Subjective preference in relation to objective parameters of music sound fields with a single echo[a)]

Y. Ando[b)]

Drittes Physikalisches Institut, Universität Göttingen, D—3400 Göttingen, Federal Republic of Germany

(Received 13 January 1977; revised 29 July 1977)

Subjective preference tests with a simulated single reflection in an anechoic chamber were conducted in order to get a knowledge of the preferred properties of sound fields. The degree of preference in relation to a long-time autocorrelation function of source signal and the interaural cross correlation of the sound field is discussed here. The preferred time delay gap between the direct sound and the first echo can be determined by the coherence of autocorrelation function and the amplitude of the echo. The preferred echo direction to a listener may be obtained by minimizing the interaural cross correlation which relates to the "subjective diffuseness." Results of the autocorrelation function, which was measured with an A-weighting filter, and the interaural cross correlation are presented for several music motifs.

PACS numbers: 43.55.Br, 43.75.-z, 43.60.Cg

INTRODUCTION

In 1950, Haas investigated the disturbance of a single echo on the subjective hearing of continuous speech.[1] Subsequently, many investigators performed such subjective works on the question of how the echo may be masked by the direct sound.[2–4] These investigators adopted, as a necessary condition for the sound field, whether or not the echo causes significant disturbance.

On the other hand, Damaske suggested that the interaural coherence is a significant factor in determining the perceived direction of a sound source, or the degree of subjective diffuseness of the sound field.[5,6] The method of evaluating the interaural cross correlation (IACC) was previously used in order to identify the desirable loudspeaker directions minimizing the degree of IACC.[7] This is applied here for sound fields with discrete echoes. Also, Keet suggested that the degree of the spatial effect was related to the cross correlation between signals at the two ears and the sound level.[8] The need of early lateral reflections was proposed by Barron for the desirable spatial effect.[9]

In order to get a knowledge of the preferred properties of sound fields, subjective preference tests with a simulated single reflection in an anechoic chamber were conducted. The degree of preference in relation to a long-time autocorrelation function (ACF) of source signal and the IACC was recorded.

I. SIGNALS AT THE TWO EARS

A. Analysis

Let $h_l(t; r|r_0)$ and $h_r(t; r|r_0)$ be pressure impulse responses between a source located at r_0 in a room and the left and the right eardrums of a listener sitting at r, respectively. Then, the pressures to the two eardrums, which must include all acoustic information to be analyzed, are expressed by the following equation:

$$f_l(t) = p(t) * h_l(t;\ r|r_0),$$
$$f_r(t) = p(t) * h_r(t;\ r|r_0), \tag{1}$$

where $p(t)$ is a source signal and the symbol $*$ indicates the convolution. The impulse responses may be divided into a set consisting of the reflecting impulse response of boundaries $w_n(t)$—see Refs. 10 and 11—and the impulse response from the free field to the eardrum $h_{nl}(t)$ or $h_{nr}(t)$—see Ref. 17. Equation (1) may now be written in the forms of

$$f_l(t) = \sum_n p(t) * A_n w_n(t - \Delta t_n) * h_{nl}(t),$$
$$f_r(t) = \sum_n p(t) * A_n w_n(t - \Delta t_n) * h_{nr}(t), \tag{2}$$

where A_n and Δt_n are the amplitude (pressure) and the delay time of the nth echo relative to the direct sound, Δt_0 being zero. A_n is determined by the $1/r$-law, A_0 being unity. Every n corresponds to a single reflection with the horizontal angle ξ and the elevation angle η.

All independent objective parameters included in Eq. (2) may be reduced to the following:

(i) the source signal $p(t)$, in terms of the long-time ACF defined by

$$\Phi_p(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} p'(t)p'(t+\tau)dt, \tag{3}$$

where $p'(t) = p(t) * h(t)$, $h(t)$ may be defined as a filter of the ear expressed in the time domain. For the sake of convenience, $h(t)$ is chosen by an A-weighting filter corresponding to the ear sensitivity (monaural criterion);

(ii) the set of the reflecting impulse responses of the boundaries which determines the initial time delay gap between the direct sound and the first reflection, the initial amplitude pattern of echoes, the reverberation time, etc. (monaural criterion);

(iii) the two sets of the head related impulse responses for the two ears $h_{nl}(t)$ and $h_{nr}(t)$ (binaural criterion). Concerning (iii), we introduce the long-time IACC[7] between the signals of $f_l(t)$ and $f_r(t)$. It is a convenient measure of the subjective diffuseness. The IACC is defined by

$$\Phi_{lr}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} f_l(t) f_r(t+\tau)dt,$$
$$|\tau| < 1\ \text{ms}. \tag{4}$$

TABLE I. Music motifs used[14] and the durations of the long-time autocorrelation function ($2T = 35$ s).

Motif	Title of piece	Composer	τ_e (ms)
A	Royal Pavane	Gibbons	127[a]
B	Sinfonietta, Opus 48; III movement Allegro con brio	Malcolm Arnold	35
C	Symphony No. 102 in B flat major; II Movement: Adagio	Haydn	65[a]
D	Siegfried Idyll; Bar 322	Wagner	40

[a]Estimated.

First, let us consider the IACC $\Phi_{lr}^{(0)}(\tau)$ of the direct sound only. The normalized IACC is given by

$$\varphi_{lr}^{(0)}(\tau) = \frac{\Phi_{lr}^{(0)}(\tau)}{[\Phi_{ll}^{(0)}(0)\Phi_{rr}^{(0)}(0)]^{1/2}} , \tag{5}$$

where $\Phi_{ll}^{(0)}(0)$ and $\Phi_{rr}^{(0)}(0)$ are autocorrelation functions (at $\tau = 0$) of the direct sound at the eardrums. If discrete echoes are added to the direct sound after the ACF of the direct sound becomes weak enough, then the normalized IACC is expressed by

$$\varphi_{lr}^{(n)}(\tau) = \frac{\sum_n A_n^2 \Phi_{lr}^{(n)}(\tau)}{[\sum_n A_n^2 \Phi_{ll}^{(n)}(0) \sum_n A_n^2 \Phi_{rr}^{(n)}(0)]^{1/2}} , \tag{6}$$

where $\Phi_{lr}^{(n)}(\tau)$ is the IACC of the nth echo, and $\Phi_{ll}^{(n)}(0)$ and $\Phi_{rr}^{(n)}(0)$ are autocorrelation functions of the nth echo at the eardrums.

B. Measurement of correlation functions

1. The long-time autocorrelation function (ACF)

Jansson and Sunberg, who measured spectra of signals for 2 to 100 s duration, recommended that the piece of music to be analyzed should be not shorter than 20 s in practice.[13] In this measurement, the long-time ACF of 35 s duration ($2T = 35$ s) was measured with the A-weighting filter. Music motifs measured are shown in Table I. Motif A was recorded by the Philip Jones Brass Ensemble and motif B, C, and D were recorded by the English Chamber Orchestra in the Building Research Station Anechoic Chamber.[14] Two examples of the normalized ACF given by $\varphi_p(\tau) = \Phi_p(\tau)/\Phi_p(0)$ are shown in Fig. 1. An effective duration of ACF is defined by the delay τ_e such that the envelope of the normalized ACF becomes smaller than 0.1 after the delay. These values for each motif are also indicated in Table I. As a result of using the A-weighting filter, it appears that they are greatly different from those measured by Fourdouiev.[15] Extreme results are seen for motif A by Gibbons ($\tau_e = 127$ ms) and motif B by Arnold ($\tau_e = 35$ ms).

2. The interaural cross correlation (IACC)

The long-time IACC ($2T = 35$ s) was measured at the eardrums of a dummy head with ear canals and built-in microphone.[16] It was made after an accurate measurement of threshold level so that the output signal of the microphones corresponded to the ear sensitivity. Measured values of the IACC as a function of delay are shown in Fig. 2 for different horizontal angle of a single sound. The geometrical arrangement is shown in the upper part of Fig. 2(b). The delay range $|\tau| > 1$ ms was not measured, because the possible interaural time delay is about 0.75 ms. The particular values of the IACC at $\tau = 0$ are indicated in Table II for each motif. The similar behavior of the IACC is commonly found for the all signals, for example, values of $\Phi_{lr}(0)$ show the maxima at the frontal incidence $\xi = 0°$ and the minima at $\xi = 54°$. Also, the large differences are obtained between the values of $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ at $\xi = 54°$, because of the level differences for the two ears. Strong periodicities corresponding to the spectrum of signals are observed at 1.3 and 3.3 kHz in Figs. 2(a) and 2(b), respectively. Comparing the two figures, however, it is seen that the maximum for a given ξ occurs at the same interaural time delay, as might be expected. Using these values, the normalized IACC of sound fields with the discrete echoes can be calculated by Eq. (6).

The calculated maximum absolute values of the normalized IACC for the delay range of $|\tau| < 1$ ms of the fields with the single echo are shown in Fig. 3 as a function of the horizontal direction of echo. A similar result is obtained when values of IACC are plotted for $\tau = 0$, as shown in Fig. 4. The behavior of the IACC as a function of the echo direction is similar for the all motifs, as well as the noise source.[7,11] Thus the magnitude of IACC is almost independent of the source signal. Inspection of Figs. 3 and 4 indicates that the IACC drops rapidly with increasing ξ to a minimum for dis-

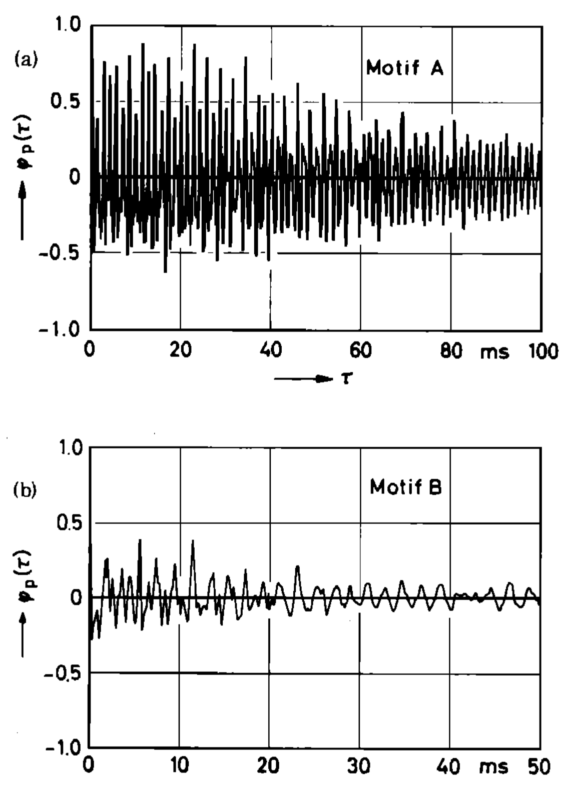

FIG. 1. Examples of the measured ACF. (a) Music motif A, (b) music motif B.

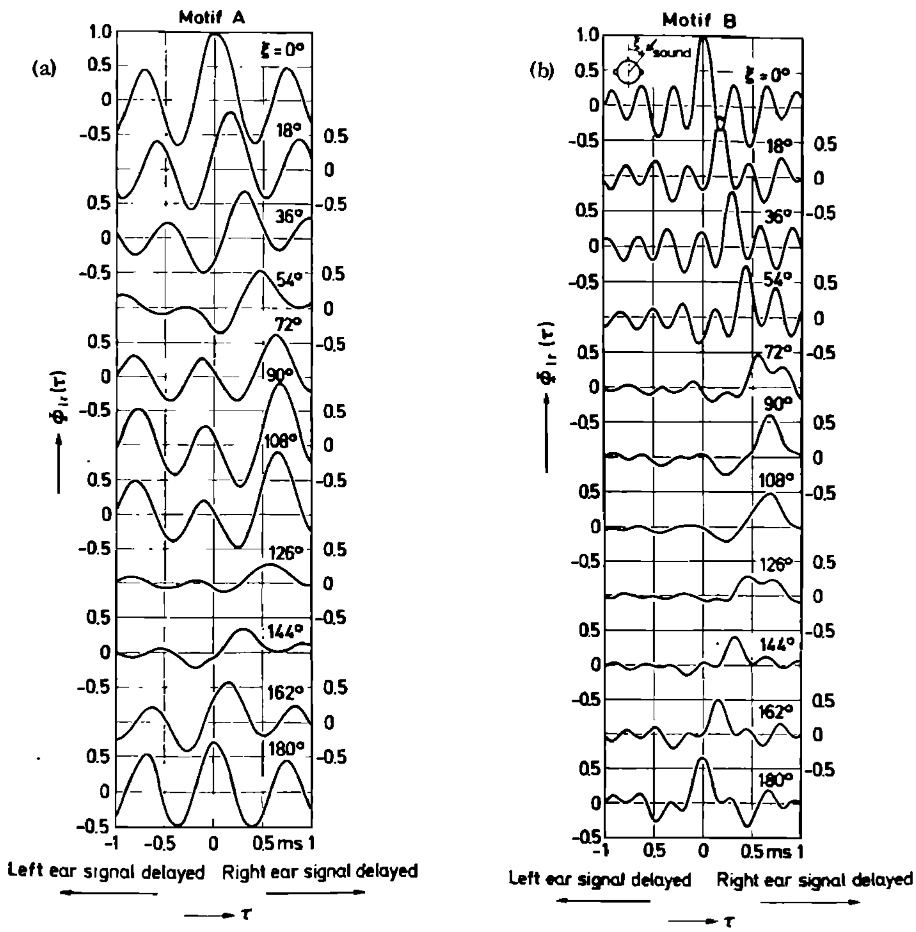

FIG. 2. Examples of the measured IACC for the different horizontal angle of incidence. (a) Music motif A, (b) music motif B.

crete echoes in the range $\xi = 15°$ to $60°$ and leveling off for higher angles. It will be shown later that this angular range is of special importance subjectively also. Moreover, the IACC decreases with increasing echo amplitude particularly for echoes from this region. On the contrary, echoes coming from the median plane $\xi = 0°$ keep it large.

To determine the minimum delay time required for Eq. (6) to hold, measured maximum values of the IACC as a function of delay time Δt_1 are plotted in Fig. 5. The calculated values are indicated at $\Delta t_1 \rightarrow \infty$. They are in good agreement with the measured values except for $\Delta t_1 < 4$ ms (motif B, C, and D) and $\Delta t_1 \leqq 15$ ms (motif A). When the echo delay approaches zero, the measured values are rapidly increased. (Therefore, the subjective diffuseness may be expected when the echo arrives after about 10 ms. [9])

TABLE II. Measured correlation functions at $\tau = 0$ for each music motif as a function of the horizontal angle of incidence $\xi(\eta = 0°)$. For the range $180° < \xi < 360°$, the values are obtained by putting $\xi = 360° - \xi$ and interchanging the suffixes 1 and r.

		The horizontal angle ξ										
Motif		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°
	$\Phi_{1r}(0)$	0.99	0.30	−0.32	−0.32	0.09	0.13	0.00	−0.07	−0.09	0.30	0.69
A	$\Phi_{11}(0)$	1.00	0.71	0.42	0.32	0.34	0.65	0.62	0.19	0.24	0.52	0.69
	$\Phi_{rr}(0)$	1.00	1.12	1.31	1.42	1.27	1.51	1.51	0.75	0.84	0.75	0.71
	$\Phi_{1r}(0)$	0.99	−0.17	0.18	−0.28	−0.04	0.04	0.00	−0.06	0.03	0.00	0.63
B	$\Phi_{11}(0)$	1.00	0.54	0.39	0.35	0.28	0.34	0.30	0.23	0.27	0.39	0.66
	$\Phi_{rr}(0)$	1.00	1.38	1.73	2.06	1.42	1.25	1.13	0.87	0.92	0.75	0.66
	$\Phi_{1r}(0)$	0.97	0.18	−0.06	−0.28	−0.16	−0.14	−0.16	−0.12	−0.04	0.22	0.61
C	$\Phi_{11}(0)$	1.00	0.66	0.42	0.40	0.42	0.60	0.57	0.26	0.30	0.45	0.72
	$\Phi_{rr}(0)$	1.00	1.17	1.34	1.50	1.39	1.39	1.30	0.87	0.85	0.74	0.58
	$\Phi_{1r}(0)$	1.00	−0.11	0.02	−0.23	0.06	0.15	0.10	−0.08	−0.06	−0.06	0.88
D	$\Phi_{11}(0)$	1.00	0.65	0.46	0.20	0.28	0.36	0.28	0.22	0.23	0.44	0.90
	$\Phi_{rr}(0)$	1.00	1.40	1.66	1.66	1.56	1.38	1.14	0.97	0.87	0.80	0.96

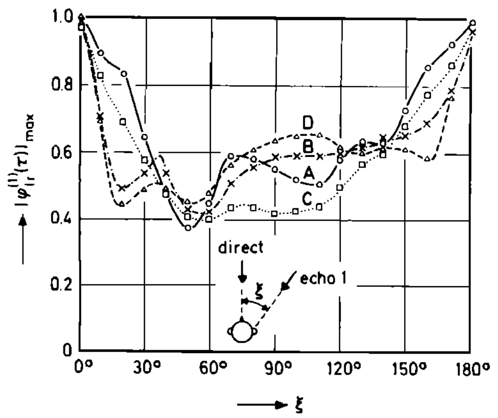

FIG. 3. Calculated maximum absolute values of the IACC of the sound fields as a function of the horizontal angle of the single echo.

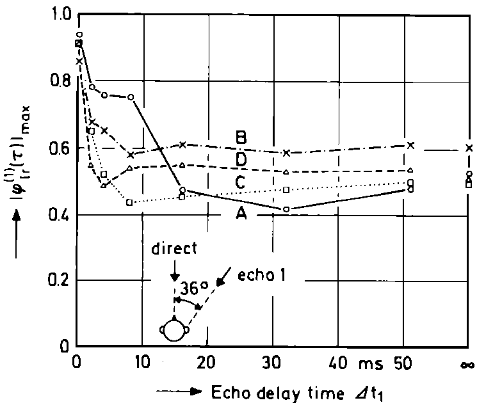

FIG. 5. Measured maximum absolute values of the IACC of the sound fields as a function of delay of the single echo at $\xi = 36°$. Values at $\Delta t_1 \to \infty$ are calculated by Eq. (6).

II. SUBJECTIVE PREFERENCE TESTS

A. Procedure

The tests were performed in order to get the degree of preference in relation to the coherence of ACF of the source signal and the maximum value of the IACC.

The sound pressure level of the direct sound ($\xi = 0°$, $\eta = 9°$) presented to a subject was adjusted to a peak value of 80 dB A. The reflecting impulse response of boundary was chosen as $w(t - \Delta t_1) = \delta(t - \Delta t_1)$.[11] The paired-comparison tests were conducted with 13 subjects judging which of the sound fields they preferred to hear. The first 12 s of motifs A and B were chosen as source signals.

Preference scores were obtained by giving scores +1 and −1 according to positive and negative judgments, respectively.[17] For no preference judgment between the sound fields, they were given zero. The normalized score may be obtained by accumulating the scores for all sound fields (F) and all number of subjects (S), and then dividing the factor of $S(F-1)$. The positive scores, for example, indicate what percent remains of the positive judgment to the sound field, after subtracting that of the negative.

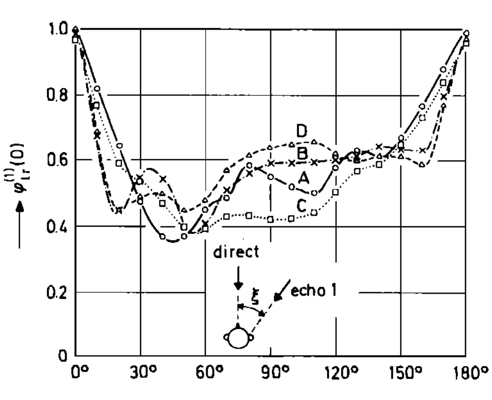

FIG. 4. Calculated values of the IACC at $\tau = 0$ of the sound fields as a function of the horizontal angle of the single echo.

B. Results and discussion

1. Preferred delay of the single echo

The sound with the single echo from a fixed direction of $\xi = 36°$ was presented, this value being selected as a typical reflection angle. In the next section, it will be noted that the preferred angle of reflection is somewhat larger, but the relation between preferred time delay and preferred angle is not thought to be critical. The delay time was adjusted in the range of 6–256 ms. The normalized scores of the sound field are shown in Fig. 6 as a function of the delay. The most preferred delay time showing the maximum score differs greatly in motifs. If the amplitude of echo $A_1 = 1$, the most preferred delays are around 128 and 32 ms for motifs A and B, respectively.

After an iteration, the desired delay is found roughly at a certain duration of ACF which is defined by τ_e such that the envelope of ACF becomes $0.1A_1$. The data collected are shown in Fig. 7 as a function of the duration

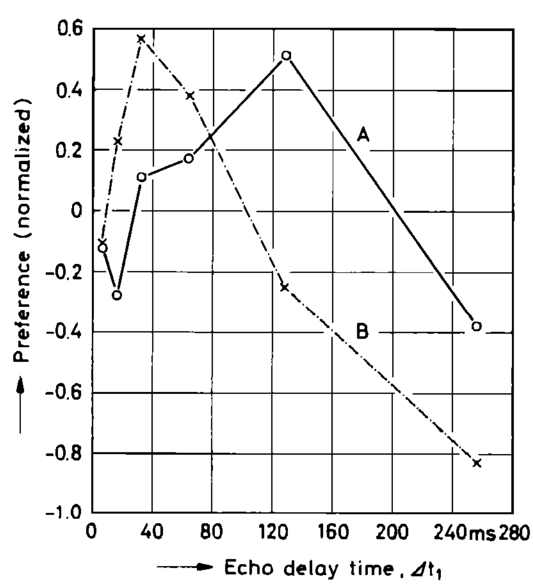

FIG. 6. Preference scores of the sound fields as a function of the delay of the single echo at $\xi = 36°$, $A_1 = 0$ dB (13 subjects).

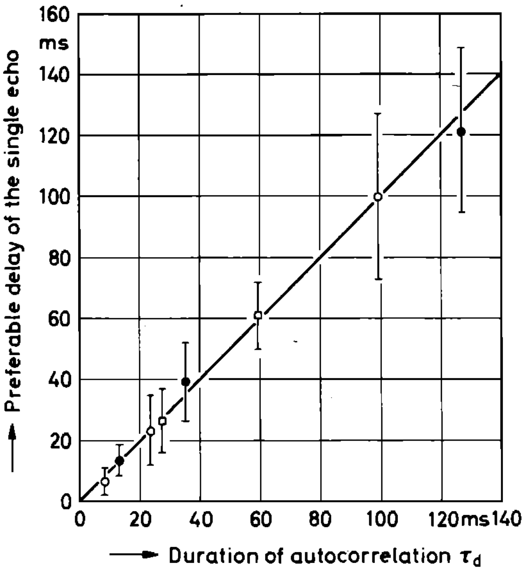

FIG. 7. Relationship between the preferable delay of the single echo and the duration of ACF such that $|\varphi_p(\tau)| = 0.1A_1$. Ranges of preferable delay are graphically obtained at a 0.1 below the maximum score. Different symbols are the center values obtained at the echo levels of +6 dB (o), 0 dB (●) and −6 dB (□), respectively (7–13 subjects).

τ_d. Thus, $\tau_d = \tau_e$ only when $A_1 = 1$. This means that the pressure amplitude of the echo is equal to ten times (not always 10 dB) greater than the amplitude of ACF at τ_d. The data with a continuous speech signal of $\tau_e = 12$ ms[11] are also plotted in the figure.

Two reasons can be considered why the preference decreases for $0 < \Delta t_1 < \tau_d$:

(1) Tone coloration effects occur, because of the interference phenomenon in the coherent time region;

(2) the IACC increases as shown in Fig. 5 for $\Delta t_1 \to 0$. On the other hand, echo disturbance effects can be observed, if Δt_1 is longer than τ_d.

2. Preferred echo direction

The delay time of the echo in this experiment was fixed at 32 ms. The echo was presented by loudspeakers located at $\xi = 18°(\eta = 9°)$, $36°, \ldots, 90°$ and $\xi = 0°(\eta = 27°)$.

Results of the preference tests are shown in Fig. 8 as a function of the echo direction ξ. The maximum values of the IACC are also plotted for comparison. No fundamental differences are observed between the curves of the motifs in spite of the great differences of source signal. The preference score increases roughly by decreasing the degree of IACC. The correlation coefficient between the score and the IACC is −0.76 (at 1% significance level). The score with motif A at $\xi = 90°$ is dropped into a negative value, indicating that the lateral reflections coming from around 90° are not always preferred. The figure shows that there is a preference for angles greater than $\xi = 30°$. On the average there may be an optimum range centered on 55°. Similar results can be seen in the data with the speech signal.[11]

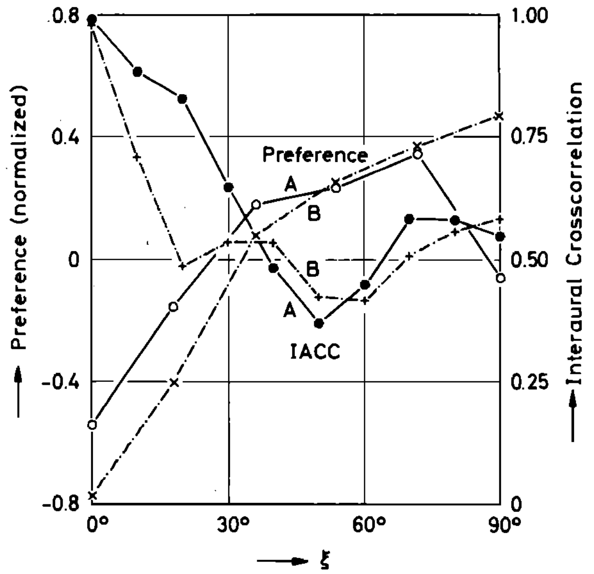

FIG. 8. Preference scores of the sound fields and the degrees of IACC as a function of the horizontal angle of the single echo, $A_1 = 0$ dB (13 subjects).

III. CONCLUSIONS

The coherence of the ACF measured with the A-weighting filter varies greatly with sources. However, the magnitude of IACC of the sound fields is almost independent of the sources. The IACC drops rapidly with increasing ξ to a minimum for discrete echoes in the range of $\xi = 15°$ to 60° and leveling off for higher angles.

The result of subjective preference tests indicates that the preferred delay of the single echo can be determined by the coherence of ACF and the amplitude of the echo. Also, the preference score of the sound field increases by decreasing the degree of IACC, so that the preferred echo directions may be found in the range centered on 55°.

These findings can be expanded to sound fields with multiple echoes.[18]

ACKNOWLEDGMENTS

The author wishes to thank Prof. Dr. M. R. Schroeder and Dr. D. Gottlob for their valuable discussions during this project. This work was supported by the Alexander von Humboldt Foundation, Bonn.

[a]A part of this paper was presented at the "DAGA '76" in Heidelberg.

[b]Present address: Faculty of Engineering, Kobe University, Kobe 657, Japan.

[1]H. Haas, "Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache," Acustica 1, 49–58 (1951).
[2]E. Meyer and G. R. Schodder, "Über den Einfluss von Schallrückwürfen auf Richtungslokalisation und Lautstärke bei Sprache," Nachr. Akad. Wiss. Göttingen, Math-Phys. Chem. Abstract No. 6, 31–42 (1952).
[3]R. W. Muncey, A. F. B. Nickson, and P. Dubout, "The Acceptability of Speech and Music with a Single Artificial Echo," Acustica 3, 168–173 (1953).
[4]J. P. A. Lochner and J. F. Burger, "The Subjective Masking of Short Time Delayed Echoes by Their Primary Sounds and

Their Contribution to the Intelligibility of Speech," Acustica **8**, 1–10 (1958).

[5]P. Damaske, "Subjektive Untersuchung von Schallfeldern," Acustica **19**, 199–213 (1967/68).

[6]P. Damaske, "Richtungsabhängigkeit von Spektrum und Korrelationsfunktionen der an den Ohren empfangenen Signale," Acustica **22**, 191–204 (1969/70).

[7]P. Damaske and Y. Ando, "Interaural Crosscorrelation for Multichannel Loudspeaker Reproduction," Acustica **27**, 232–238 (1972).

[8]W. V. Keet, "The Influence of Early Lateral Reflections on the Spatial Impression," Proc. 6th ICA Tokyo, paper E-2-4 (1968).

[9]M. Barron, "The Subjective Effects of First Reflections in Concert Halls—The Need for Lateral Reflections," J. Sound Vib. **15**, 475–494 (1971).

[10]Y. Ando and K. Kato, "Calculations on the Sound Reflection from Periodically Uneven Surfaces of Arbitrary Profile," Acustica **35**, 321–329 (1976).

[11]Y. Ando and K. Kageyama, "Subjective Preference of Sound with a Single Early Reflection," Acustica **37**, 111–117 (1977).

[12]S. Mehrgardt and V. Mellert, "Transformation Characteristics of the External Human Ear," J. Acoust. Soc. Am. **61**, 1567–1576 (1977).

[13]E. V. Jansson and J. Sunberg, "Long-Time-Average-Spectra Applied to Analysis of Music. Part I: Method and General Applications," Acustica **34**, 15–19 (1975).

[14]A. N. Burd, "Nachhallfreir Musik für akustische Modelluntersuchungen," Rundfunktechn. Mitteilungen **13**, 200–201 (1969).

[15]V. Fourdouiev, "Evaluation Objective de l'Acoustique des Salles," Proc. 5th ICA Liège, 41–54 (1965).

[16]V. Mellert, "Construction of a Dummy Head after New Measurements of Thresholds of Hearing," J. Acoust. Soc. Am. **51**, 1359–1361 (L) (1972).

[17]M. R. Schroeder, D. Gottlob, and K. F. Siebrasse, "Comparative Study of European Concert Halls: Correlation of Subjective Preference with Geometric and Acoustic Parameters," J. Acoust. Soc. Am. **56**, 1195–1201 (1974).

[18]Y. Ando, U. Eysholdt, and D. Gottlob, "Subjective Preference Tests of Sound Fields with Multi-Echoes," 9th ICA Madrid, paper B8 (1977).