studies of this activity are difficult to find.

While research on clapping may not be of the highest priority, the topic offers a surprising variety of aspects to investigators who, prompted by curiosity, might wish to explore a little-studied human behavior. Thus

---

sociologists and historians might be interested in the role of clapping in different cultures and in the evolution of conventional applause in Western society (see Jenniches, 1969; Victoroff, 1959). Musicologists might want to explore the use of clapping in various kinds of folk music. Acousticians might be challenged to explain the generation of clapping sounds by applying acoustical theory. Students of motor behavior might wish to study clapping as a skill requiring precision, bimanual coordination, and auditory feedback.

For psychologists (represented by the author) two different aspects of clapping behavior seem of interest. The first, more obvious one, is the communicative function of clapping. Thus it might be asked how people convey their degree of enthusiasm for a performance, how their clapping behavior varies as a function of the stimulus and their state of mind, how a performer judges an audience's reaction from the applause, etc. While these topics are worthy of study, they are not the ones explored in the present investigation. This study, rather, pursues questions that arise when clapping is viewed as an individual articulatory activity, not unlike certain events occurring in the course of speaking.

To be sure, clapping and speaking have only few things in common. Communicative aspects of clapping may have certain parallels in paralinguistic features of speech, conveyed by parameters such as rate and loudness, which modulate the basic articulatory activity. Here we are concerned with another commonality: In both activities, the sound produced at any instant in time reflects the configuration of adjustable articulators that are part of the human body: the two hands in one case, and the various parts of the vocal tract in the other. The analogy is closest when brief transients in speech are considered, such as stop consonant release bursts or clicks, whose durations are similar to those of claps (see, e.g., Ladefoged & Traill, 1984; Repp, 1983; Fre Woldu, 1985). The dependency of sound properties on the configuration of the source mechanism follows from acoustical theory: Variations in the configuration will have systematic acoustical consequences. To the student of perception, be it of clapping or of speech, this means that the sound carries information about the momentary state of the articulators (as well as about their dynamic change, if the brief signal permits it) that can be apprehended by listeners who have (innate or acquired) knowledge of the constraints under which the source mechanism operates (cf. Gibson, 1966; Liberman & Mattingly, 1985; Neisser, 1976). Human listeners almost certainly have such knowledge available about the vocal tract and about a variety of environmental events (Jenkins, 1985); human hands should be no exception. Just as stop consonant release bursts convey information about vocal tract size (presumably) and configuration (e.g., Blumstein & Stevens, 1980), so claps may convey information about hand size and configuration.

This idea provided a useful point of departure for this preliminary investigation of the production and perception of claps. More specifically, the questions addressed were: What sorts of sounds are claps? What different ways are there of producing them? How much information do their acoustical properties contain about hand size and configuration? How sensitive are listeners to that information in the acoustic signal? Answers to these questions would not only increase our knowledge about a little-studied human activity but also would be relevant to the theoretical notion that there are general principles of perception-production relationships that extend across both speech and nonspeech domains.

Being a first exploration, the present study was fairly broad in scope but crude in some aspects of execution. The focus was on spectral properties of claps; rate and intensity (which are of much greater relevance to social communication) were considered only in passing. Analyses of clap spectra were conducted to determine how, and how consistently, information about hand size and about different hand configurations is acoustically represented. As a number of subjects were employed, the question of individual differences in clapping style necessarily entered the picture. A computer classification was conducted to explore the extent of intra- versus inter-individual variability in clap spectra, and two subsequent perceptual studies tested human listeners' ability to extract from claps information about hand size and hand configuration, respectively.

I. Production Study

A. Methods

1. Subjects

The subjects were 10 male and 10 female individuals between the ages of 25 and 45, all researchers, graduate students, or technicians at Haskins Laboratories.

2. Recording Procedure

Subjects were seated, one at a time, in a sound-insulated booth, with their hands about 60 cm from a Sennheiser microphone.[1] An Otari MX5050 tape recorder with peak indicator lights was located in an adjacent booth. Care was taken to set the recording level so that no peak distortion occurred. Each subject was asked to clap at his or her most comfortable rate, "the way you would normally clap after an average concert or theater performance," for about 10 seconds. The length and width of the subject's left hand were then measured with a ruler, from the wrist to the tip of the middle finger and across the palm above the thumb, respectively, and notes were taken on the hand configuration observed during clapping.

3. Acoustical Measurement Procedures

All recordings were digitized at a sampling rate of 10 kHz, with low-pass filtering at 4.9 kHz. From each subject's recording, a sequence of 10 consecutive claps was excerpted, starting a few claps into the series. Clap onsets were located using an automatic thresholding procedure, and onset-to-onset intervals (OOIs) were measured. The mean OOI and its standard deviation within a series provided measures of a subject's clapping speed and rhythmicity, respectively.

The FFT spectrum of each individual clap was calculated from the first 10 ms of the waveform, which generally occupied about 20 ms.[2] Subsequently, the spectra (each quantized in computer memory as a series of levels in 20-Hz bands) were averaged arithmetically over the 10 claps in a series to yield a subject's average clap spectrum. These average spectra were subjected to further analysis, as described below.

The relative amplitudes of the individual claps were estimated by the following rough procedure: A 20-ms Hamming window was moved in 10-ms steps across each subject's file of 10 digitized claps, and the maxima in the resulting series of dB values were taken to represent the clap amplitudes. Since some individuals were recorded on different days, and distance from the microphone was not precisely controlled, these amplitudes did not accurately reflect individual differences in clap intensity but merely represented the relative intensities of the claps as recorded (and as played back to the subjects in the perceptual experiments). The mean amplitude and its standard deviation within a series provided measures of a subject's recorded clapping strength and regularity, respectively.

B. Results and Discussion

1. Rate and Amplitude Measurements

Although rate and amplitude measures were not of primary interest, they are reported here for the sake of completeness and because they played a role in the perceptual experiments. The average "comfortable" rate of clapping was 4/s (mean OOI = 250 ms). Individual rates ranged from 2.7/s (OOI = 366 ms) to 5.1/s (OOI = 196 ms). There was a nonsignificant tendency for males (OOI = 265 ms) to clap slower than females (OOI = 236 ms), $t(18) = 1.59$, $p < .10$. If real, this difference could either be due to the fact that males, because of their generally larger arms and hands, have a larger mass to move in clapping, or it could represent a sex difference that is independent of size. The male subjects indeed had substantially larger hands (length x width = 162 cm^2 on the average) than the female subjects (126 cm^2), $t(18) = 6.89$, $p < .001$. The overall correlation between hand size and OOI reached significance ($r = 0.44$, $p < .05$). Computed separately for each sex, however, the correlation tended to hold up only for males ($r = 0.55$, $p < .10$), not for females ($r = 0.09$). In any case, only a small fraction of the individual differences in rate was accounted for by this factor.

Temporal variability was 6.8 ms on the average (range: 2.8 to 13.6 ms), or 2.7 percent of the mean OOI (range: about 1 to 5 percent). It should be noted that the subjects had not been instructed explicitly to clap as regularly as possible, and greater regularity could probably be achieved by most subjects under more controlled conditions. Even so, the lowest standard deviations probably are close to the maximum regularity attainable in clapping. Temporal variability showed neither any significant difference between males and females nor any relation to hand size.[3]

Clap amplitudes as recorded did not differ significantly between males and females. Amplitude standard deviations within a series ranged from 0.7 to 5.2 dB across subjects. They showed no sex difference and did not correlate with temporal variability ($r = -0.02$).

2. Spectral Analysis

The average clap spectra of the 20 subjects are shown in Figure 1. Whereas the averages are quite representative of the individual clap spectra (see section I.B.4), there is considerable variability in spectral shapes across individuals. In the figure, the spectra are arranged roughly according to visual similarity. The shapes range from a rather flat, rising type to those with a pronounced mid-frequency peak (between 2 and 3 kHz), those with
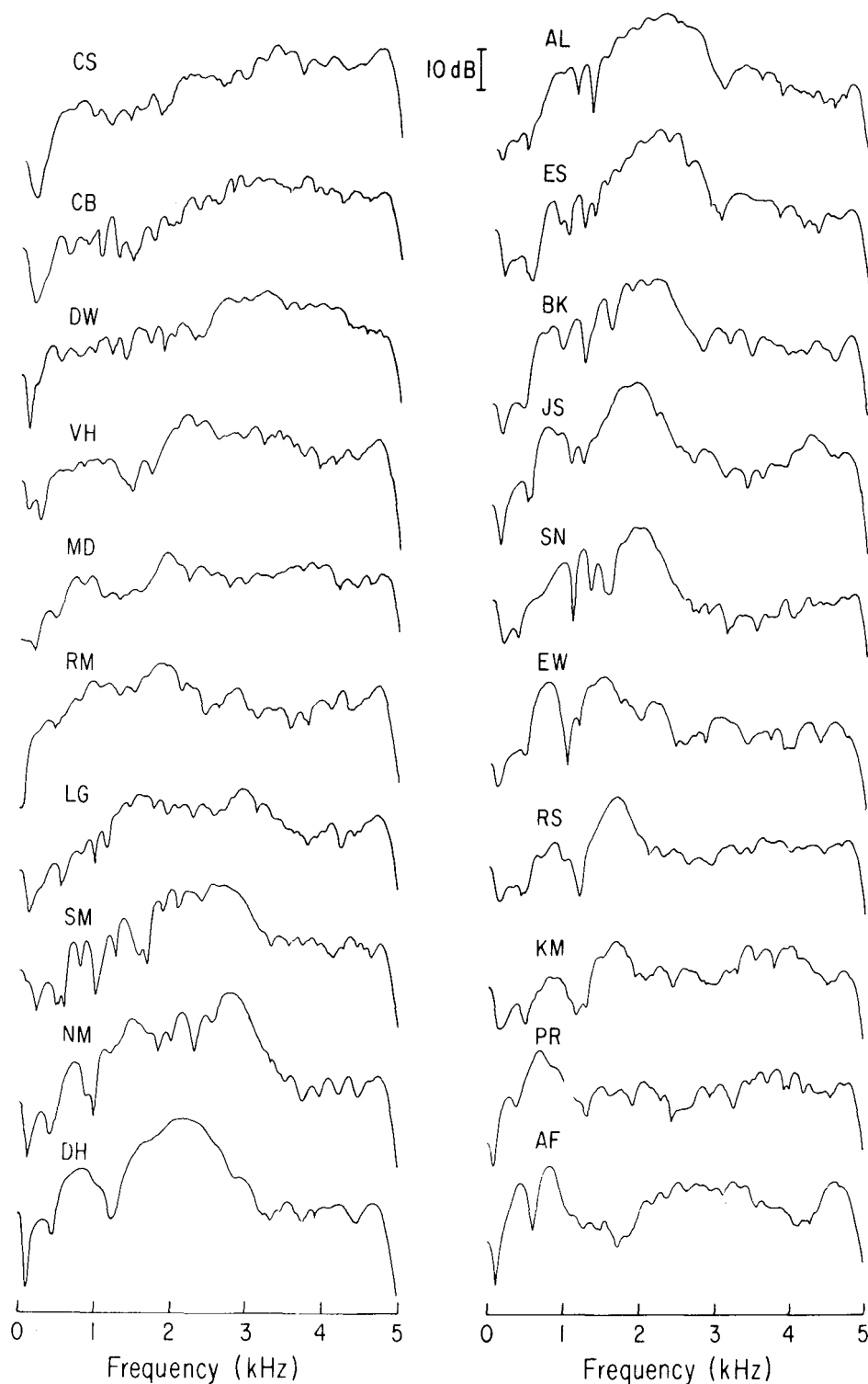
Figure 1.   Average FFT spectra of claps from 20 subjects.   Each   spectrum   is
the   arithmetic   average   of   the   spectra   (levels   in   dB)   of 10
individual claps, computed over a 10-ms   window   starting   at   clap
onset.   The spectra have been amplitude normalized and include high
frequency pre-emphasis.   They   are   arranged   roughly   according   to
visual similarity.

an emerging second peak below 1 kHz, and finally some with only this low-frequency peak.

For purposes of statistical analysis, it was desirable to quantify spectral shape in some way. A principal components factor analysis with Varimax rotation (which maximizes the variance of factor loadings for each input spectrum; see Harman, 1967) was conducted for this purpose. The input to the analysis was the set of 20 average spectra, each represented by 256 numbers (levels in 20-Hz bands). The 20 x 20 intercorrelation matrix was computed, and its linear decomposition yielded four significant factors (i.e., with eigenvalues greater than 1), which together accounted for 88 percent of the variance among subjects' clap spectra. These factors represent prototypical spectral shapes whose linear combinations (weighted by the factor loadings specific to each subject) approximate the 20 input spectra.[4]

The spectral shapes of the four factors are plotted in Figure 2, and the factor loadings of the 20 input spectra (i.e., their correlations with the factors) are listed in Table 1 in the order corresponding to Figure 1. The first factor, which accounts for 39 percent of the variance, is characterized
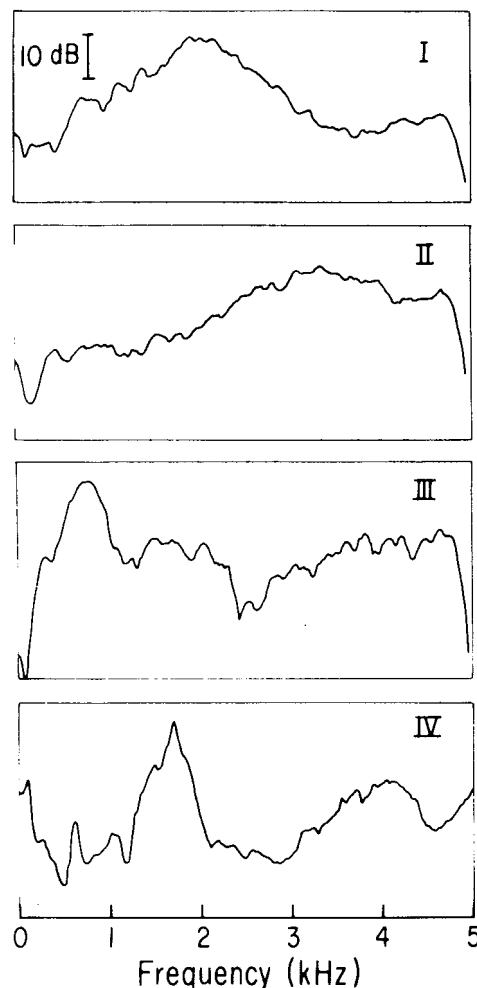

Figure 2. Spectral representation of the four principal factors, obtained by converting the (standardized) factor scores into levels (dB).

Table 1

Factor loadings (I-IV) of the 20 subjects' average clap spectra,
with observed hand configuration (Hands) and listeners' hand
configuration judgments (Ratings). See text for explanation.

Subject	Sex	I	II	III	IV	Hands	Ratings
CS	F	0.068	0.929	0.216	0.026	A2	2.86
CB	F	0.080	0.959	-0.020	-0.039	A2	2.82
DW	M	0.154	0.919	0.100	-0.020	A3	1.95
VH	F	0.530	0.725	0.052	-0.290	P2.5	1.95
MD	F	0.435	0.738	0.348	0.119	a2	2.36
RM	M	0.683	0.172	0.519	0.087	A2.5	1.95
LG	M	0.686	0.599	0.014	0.146	A2.5	2.09
SM	F	0.772	0.504	-0.177	-0.126	A3	2.23
NM	F	0.828	0.346	-0.074	-0.086	P2	2.73
DH	M	0.934	0.060	0.127	-0.047	A3	2.23
AL	F	0.889	0.358	0.089	0.045	P3	1.77
ES	M	0.890	0.341	-0.040	0.129	A2.5	2.45
BK	M	0.846	0.280	0.286	0.175	a3	1.91
JS	M	0.798	0.046	0.456	0.304	A3	1.09
SN	F	0.795	-0.210	0.297	0.182	A2	2.00
EW	M	0.610	-0.035	0.591	0.147	A2	1.18
RS	F	0.478	0.369	0.354	0.649	a2	2.59
KM	M	0.301	0.663	0.234	0.533	A3	2.41
PR	M	-0.001	0.294	0.903	-0.089	a1	1.18
AF	F	0.093	0.500	0.347	-0.702	A1	1.05

by a broad spectral peak in the vicinity of 2 kHz. More than half of the
input spectra have substantial loadings in this factor, with subject DH being
the closest match (cf. Fig. 1). The second factor, which accounts for 29
percent of the variance, represents spectral upward tilt, or strong
high-frequency components without any pronounced peaks. A number of spectra
have high loadings in this factor, with subject CB being the closest match.
Some spectra, such as that of subject LG, represent a mixture of the two
factors.   The third factor, which accounts for 12 percent of the variance,
represents a narrow peak below 1 kHz together with a notch around 2.5 kHz.
Only one spectrum, that of subject PR, has a high loading on this factor;
several others have moderate loadings. Some spectra, such as that of EW,
constitute mixtures of factors one and three. Note that not all spectra with
peaks below 1 kHz load on the third factor, only those without a pronounced
mid-frequency peak. Finally, the fourth factor, which accounts for 8 percent
of the variance, represents a narrow peak below 2 kHz and a broader peak
around 4 kHz.   There are no clear instances of this pattern among the input
spectra, but several spectra have moderate loadings, including one (subject
AF) with a negative loading (i.e., an inverted pattern). Subject RS has the
most eclectic pattern, with moderate loadings in all four factors. (Note that
the Varimax rotation, which aims for "simple structure," minimized the
occurrence of such cases.) The individual spectrum with the smallest amount of
variance accounted for by the four factors (74 percent) is that of subject EW.

The factors extracted, especially the first three, provide a useful framework for characterizing the shapes of clap spectra. In addition, they furnish numerical indices (the factor loadings) of the degree to which individual spectra resemble the factor prototypes. This quantification of spectral features permits statistical analyses to be conducted that would otherwise be impossible. Thus a multivariate analysis of variance was performed on the factor loadings to determine whether spectral shapes differed between males and females. There was no significant sex effect overall or for any of the four factors individually. This implies not only that males and females clapped similarly, but also that hand size had no important influence on the clap spectrum.

3. The Relation of Clap Spectra to Hand Configuration

The absence of a sex difference in clap spectra suggests that hand configuration, rather than hand size, is the most important determinant of the sound pattern and accounts for the individual differences observed. As a first step toward a better understanding of this variable, the author recorded himself clapping in eight different ways ("modes"), which are illustrated in Figure 3. Modes P1-P3 kept the hands parallel and flat but changed their vertical alignment from palm-to-palm (P1) to fingers-to-palm (P3), with P2 halfway between these extremes (i.e., with the right hand lowered by about 4 cm). Modes A1-A3 varied alignment in a similar way, but with the hands held at an angle. (Note that modes P1 and A1 differ in that the fingers of the two hands strike each other in P1 but not in A1. Modes P3 and A3 are more similar to each other.) Since the hands automatically tended to be more relaxed (slightly cupped) in the A modes than in the P modes, two additional versions of A1 were recorded, with the hands either very cupped (A1+) or flat (A1-), so as to examine the effect of this variable. Three parameters were thus manipulated in a semi-independent fashion: hand alignment, angle, and curvature.

All recordings were digitized, 10 consecutive claps were excerpted from each, and average spectra were calculated, which are shown in Figure 4. The spectral variation observed was somewhat smaller than expected, but nevertheless informative. Mode P1 yielded a rather flat spectrum, but a mid-frequency peak started to emerge, and low-frequency energy decreased, as the parallel hands became increasingly misaligned (modes P2 and P3). Similarly, displacement of the hands held at an angle (going from A1 to A3) led to a relative increase in mid-frequency energy and to a decrease of low-frequency energy. The palm-to-palm claps (P1, A1, A1-, A1+) all showed peaks below 1 kHz but no mid-frequency peak. Extreme cupping (A1+) or stretching of the hands (A1-) had relatively little effect on the spectrum.

These visual impressions were confirmed by entering the eight average clapping mode spectra together with the four factor shapes from the earlier analysis into another principal components analysis, in which the earlier (orthogonal) factors served as "marker variables." The factor loadings that emerged from this analysis are listed in Table 2. Again, four factors accounted for 88 percent of the variance. As can be seen from the factor loadings of the marker variables, the original factor I was second in the present analysis, the original factor III came out first, and the original factor II was third. The reason for these shifts in relative importance was the absence of very strong mid-frequency peaks (factor I) in the author's clap spectra, whereas low-frequency peaks (factor III) were very consistently

94

Figure 3. Eight clapping modes (see text). These still photographs were posed after the recording session.
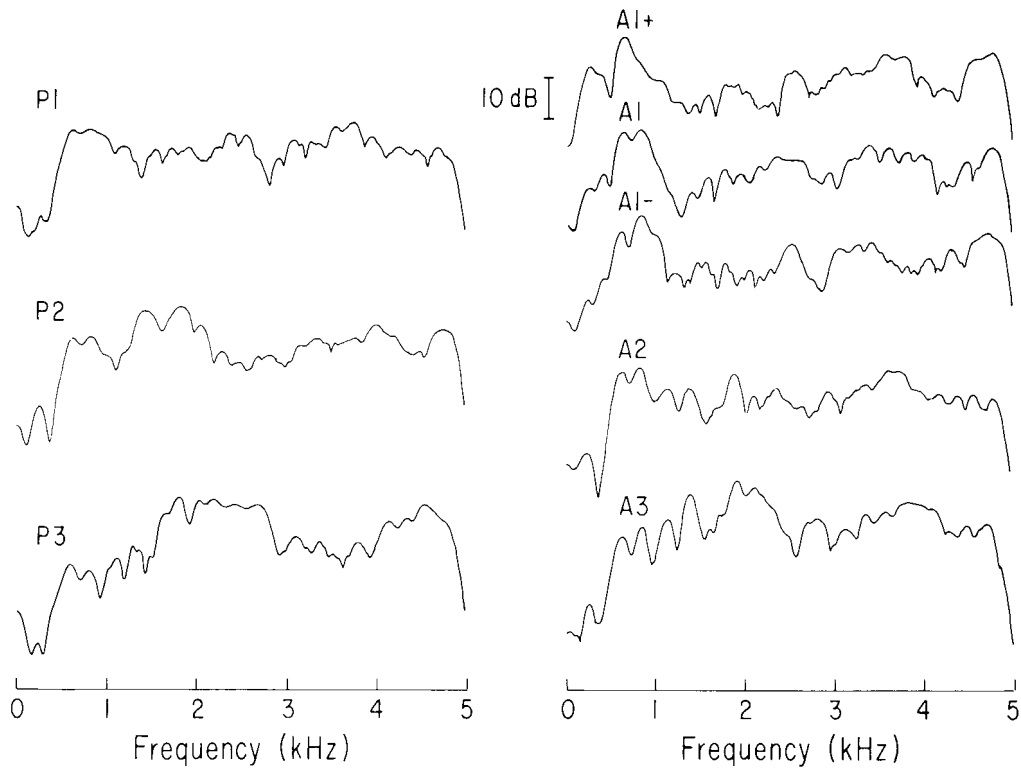
Figure 4.   Average amplitude-normalized FFT spectra of the author's   claps   in
eight different clapping modes (see Fig. 3).

Table 2

Factor loadings (I-IV) of author's clap spectra from eight different
clapping modes.   Factors from earlier analysis (Table 1)
serve as marker variables (FI-FIV). Also shown are
subjects' hand configuration judgments (Ratings).

Mode	III	I	II	IV	Ratings
P1	0.710	0.495	0.305	0.104	2.32
P2	0.475	0.614	0.164	0.553	2.36
P3	0.077	0.762	0.500	0.108	2.95
A1	0.861	0.107	0.262	-0.180	1.55
A2	0.676	0.536	0.240	0.185	1.91
A3	0.362	0.766	0.294	0.307	2.64
A1 -	0.820	-0.272	0.223	-0.155	2.00
A1 +	0.840	0.149	0.199	-0.085	1.00
FI	-0.121	0.939	-0.132	-0.119	
FII	0.255	0.172	0.926	0.047	
FIII	0.856	0.122	-0.317	0.208	
FIV	-0.182	0.042	0.021	0.947	

present.   (The original numbering of the factors has been maintained in the table to avoid confusion.) The modes with high loadings in the low-frequency peak factor (III) were P1, A1, A1-, and A1+--those in which the two palms struck each other. Modes P2 and A2, with partial contact between the palms, had moderate loadings in this factor, and modes P3 and A3, where the palms did not touch, had the smallest loadings. These latter modes, however, had the highest loadings on the mid-frequency peak factor (I); modes P2 and A2, in which there was partial contact between the fingers of the right hand and the palm of the left hand, correlated moderately with this factor, and so did mode P1. No modes had high loadings on factors II and IV; moderate loadings were exhibited by modes P3 and P2, respectively.

This analysis leads to the conclusion that the low-frequency peak represents the palm-to-palm resonance, and the mid-frequency peak represents the fingers-to-palm resonance. The interpretation of the other two factors is less clear. The spectral upward tilt factor may simply represent a failure to achieve strong resonances due to insufficient force or lack of a sufficient seal around the hand contact areas, which is most likely to occur at intermediate hand alignments. It may also represent a fingers-to-fingers resonance.

We may now return to the 20 subjects' data and examine whether the same relation between factor loadings and hand configuration holds for them. Table 1 presents, following the factor loadings, a rough classification of the subjects' hand configurations, as observed at the time of recording. (Lower-case "a" denotes a small angle, and 2.5 a position close to 3.) The correlations between the factor loadings and the numerical hand position scores (neglecting hand angle) were, in order of magnitude: I (r = 0.57, p < .01), III (r = -0.54, p < .01), IV (r = 0.38, p < .10), II (r = -0.01). Thus fingers-to-palm clappers tended to show mid-frequency peaks (factor I) but not low-frequency peaks (factor III), as predicted from the analysis of the author's clapping modes. Because of other sources of variability, the relationship was less tight in this group of subjects. In a stepwise multiple regression analysis of the same data, factor I accounted for 33 percent of the variance, and factor II, though initially uncorrelated with the hand position scores, accounted for an additional 20 percent, while factors III and IV made no further contribution. Factor II thus seems to represent an aspect of hand configuration that is independent of factors I and III, whose loadings tend to be negatively correlated.

On the whole, it appears that the observed variations in hand configuration are responsible for about half of the spectral variability among individuals. The unexplained variation may derive from such factors as hand curvature and stiffness, fleshiness of the palms, tightness of the fingers, precision, and striking force, that could not be assessed accurately in this exploratory study. A more careful assessment of the roles of hand angle and finger contact also remains to be conducted.

4. Automatic Classification of Clap Spectra

The foregoing analyses were conducted on the subjects' average clap spectra. No attempt was made to assess quantitatively the amount of intra-individual spectral variation. Nevertheless, it seemed important to determine whether subjects were sufficiently consistent from one clap to the next to maintain distinctive individual characteristics. For that purpose,

the correlations between the 200 individual clap spectra and the 20 average spectra were computed. Whenever an individual clap's spectrum was most highly correlated with the same subject's average spectrum, this was considered a correct identification. The computer thus simulated the "clapper identification" performance of an ideal human listener who is thoroughly familiar with each subject's characteristic way of clapping. Of the 200 claps, 181 or 90.5 percent were classified correctly in this way. No two individuals were consistently confused; the errors that occurred did not follow any particular pattern. This must be considered a remarkably high success rate, indicating that subjects maintained distinctive individual characteristics in their clapping, despite a certain amount of variability from one clap to the next, and despite often similar hand configurations across individuals. In the present sample of 20 subjects, at least, no individual made exactly the same sounds as any other.

II. Perception Studies

A. Perception of Hand Size, and Self-recognition

In contrast to the computer of the foregoing simulation, humans generally know little about each other's ways of clapping, so they cannot be expected to recognize individuals from their clapping sounds. If the following experiment was nevertheless presented to the subjects as one of individual clapper identification, it was primarily for the subjects' amusement. The primary purpose of the study was to determine whether subjects could extract some information about the clappers' sex and thus about their hand size. (The experiment was conducted before the results of the acoustical analyses became available, which suggested that there is little hand size information in the spectrum.) In addition to spectral information, the present listeners also had rate and loudness available as possible (but probably unreliable) cues to a clapper's physical size. A secondary purpose of the experiment was to find out whether listeners could recognize their own clapping.

1. Methods

Eighteen of the 20 subjects used in the production study served as listeners; two females (CS, NM) who were unavailable were replaced by Haskins colleagues of the same sex. All subjects were known to each other, with one exception (CS), who did not participate as a listener. The stimuli consisted of the 20 clapping excerpts (10 successive claps each) in random sequence, with 5 seconds of silence in between. The subjects were seated individually in a sound-insulated booth and listened to the test tape monaurally (right ear) over THD-39 headphones at a comfortable intensity. Each subject first listened to the whole stimulus sequence without responding, for purposes of familiarization. Then the tape was presented a second time, and subjects were asked to guess who had been clapping by writing down the initials of three different individuals for each excerpt, in order of confidence. An alphabetic list of the names of the 20 clappers was provided on the answer sheet. Subjects were permitted to use each name as a response as often as they liked or not at all; in fact, however, they tended to be fairly even-handed in their response choices.

2. Results and Discussion

In the analysis of the data, three points were assigned to a correct first guess, two to a correct second guess, and one to a correct third guess. Thus overall percent correct scores were calculated with respect to a possible maximum score of 60. Chance performance was at 5 percent correct.

Overall, clapper recognition was 11 percent correct with self-recognition scores excluded (13 percent correct otherwise), which is poor but significantly above chance ($\underline{t}(19)$ = 3.74, $\underline{p}$ < .001). Self-recognition, however, was much higher: 46 percent correct. That almost half of the 18 relevant subjects were able to recognize their own clapping among 20 excerpts indicates that clapping does convey stable individual characteristics, as did also the automatic classification exercise described earlier. Memory for their specific behavior during the recording session may have aided some subjects.

The question of primary interest was whether subjects were able to determine the clappers' sex and thus the size of the hands that produced the sounds. For this purpose the data were rescored in terms of "male" responses, disregarding the specific initials put down. The chance level for this score is 50 percent correct. The obtained score, with self-judgments excluded, was 54 percent correct (56 percent correct otherwise), which is barely above chance.[5] The correlation of average judged masculinity with clappers' measured hand size ($\underline{r}$ = 0.36, $\underline{p}$ < .10) fell short of significance.

The low sex recognition scores might suggest that subjects' responses were largely random. This was not the case, however. Subjects were very consistent in thinking that certain clappers were either male or female, though they were often wrong. The most striking instance was the clapping of the smallest female in the group (AF), which was judged as "male" 99 percent of the time. What variables influenced the subjects' responses?

To answer this question, the average percentages of "male" judgments for the 20 clappers were entered into a stepwise multiple regression analysis together with eight independent variables: average OOI, temporal variability, average amplitude, amplitude variability, and the factor loadings on the four spectral shape factors (I-IV). Four of these variables made a significant contribution to the regression equation and together accounted for 85 percent of the variance. OOI emerged as the most significant factor, accounting for 44 percent of the variance ($\underline{r}$ = 0.67). Subjects thus <u>expected</u> males to clap slower than females--an expectation that, however, was only weakly supported by the actual temporal measurements (hence the low accuracy of sex recognition). Second in importance, accounting for an additional 14 percent of the variance, was amplitude: Louder claps were considered more "male." (In fact, there was no such sex difference in the recordings.) The variable third in importance was factor IV, whose inclusion in the regression equation increased the variance accounted for by another 15 percent. This effect was probably due largely to AF's clapping which, it will be recalled, had the highest (negative) loading on factor IV and was overwhelmingly identified as "male." Finally, factor III added another 11 percent to the variance accounted for, indicating a tendency of subjects to consider low-frequency resonances as "male." It will be recalled that loadings in this factor did not differ between male and female clappers.

All these response trends may reflect general sex stereotypes (males: slow, loud, low-pitched; females: fast, soft, high-pitched) rather than tacit or explicit knowledge of sex differences in clapping behavior, of which there was no evidence in the present subject sample. It is conceivable, of course, that this sample was not representative, and that subjects' judgments do reflect expectations based on actual differences in clapping behavior in the population-at-large. All that can be concluded from the present data is that listeners are sensitive to a variety of physical parameters of claps, not only rate and intensity but also spectral aspects.

B. Perception of Hand Configuration

In hindsight, after the acoustical analyses revealed no effects of hand size on the clap spectrum, the poor recognition performance of the subjects in the preceding experiment is not surprising. The demonstration that subjects are sensitive to physical parameters of claps, however, leads to the question of whether subjects can judge hand configuration from the sound of claps, since that variable was a major determinant of the spectrum. A second perception experiment was conducted for this purpose.

1. Methods

Twenty-two new subjects participated in this study, partially Yale student volunteers (for whom the brief test was tacked on to the end of a paid experimental session) and partially Haskins researchers who were unfamiliar with the previous clapping experiments. The same stimulus sequence as in the preceding experiment was used. In addition, however, the author's eight clapping mode excerpts were recorded in two different randomizations. The first of these served as familiarization, without any responses being required. For the second randomization, and for each of the following 20 excerpts, the subjects judged which hand configuration was used by choosing from the numbers "1,2,3." The three configurations corresponding to these judgments were illustrated by photographs of the author's hands in modes A1, A2, and A3, respectively (see Figure 3), which remained visible to the subjects throughout the experiment. The subjects were told that the first 8 excerpts represented a single person clapping in different ways, whereas the following 20 excerpts derived from different people, each clapping in his or her most comfortable way. The instructions also mentioned specifically that hand configuration affects the sound of claps, but not in which way.

2. Results and Discussion

The data were reduced by computing the average rating of each excerpt by the 22 subjects. An average score of 1.0 thus means that all subjects judged these claps as having been produced in a palm-to-palm position, a score of 3.0 means complete agreement on a fingers-to-palm position, and intermediate scores represent either agreement on an intermediate position or various amounts of disagreement among subjects. In fact, subjects' judgments were quite systematic, and while there was some variability, no excerpt received a bimodal response distribution (i.e., more "1" and "3" judgments than "2" judgments).

Let us consider first the responses to the author's eight clapping modes. The average ratings are shown in the last column of Table 2. It is evident that the subjects were able to recognize the different hand configurations.

They seemed more accurate with modes A1-A3 than with modes P1-P3, which all sounded more like fingers-to-palm to them, perhaps because of the greater flatness of the hands and the added finger contact in P1 and P2. (This may also have been an artifact of illustrating the hand positions with photographs of modes A1-A3.) Subjects were also able to distinguish the three versions of the A1 mode, despite the relatively small spectral differences among them (see Figure 4), by translating degree of cupping into hand configuration estimates.

Analysis of variance confirmed these impressions. In one repeated-measures analysis, modes A1+ and A1- were omitted, and hand angle and position were the two crossed factors. There were highly significant main effects for both angle, $F(1,21) = 43.35$, $p < .0001$, and position, $F(2,42) = 22.94$, $p < .0001$, but no significant interaction, $F(2,42) = 1.59$, $p = .2158$. Thus, although it seemed that subjects were better at distinguishing hand configurations when the hands were held at an angle, this tendency was not reliable. In a second analysis, the three degrees of hand cupping for the A1 mode (A1+, A1, A1-) were compared. The main effect of this variable was highly significant also, $F(2,42) = 22.48$, $p < .0001$.

The average hand position ratings were entered into a stepwise multiple regression analysis, with the loadings in the four spectral factors (Table 2) as independent variables. Factor III alone accounted for 73 percent of the variance in subjects' judgments, with high factor loadings corresponding to low (palm-to-palm) hand configuration ratings ($r = -0.85$). None of the other three factors made a significant additional contribution, even though the loadings in each of them correlated positively with subjects' ratings. The principal determinant of subjects' judgments, then, seemed to be the presence and extent of low-frequency peaks in the spectrum.

For the ratings of the 20 subjects' excerpts a similar regression analysis was conducted, with OOI, temporal variability, amplitude, and amplitude variability as additional independent variables.[6] Although these variables were not considered relevant to the judgment of hand position, they were included because of their perceptual salience, and also to make the analysis comparable to that conducted earlier on the masculinity scores. Three variables made a significant contribution, explaining 72 percent of the variance. Surprisingly, OOI came out first, explaining 44 percent of the variance (longer OOIs, or slower rates, leading to more palm-to-palm judgments); factor III accounted for a further 16 percent, and factor IV for another 12 percent. These results resemble those of the immediately preceding analysis in that they reveal a significant influence of the low-frequency peak factor (III) on subjects' judgments. However, they also resemble the results of the analysis of the masculinity scores, with the main difference being the total absence of any correlation of hand position ratings with amplitude.

The last-mentioned similarities raise the questions of whether masculinity and hand configuration judgments were related, and whether OOI had any true relation to hand configuration. Indeed, the correlation between the two types of judgments was high ($r = -0.82$, $p < .001$), which confirms that the listeners (different groups in the two tests) relied largely on the same acoustical information in judging sex (hand size) and hand configuration. The spectral information did reflect hand configuration, at least in part, whereas it had no obvious relation to hand size. Average OOI, however, was not related to either clappers' sex or hand size. Actual hand configuration (derived from the "Hands" column of Table 1) was likewise uncorrelated with

OOI ($\underline{r}$ = 0.05). Unfortunately, actual hand configuration was also uncorrelated ($\underline{r}$ = 0.19) with judged hand configuration. Therefore, it is not clear whether the listeners were really able to perceive or infer what the 20 clappers did with their hands. The large variations in the irrelevant rate parameter may have diverted subjects' attention from the relevant spectral properties. Subjects' success in the preceding test based on the author's clapping modes suggests that they would perform more accurately if irrelevant variation were reduced.

III. General Discussion

A. Methodological Shortcomings

The present study was a first exploration of a hitherto little-studied subject, and it was conducted under time constraints. As such, it suffers from a number of methodological weaknesses that need to be improved upon in a more thorough follow-up study. These weaknesses shall be acknowledged before proceeding to the conclusions.

First, the recording procedure was far from optimal. Future studies will have to avoid reverberation by using a sufficiently large or anechoic chamber, and distance from the microphone will have to be controlled more carefully. The data, however, provide no indications of serious artifacts due to these factors.

Second, the spectral analysis was based on low-pass filtered signals. (See also Note 2.) Future analyses may reveal that there is additional spectral information in frequencies above 5 kHz.

Third, the registration of subjects' hand configuration was casual and possibly inaccurate (except for the author's own clapping modes). More precise ways will have to be found for recording hand position (as well as angle, degree of cupping, etc.) by means of measurements in situ, from still photographs, or from video tapes.

Fourth, by asking a number of subjects to clap in their most comfortable ways, differences in hand configuration were confounded with a variety of other individual differences. It would be desirable to separate these aspects in a future study by asking each individual to clap in different, precisely specified "modes," as was done here with a single subject (the author).

Finally, subjects' ability to infer hand configuration from the sound of claps was probably impaired by the presence of irrelevant but salient variations in rate and loudness, as well as by the elimination of the higher frequencies in the spectrum. To test subjects' full ability, it would be desirable to present high-quality recordings in which rate and loudness variations are neutralized.

B. Conclusions and Further Questions

With these caveats, then, what conclusions can be drawn from this pilot study, and what questions do they raise or perhaps even answer?

First, it is evident that different individuals clap in different ways.
This simple fact raises interesting questions about the origin of these
individual differences--questions that the present study could not even begin
to address, but that are worth listing here: To what extent are individual
differences in clapping anatomically conditioned, and to what extent to they
represent learned behavior patterns? If an individual's preferred hand
configuration, in particular, is learned, when and how did this learning take
place? How consistently do individuals employ a particular way of clapping,
and to what extent do they vary their behavior across different situations?
The assumption here has been that situational factors lead primarily to
adjustments in clapping rate and loudness--parameters that are relevant to the
ordinary communicative function of applause--but not to changes in
characteristic hand configuration. There may be some people, however, who do
vary their hand configuration systematically or randomly, so that they could
not be said to have a characteristic way of clapping at all. It is also
possible that adjustments in hand position are contingent on large changes in
rate (see Note 3) and loudness.

Second, apart from variations in rate and loudness, which are of
secondary interest here, different individuals produce different clapping
sounds. A considerable part of that spectral variability appears to be due to
differences in hand configuration. Other factors must contribute to the
spectral shapes, however, or else it would not have been possible to classify
over 90 percent of individual clap spectra correctly by computer. What these
factors are is not clear at present. The success of the computer
classification analysis suggests that individuals may have a "clap
signature"--a characteristic spectrum that distinguishes them from many other
individuals. To support this suggestion, however, it will be necessary to
assess intra-individual variability over a wider range than merely a train of
10 consecutive claps, and also to eliminate possible artifactual contributions
from variations in recording conditions.

Third, no sex differences in clapping were evident in the present group
of subjects. While sex differences as such were not of particular interest
here, the finding does contradict popular opinion that "ladies clap
differently from gentlemen." The present subjects, all Ph.D.'s or graduate
students, did not seem to fit these traditional categories of social demeanor.
It remains to be seen whether a sample drawn from the population-at-large will
show the differences that are often attributed to the sexes, and/or whether
such differences emerge only in real-life situations. More to the point of
the present study, however, it appears that hand size--which exhibits clear
sexual dimorphism--does not have any influence on the sound of claps. This is
an unexpected finding.

Fourth, the spectral differences among claps, as well as their rate and
loudness, were readily discriminated by listeners and were systematically
related to their judgments of clappers' presumable sex and hand configuration.
The most salient parameter was rate: Slower rates were considered to
represent a male clapper and a palm-to-palm hand position, even though rate
was in fact unrelated to both sex and hand configuration in the present sample
of subjects. Thus the listeners relied on expectations or stereotypes that
linked these variables. Spectral properties of claps, which were correlated
with actual hand configurations, also contributed to listeners' judgments. In
the case of a single clapper (the author), it was quite clear that subjects
were able to determine hand configuration from the sound of claps. In the

case of the more heterogeneous sample of 20 clappers, the evidence was not conclusive.

C. Theoretical and Practical Issues

At the theoretical level, the results of the present study give some support to the hypothesis that sound emanating from a natural source, particularly one involving parts of the human body, conveys perceptible information about the configuration of that source. The prime example of the principle embodied in this hypothesis is speech, whose sounds convey the changing states of the vocal tract. In the case of listening to continuous speech, there is little awareness of the pure sound qualities (the proximal stimulus), and perception is focused on the distal events. It has been argued that the distal speech events are perceived directly, without mediation by an auditory representation of the input (Fowler, 1986; Liberman & Mattingly, 1985). This argument is less convincing, however, when applied to the common laboratory situation of individual speech sounds (e.g., fricative noises or stop consonant release bursts) that are removed from their context and presented in isolation (e.g., Blumstein & Stevens, 1980; Repp, 1981). Listeners then do perceive characteristic auditory qualities as well as the articulatory information behind them, so the former could, in principle, be used to infer the latter. Listening to claps is like listening to isolated stop release bursts in that auditory, pitch-like qualities are perceived together with, presumably, the "place of articulation" on the clapper's hands. It is a moot point whether listeners arrive at judgments of hand configuration from claps directly, as it were, or via an inferential process based on perceived sound quality. Actually, this question becomes unnecessary if perception itself is viewed as involving unconscious inference (Rock, 1983). It seems plausible to assume that perception of isolated speech sounds differs from clap perception only in the availability of well-established phonetic categories to classify speech stimuli. The perceiver's tacit knowledge of the constraints under which parts of the body operate, and the consequent possibility of deriving articulatory information even from static spectral properties (cf. Stevens & Blumstein, 1981), may be similar in the two cases. Of course, when it comes to longer stretches of (time-varying) speech, the information to be perceived becomes much more complex than that in isolated sounds.

It is more difficult to say anything convincing about the practical utility of the present research. After all, it focused precisely on those parameters of clapping that presumably play no role in the communicative function of applause. Two aspects, however, may be of slight interest to the pragmatist. The possibility of an individual "clap signature," though it is in need of much stronger empirical support, may be of interest to those concerned with automatic recognition of individuals from acoustic signals. Devices are on the market now that are said to respond to claps, and it might be suggested that they could be tuned to respond selectively to different individuals or to different hand configurations of the same individual. Another possible application of knowledge gained from a study of clapping might be in music performance. The hands might be considered as a percussion instrument with the capability of producing two or more timbres, and while this is not an impressive range, the instrument is cheap, portable, easy to maintain, and readily mastered. Apart from the universal use of clapping for purely rhythmic purposes, the capability of the hands to produce different timbres may in fact already have been discovered by some folk musicians.[7] If

so, more detailed knowledge about the production and perception of clapping may help in analyzing such existing practices, and also may lead to their deliberate introduction into some contemporary art music as a welcome humanizing element.

References

Blumstein, S. E., & Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. Journal of the Acoustical Society of America, 67, 648-662.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. Journal of Phonetics, 14, 3-28.

Fre Woldu, K. (1985). The perception and production of Tigrinya stops. RUUL 13. Uppsala, Sweden: Uppsala University, Department of Linguistics.

Gibson, J. J. (1966). The senses considered as perceptual systems. Boston, MA: Houghton Mifflin.

Harman, H. H. (1967). Modern factor analysis. Chicago, IL: University of Chicago Press.

Jenkins, J. J. (1985). Acoustic information for objects, places, and events. In W. H. Warren & R. E. Shaw (Eds.), Persistence and change. Proceedings of the First International Conference on Event Perception. Hillsdale, NJ: Erlbaum.

Jenniches, K. M. (1969). Der Beifall als Kommunikationsmuster im Theater. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 21, 569-584.

Ladefoged, P., & Traill, A. (1984). Linguistic phonetic descriptions of clicks. Language, 60, 1-20.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.

Neisser, U. (1976). Cognition and reality. San Francisco, CA: Freeman.

Repp, B. H. (1981). Two strategies in fricative discrimination. Perception & Psychophysics, 30, 217-227.

Repp, B. H. (1983). Coarticulation in sequences of two nonhomorganic stop consonants: Perceptual and acoustic evidence. Journal of the Acoustical Society of America, 74, 420-427.

Rock, I. (1983). The logic of perception. Cambridge, MA: MIT Press.

Stevens, K. N., & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (Eds.), Perspectives in the study of speech. Hillsdale, NJ: Erlbaum.

Victoroff, D. (1959). El aplauso, una conducta social. Revista Mexicana de Sociologia, 21, 703-739.

Zahorian, S. A., & Rothenberg, M. (1981). Principal-components analysis for low-redundancy encoding of speech spectra. Journal of the Acoustical Society of America, 69, 832-845.

Footnotes

[1]The recording environment and procedure were not optimal but were deemed adequate for this pilot study. Distance from the microphone was not controlled precisely, and some reverberation was present.

[2]A short window was used to exclude reverberation as much as possible. The FDI program of the ILS package (Version 4.0, Signal Technology Inc.) was used to compute the spectrum. This program employs a fixed window of 25.6 ms duration and fills the unused portion with silence. The program also uses a Hamming window by default, which was maintained (unnecessarily) in the present

analyses. Reanalysis of several claps without the Hamming window and/or using a window of longer duration revealed only minimal changes in the spectrum.

[3]A temporal analysis was also conducted of each subject clapping as fast as possible. The clapping rates achieved under these instructions ranged from 5.4/s (OOI = 184 ms) to 8.1/s (OOI = 123 ms), with an average of 6.6/s (OOI = 152 ms). Although the instructions requested that the hand configuration remain the same, many subjects stiffened their hands and reduced hand excursion to an extent that would rarely be encountered in natural applause. There was no difference between the fast clapping rates of males (OOI = 149 ms) and females (OOI = 154 ms), nor was there any relation to hand size (r = -0.28), even though limitations imposed by the mass of the limbs might have been expected to be revealed more clearly in this extreme situation. The average variability of fast clapping was 6.2 ms (4.1 percent), with no sex difference and no significant correlation with OOI (r = 0.23). The correlations between normal and fast clapping rates (r = 0.35) and between variability measures at normal and fast rates (r = 0.30) were nonsignificant.

[4]It should be noted that this analysis differs from the type of principal components analysis commonly conducted on speech spectra (e.g., Zahorian & Rothenberg, 1981), in which the correlations are computed for all pairs of frequency bands across a number of different spectra. (In the present case, this would have resulted in a 256 x 256 intercorrelation matrix.) The factors emerging from such an analysis represent spectral components such as formant peaks, whereas the present factors represent full spectra that instantiate the types of spectral shapes observed for a group of subjects. In other words, the more common analysis is meant to uncover dimensions underlying spectral shape, whereas the present analysis was employed primarily as a data reduction procedure.

[5]A significance test of the difference from chance becomes meaningless in view of the enormous variation of scores (from 1 to 97 percent correct) across stimuli, to be discussed below.

[6]These variables were not analyzed for the author's clapping modes. Although subjects had them available in that test also, their range of variation was much more restricted.

[7]The author has not yet come across any relevant recordings or literature and would welcome pertinent information, also about any other literature on clapping that may exist.