The Influence of a Single Echo on the Audibility of Speech*

HELMUT HAAS

A thorough investigation was made of the influence of a single echo as a function of different parameters on the audibility of speech. An apparatus was built for the artificial generation of echoes and measurements were made with a large number of observers under specified conditions.

Editor's Note: Through the efforts of Edward M. Long, whose paper appears earlier in this issue, we are pleased indeed to publish the English translation of Dr. Haas' now-famous paper. To the best of our knowledge, this dissertation has never been published in the United States, and we are certain that many working in the field would benefit from reading the "Haas Effect."

INTRODUCTION: Sounds produced in rooms suffer numerous reflections from boundary surfaces and furniture. A reflected sound reaches the observer later than the direct sound, and major path differences or delay time (transit time) differences may impair or even spoil the acoustical effect. A sufficiently long time difference causes the direct sound and the reflected sound to be heard separately, that is to say, one hears an echo in the usual sense of the word. In the present paper the term "echo" will cover all reflections of sounds, independent of the magnitude of the delay time.

The threshold value of the delay time, above which a noticeable deterioration of the acoustical impression occurs, has been called "threshold of masking" by Petzold [1]. He gave as the value for speech.

$$t = 0.05 \pm 0.01 \text{ second}$$

corresponding to a path difference in air $s = 17 \pm 3$ m.

In rather small rooms such a difference would occur only after many reflections. No disturbance due to echo is to be expected because the intensity of the often reflected sound is very much reduced after 50 ms. Besides, the ear receives in quick succession a great number of reflections, decreasing in intensity, and the hearer is not capable of perceiving them separately. This gradual decay of sounds in rooms is called reverberation. A certain reverberation time is even advantageous for audibility, since it entails an intensification of the sound as well as a pleasant modification of the character of the tone.

The critical delay time can easily be exceeded in large rooms, after only one reflection, or after a few reflections. A bad effect upon the representation is likely to result if the energy of the reflected sound is sufficiently high to accentuate certain points in the reverberation curve of a gradually decaying sound. In the extreme case a real echo will be heard.

The possibility of estimating in advance the extent of likely disturbances due to echoes is of paramount importance in the design and construction of buildings. Irritating echo phenomena caused by sufficiently great time differences can also occur in connection with the arti-

* This paper was submitted as a dissertation toward the doctor's degree to the University of Gottingen, Gottingen, Germany, under the title, "Über den Einfluss des Einfachechos auf die Horsamkeit von Sprache." The present translation from the German by Dr. Ing. K.P.R. Ehrenberg is reprinted with permission from the Building Research Station, Watford, Herts., England, Library Communication no. 363, December 1949.

ficial amplification of sounds indoors, or with the transmission of sounds by means of several loudspeakers placed at distances from each other. Lastly, echo phenomena occur with disturbing effects under certain conditions in telecommunications.

The object of the following investigation is to determine the influence of echoes on the audibility of speech in binaural hearing, as a function of various parameters, viz., delay time, intensity, timbre, and direction. The investigation has been confined to single echoes, i.e., only one repetition following the direct sound, in order to obtain simpler conditions, and to limit the experimental requirements. Measurements have been made as far as the results are of practical significance. No previous investigations on such a scale with a large number of observers taking part have been made in this field. Decker [2] studied echo phenomena occurring in telephone connections over long cables. When the intensity of the echo was equal to that of the primary sound, 50% of 50 pairs of observers felt disturbed in the flow of their conversation when the delay time difference was 100 ms. The critical difference rose to 150 ms when the echo amplitude was only 60% of the amplitude of the primary sound.

Stumpp [3] studied in the open air the influence of different directions of the incidence of a single echo equal in intensity to the primary sound. The critical delay time difference was 80 ms for speech and echo coming from the same lateral direction, but 50 ms for opposite lateral directions.

DESCRIPTION OF EQUIPMENT FOR ARTIFICIAL GENERATION OF ECHOES

An accurate control of the properties of reflected sound necessitates a device for the artificial generation of echoes, capable of repeating a sound phenomenon after a fixed time. This problem can be tackled in different ways.

1) Delay by electrical network. This solution is excluded because of the order of magnitude of the time differences required, and because of the wide frequency range involved.

2) Delay by using the finite velocity of sound in a medium, e.g., air in a long tube. Such an apparatus would be very cumbersome, varying the delay time by small steps would be difficult, but, most important, already an upper limiting frequency of 5,000 Hz and a delay time of 200 ms (length of the tube conduit 68 m) would cause very high damping [4], which hardly could be compensated without special expenditure.

3) Delay by using two pickups from a single record, or by staggered pickup from two identical records. At first sight this simple solution appeared appropriate, but further examination proved the apparatus to be rather complex, if convenient and exact working was to be achieved, even apart from the poor frequency range of the phonograph method.

4) Delay by the magnetophone method. Although this is expensive, it has been chosen because it best fulfills the requirements. The idea is that the sound process to be delayed is magnetically recorded on a tape, from which the direct sound and the echo can be picked up with any delay needed in practice.

Fig. 1 shows the block diagram of this installation. The microphone M transforms the sound, in the following investigation chiefly speech, into voltage modulations, to be amplified and conducted to the recording channels of the magnetophone apparatus. If the same text is to be used repeatedly for certain measurements, it can be recorded and played back by means of a magnetic tape apparatus. The changeover switch after the microphone amplifier is used to select direct or indirect reproduction. The level is controlled, and kept constant by means of a level monitor.

The actual delay of sound is produced by an endless-tape magnetophone apparatus, sketched on top of Fig. 1, and registered by the high-frequency method.

Two pivoted guide rollers R, 150 mm in diameter, carrying a loop of normal magnetic tape, are fixed at a distance of 1 m. A rubber roller presses the tape against the shaft of a synchronous motor (n = 1500 r/min),

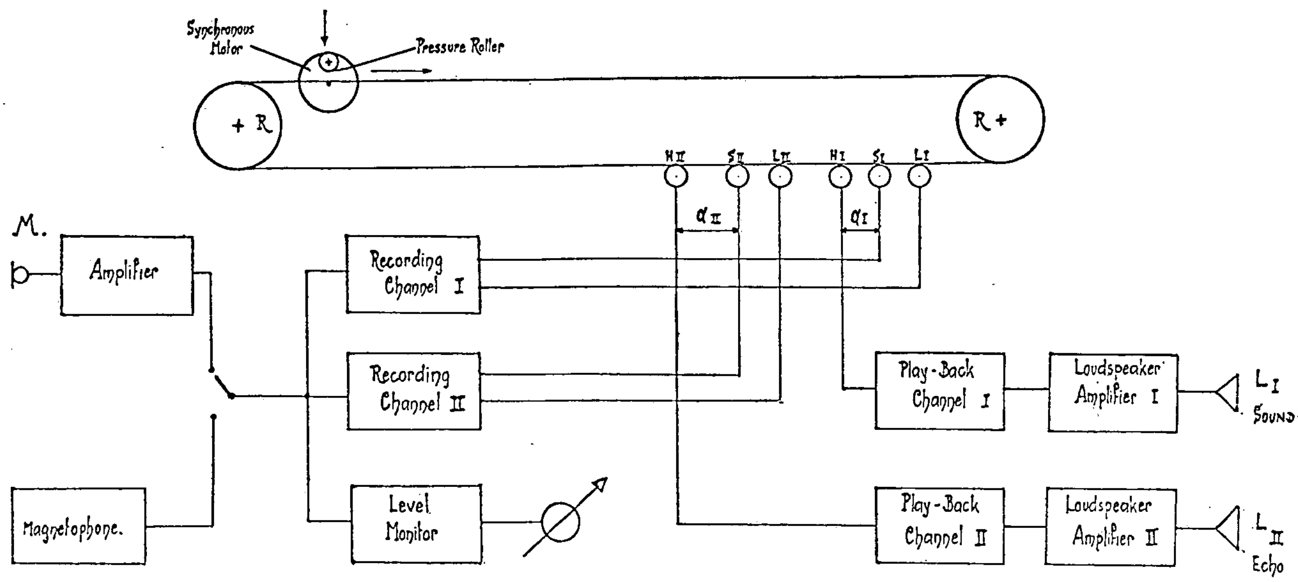

Fig. 1. Block diagram of echo apparatus.

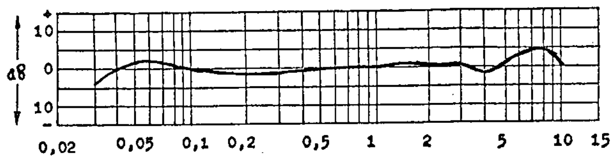

Fig. 2. Frequency response curve of microphone (according to manufacturer).

which drives it at a speed of 77 cm/s in the direction of the arrow. Circles indicate the wiping heads L_I and L_{II}, playback heads H_I and H_{II}, and the recording heads S_I and S_{II}, all fixed on an optical bench, H_{II} being movable along the line of the tape.

The equipment operates in the following manner. L_I and L_{II}, fed with high frequency (about 80000 Hz) from the recording heads, wipe out the magnetic records on the tape. S_I and S_{II} are excited from the recording channels I and II, which on the input side are connected in parallel. This signal is then picked up by the playback heads H_I and H_{II}. The delay difference Δt is determined by

$$\Delta t \approx (a_{II} - a_I)/v$$

where a_I denotes the distance between the gaps of S_I–H_I, a_{II} the corresponding distance S_{II}–H_{II}, and v the speed of the tape. The transit difference is $\Delta t = 0$ for $a_I = a_{II}$, i.e., the sound phenomena are picked up simultaneously as they were recorded. The echo delays are adjusted by moving H_{II} along a scale calibrated in milliseconds.

It was originally thought that an echo could be obtained from the time difference between the direct sound and a simple recording via the playback head, but it would not then have been possible to reduce a_I, and with it the delay difference, to a very small value. Double recording was therefore necessary.

The levels induced by the moving magnetic tape at H_I and H_{II} are fed to two separate playback channels I and II and amplified. The loudspeaker L_I emits the original sound, and L_{II} the same sound delayed in time, i.e., the echo.

The first sound phenomenon, hereafter called "direct sound" or "primary sound," is also produced from a loudspeaker. This appeared expedient in order to prevent acoustical feedback and to avoid any disturbing effect of the echo on the speaker.

Special consideration was given to appropriate dimensions with a view to excluding as far as possible inaccurate results due to the reproducing apparatus.

The recording microphone was a dynamic one, with a frequency response curve, according to information obtained from the manufacturer, as shown in Fig. 2 (output level across 200 ohms for constant sound pressure at the microphone as a function of frequency).

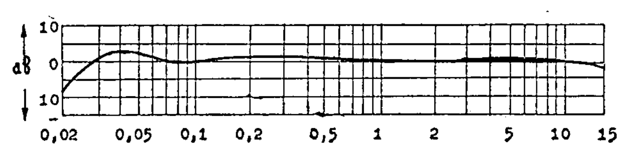

Fig. 3. Frequency response curve of magnetophone apparatus.

The reproduction is most favorable for high frequencies above 5000 Hz, provided that the direction of the speech is suitable.

The frequency response characteristic of the microphone amplifier is perfectly flat over the relevant range. The magnetic recording and reproducing equipment is also flat. The overall frequency response characteristics of both magnetophone channels are practically identical; they are given in Fig. 3 for the use of a *c* tape. The output level of the playback head for constant level of the recording channel is measured as a function of frequency. Curve I on Fig. 4 is the frequency response curve of the loudspeaker amplifier, the output level being plotted for loading and a constant input level. The resistance frequency characteristic I of this amplifier can be altered in steps according to curves II and III. Again care was taken to obtain identical frequency response curves for both amplifiers.

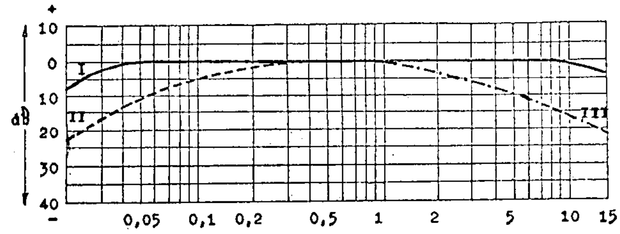

Fig. 4. Frequency response curve of loudspeaker amplifier.

The weakest link in the system is the loudspeaker. Numerous types, made by different manufacturers, were tested. Finally a dynamic 12.5-W loudspeaker (30-cm diameter of diaphragm) with a permanent magnet was chosen, and the two specimens selected were as near to uniformity as possible. These loudspeakers were placed in small, square baffle boards (40 by 40 cm), supported by easily transportable racks so that they were at the height of the observer's head. The baffle boards were used to prevent reflections from large walls when the loudspeakers stood opposite each other. The large diameter of the diaphragm was chosen in order to obtain sufficient emission of low frequencies, even with small baffle boards. A special system for the emission of high frequencies was not used, since the expenditure was not thought to be justified. Such piezoelectric loudspeakers as were tested showed a great number of very pronounced resonance points within the important range. The emission of high frequencies from small dynamic loudspeakers was very little superior to that of the large type actually used.

Fig. 5 shows the sound pressure as a function of frequency for one of the loudspeakers employed for the measurements described hereafter, at a distance of 2 m on the axis and in the 40- by 40-cm baffle board. The loudspeaker was driven from the loudspeaker amplifier, the input terminals of which were fed with an ac voltage of constant amplitude.

As no damped sound room was available, the measurements were made in the open air, on the roof of the Institute. Interference due to reflections from the floor was eliminated by placing the baffle board immediately upon the roof, viz., so that the axis of the loudspeaker

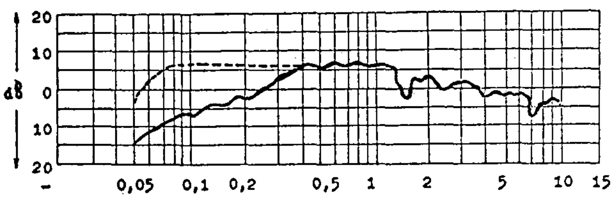

Fig. 5. Frequency response curve of loudspeaker (sound pressure on median line).

was focused upon the measuring microphone which was close to the floor. The sound pressure was measured by a calibrated spherical condenser microphone, 3.5 m in diameter, with its sound pressure response compensated by electrical filters. The sound pressure as a function of frequency was taken with a Neumann level recorder.

The dashed curve in Fig. 5 indicates the sound pressure curve for the same loudspeaker when used in a baffle board of infinite size (the wall of a house). The distance to the microphone was again 2 m in the axis.

The overall frequency response characteristic of the echo system, combined from Figs. 2–5, is shown in Fig. 6. The sound pressure at a distance of 2 m from the playback loudspeaker in the direction of its axis is plotted against frequency for constant sound pressure at the receiving microphone. It is seen that the greatest deviations within the frequency range from 100 to 10 000 Hz amount to ± 7 dB.

No nonlinear distortions could be detected in the oscillogram of a pure tone of 1000 Hz. Since the loudspeaker amplifier and the loudspeaker have been dimensioned with a sufficient margin, it follows that the magnitude of the distortion factor is determined essentially by the magnetophone equipment. The value is given in the literature as < 3% at 800 Hz.

The results obtained from objective measurements made it appear that the installation was quite suitable for the transmission of natural speech. This has been confirmed by subjective measurements. A trained team reached 96% intelligibility for syllables across the whole installation. This value is also the maximum for direct transmission between mouth and ear. The remaining 4% are explained by occasional inattention of observers or poor articulation by the speaker.

The slight linear distortions of the system can affect the timbre of the speaker, but they do not reduce the intelligibility and are therefore not detrimental.

EXPERIMENTS WITH SMALL DIFFERENCES

Auditory performances in smaller rooms are always connected with echoes following the direct sound after a brief interval. Let us imagine a conversation in a normal living room where the smooth surface of the ceiling will reflect sound. For simplification, the walls and the floor can be assumed to be more or less sound absorbing. The regular decay of the reverberation will then be replaced by individual echoes. Such echoes arriving after a short time are not disturbing, in spite of the fact that the energy of the reflected sound is not appreciably lower than that of the direct sound. On the contrary, conversations without such small reflections are felt to be more tiring and less natural; this can be observed, for instance, in a damped room, or outdoors in the presence of a deep cover of fresh snow.

This experience shows that our organ of hearing accepts as an entirely natural process the integration of the direct sound impressions together with their immediately following reflections. The question arises then, how the hearing of a sound without reflections differs from one with slightly delayed echoes. The following experimental arrangement has been devised for the solution of this problem.

In the absence of a dead room the investigation had to be made outdoors, and the flat roof of the Institute, a detached building, proved suitable. Some projecting parts were covered with inclined timber walls to deflect the incident sound upwards.

The positions for loudspeaker and observer were chosen with a view to reducing to a minimum echoes from an annex about 20 m distant. Points were determined by giving short sound impulses from a loudspeaker, and receiving the reflections in a microphone at the observer's eventual stand, with oscillographical registration.

The influence of reflections from the surface of the roof was eliminated by placing the loudspeaker, focused on the observer's head, immediately upon the roof in a 40- by 40-cm baffle board. However, this reflection from the ground was found to have no effect on the results of the measurements, so that afterwards the loudspeakers were always placed at the height of the observer's head, a procedure closer to practical conditions.

Two loudspeakers of the same type were placed 3 m distant from the observer at an angle of 45°, half to the right and half to the left of him. The only reason for this particular arrangement was that it proved to be the best for balancing. The loudspeaker axes were focused on the observer in order to prevent a fall in the high frequencies owing to the directional characteristics of the loudspeakers. Both loudspeakers were then at the same intensity; the loudness at the observer's position was about 50 phon, checked by comparing the loudness with the reference tone and a calibrated microphone.

Prior to the start of the measurements, the delay difference is set at zero, and the position of the head of the observer is chosen so that the sound appears to arrive directly from the front. (See also the remark on directional hearing below.) The observer maintains his position during the subsequent measurements. The imaginary sound source which for the zero difference had been located directly in front was found to move gradually in the direction of the primary loudspeaker when differences between zero and 1 ms were tried, i.e., very small differences influence the directional impression. Now let the speech from the echo loudspeaker be delayed by 10 ms, corresponding to a path difference in air of approximately 3.4 m. This has the remarkable effect that

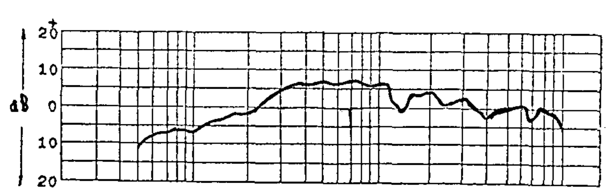

Fig. 6. Total frequency response curve of echo apparatus.

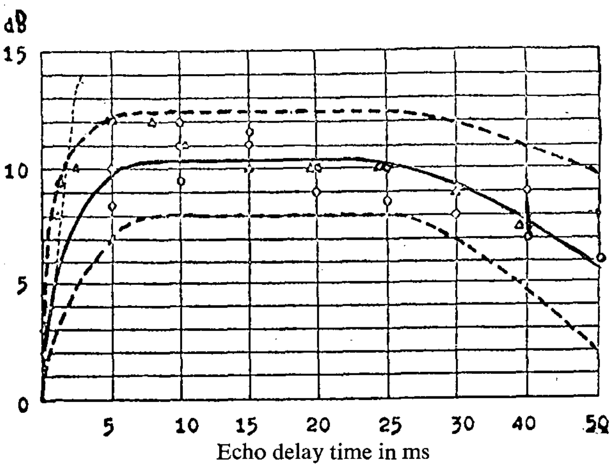

Fig. 7. Echo suppression effect as a function of delay time for speech.

the echo loudspeaker is not heard at all, although its energy is equal to that of the primary loudspeaker. Apparently the speech issues only from the latter. An explanation for this phenomenon cannot be given. It is probably due to a function of our central nervous system, and a possible explanation may be some kind of blocking reaction, caused by the first sound impression, preventing the separate perception of repetitions of the same impression following after brief intervals. Possibly the rise with time of the nerve reaction currents contributes to directional hearing.

The factors which acoustically enable us to state the direction of a sound reaching the ears have been studied thoroughly by v. Bekesy [5], Reich and Behrens [6], de Boer [7] and Warnecke [8]. These authors discovered that the directional discrimination is produced by differences of transit time and intensity as the sound reaches the two ears. However, the differences observed by these workers were mostly in the range between zero and 0.62 ms, corresponding to a path difference in air up to 21 cm. This is nearly the largest path difference with which sound impressions from a single source can reach our ears. Further increases of the path difference do not affect the directional impression.

What then can be observed when the difference is further increased? Hearing both loudspeakers with differences from 1 to 30 ms differs from hearing only one of them by a modification of the quality of the sound and a greater loudness, which will be discussed at a later stage. The impression is that of more "liveliness," the sound is growing in "body," and the sound source gains volume. This "pseudostereophonic" effect has been known for some time and was utilized as early as 1926 for the construction of the ultraphone [9], that is a normal disc recorder with two mechanical pickup heads arranged to pick up from the same disc with a time difference of 1/15–1/30 s. The same effect can be achieved with two loudspeakers operating simultaneously while the transit time difference is obtained by different distances between the two loudspeakers and the observer.

This condition is obtained without marked modification between 1- and 30-ms differences. Only when the difference reaches a value around 40 ms is the loudspeaker recognized as an additional sound source, but the sound source is still located at the position of the primary loudspeaker. A further increase beyond 50 ms makes it possible to discriminate a separate echo, but the "center of gravity" of the sound emission still rests on the primary loudspeaker.

The following method was employed for the quantitative determination of the subjective "suppression effect" observed for small delay differences, as a function of the echo delay time.

Both loudspeakers, arranged according to the above description, gave a continuous spoken text. The delay difference was statistically altered, in steps, from 1 to 40 ms.

An attenuator, calibrated in dB, enabled the observer to reduce the intensity of the primary loudspeaker until both appeared to be heard at the same loudness. This does not produce any "direct-in-front" impression as in the case of very small differences. Rather the sensation of two sound sources emitting from different directions, coinciding with the position of the loudspeakers, is produced. No difficulty was encountered by the observers in adjusting both sound sources to equal loudness of a spoken text, and scattering was relatively small. The location of both sound sources at an angle of 45° to the left and right side of the observer proved particularly favorable for loudness balancing.

The results obtained with 15 observers are shown in Fig. 7, where the difference of intensities in dB is plotted as a function of the delay time, that is, the reduction of the intensity of the primary loudspeaker required to produce the sensation of equal loudness for both sources. The full-line curve gives the mean value of all measuring points, the dashed lines show the limits of the deviations.

A distinction has been made between the points obtained from two different observers, with the purpose of showing the quantity of the deviations for individual observers.

De Boer gave the data on which the thin dotted line was based, but he obtained them by a different, indirect, method. His starting point was the perception that differences of time as well as of intensity reaching the ears of the observer bring about a directional impression. He placed two loudspeakers 3.5 m from each other, behind a canvas screen, focused on an observer standing in the direction of the median line to the connecting line of the two loudspeakers, at a distance of 3 m from the screen, which was used only in order to conceal the

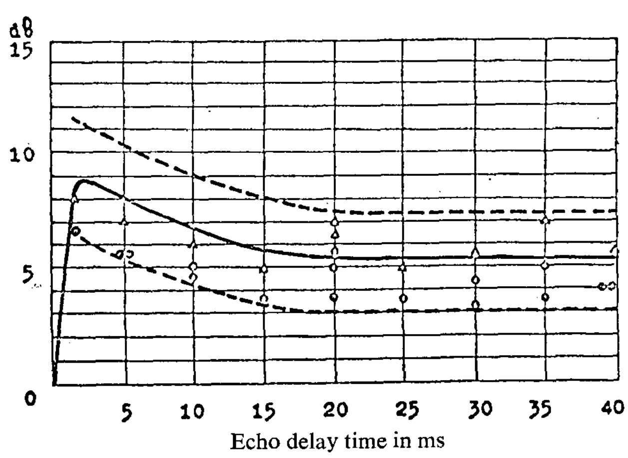

Fig. 8. Echo suppression effect as a function of delay time for noise with frequencies between 200 and 400 Hz.

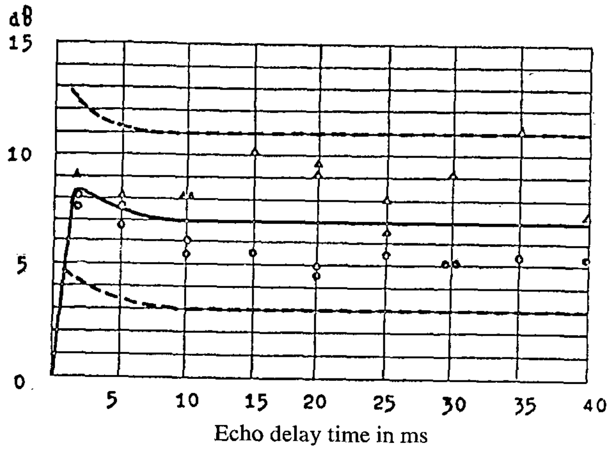

Fig. 9. Echo suppression effect as a function of delay time for noise with frequencies from 3200 to 6400 Hz.

loudspeakers from the observer. Both loudspeakers were excited simultaneously from a record and were adjusted to equal loudness, and in a first test the intensity of one loudspeaker was varied by means of a calibrated potentiometer. It was found that the hearer received different directional impressions according to the ratio of the intensities. The particular directional sensation belonging to each intensity ratio was drawn in a diagram.

A second test was based on equal intensities, where various small time differences were obtained by one loudspeaker being moved away from the observer. Again the different directions corresponding to delay differences were drawn in a diagram.

From both sets of measurements de Boer derived his curve shown in Fig. 7. He eliminated the directional impression and plotted intensity difference against time difference belonging to the same directional impression.

The agreement with our curve, obtained by a different method, is quite satisfactory, taking into account the wide scattering to be expected for subjective measurements. It appears from Fig. 7 that for differences from 5 to 30 ms the intensity of the echo loudspeaker must be ten times greater than that of the primary loudspeaker, i.e., against 10 dB, in order to create the impression of equal loudness. Above 30 ms the intensity difference drops slightly and the points are more scattered.

The experiment was repeated with the speech replaced by two noise spectra of different frequencies. For this purpose a white-noise generator was constructed with a gas-filled triode. Frequency ranges, as required, were obtained from the spectrum with an octave bandpass filter; they were then transmitted to the loudspeakers over the echo apparatus. The results are shown in Fig. 8 for a frequency range from 200 to 400 Hz and in Fig. 9 for a frequency range from 3200 to 6400 Hz. The mode of presentation and the measuring points correspond to Fig. 7. It is seen that the shape of the curves differs considerably from that in Fig. 7. The explanation may be that the effect in question is conditioned by a pronounced character and transit time difference of the sound impressions. These demands are fulfilled far less by noise spectra than by speech, and that may account for the rapid decline of the curves in Figs. 8 and 9, especially for high frequencies. Hence the balancing of the loudness was much more difficult than for speech, especially for the high frequencies, as is obvious from the wider scattering of the measurements. It is likely that this effect of the auditory suppression of echoes with a brief time difference enables us to discriminate directions in enclosed spaces where the hearing mechanism integrates only the first impression, that from the direct sound.

This effect is particularly conspicuous when loudspeaker arrangements operate in the open air, as for instance in large public meetings. As a rule, several loudspeakers of the same type, emitting equally in all directions (for instance mushroom loudspeakers), are fed simultaneously with the same signal. An observer moving from the vicinity of one loudspeaker in the direction of another one hears at first only the nearer loudspeaker. This may be explained by the ratio of energies. The impression persists unambiguously when the hearer moves on toward the second loudspeaker, although at this stage the energies emitted from both sources are almost equal, so that both should be heard.

However, as soon as the middle of the distance has been passed, if only by a few centimeters, the sensation passes suddenly to the other loudspeaker, with the impression that the sound comes from this direction only. This abrupt change of the impression is due entirely to delay phenomena.

A further test, based on speech, concerned the possible influence of the direction of the echo on the suppression reaction. No influence was found to exceed the limits of deviations in Fig. 7. This agrees with experience. Briefly delayed echoes in enclosed spaces may come from all directions without a disturbing effect. We mention in this connection a technical application of the suppression effect of briefly delayed echoes pointed out already by Cremer [10]. It is quite feasible to amplify sounds by 10 dB, possibly more, without the amplifying source being perceived, if the hearers are at the same time distracted by a further, e.g., optical, impression.

We will now discuss the increase in loudness caused by echoes with brief delay differences. In our case the echo has the same intensity as the direct sound, and the hearing organ receives therefore twice a certain, equal, energy, separated by the transit time difference. Our susceptibility to sounds is conditioned by a certain inherent inertia. For instance, according to v. Bekesy [11] the end value of the loudness perceived is reached only 200 ms after the tone has been switched on. On the other hand, investigations made by Steudel [12] showed that after the sudden switch-off of a tone the subjectively felt loudness decays only gradually. Bürck *et al.* [13], in conjunction with their development of an objective registering loudness meter, determined the time constants of the subjective building-up and decay sensation from the measurements made by v. Bekesy and Steudel. The mean values were 130 ms for v. Bekesy and 50 ms for Steudel. The conclusion suggests itself that our hearing mechanism integrates the sound intensities over short time intervals similar, as it were, to a ballistic measuring instrument. If we apply the law of the addition of energies to the echo with a short time difference, we obtain for double sound intensity a loudness increase of 3 phon, taking into account the nearly logarithmic sensitivity of the ear. Measurements with speech and music undertaken by Aigner and Strutt [14] confirmed this, provided that primary sound and echo had the same timbre. However, when their timbres were markedly distinct, the authors reported an increase in subjectively felt loudness exceed-ing by 6–9 phon the value according to the law of the addition of energies.

Lübcke [15] repeated with noises the tests made by Aigner and Strutt and found, even in the case of equal timbre, for small path differences and equal intensities of primary sound and echo a subjective increase in loudness of 5–6 phon, i.e., 2–3 phon above the value to be expected according to the said law.

The first two authors used radio receiving sets as sound sources. The difference in transit time was produced either by different distances from the observer or, with equal distances, by double staggered pickup from discs. Lübcke placed the loudspeakers at the same distance and obtained the difference in time by the insertion of an air interval in a transmission channel.

The measurements of the subjective loudness were made in both investigations with a Barkhausen noise meter [16], i.e., by comparing the loudness with a 800-Hz sound.

Aigner and Lübcke stated that the magnitude of the said effect was very little influenced by the delay difference, provided that this was below the threshold of masking. Nevertheless, it is proposed to discuss here also the influence of different transit time differences. We used the same apparatus as for the study of the suppression effect. The increase in loudness was not measured with a loudness meter, but by direct comparison of the sound impressions with and without echo.

To this end the observer was provided with a switch to disconnect the echo loudspeaker and to vary the energy of the primary loudspeaker in this position of the switch. By switching back he restored the initial state, i.e., both loudspeakers emitting the same intensity. It proved quite feasible to balance the subjectively felt loudness for both cases by repeated operation of the switch.

The effect of an echo-loudspeaker sounding with small delay differences has been defined before as an alteration of the timbre without alteration of the directional impression, i.e., without acoustical perception of the second loudspeaker.

Our tests produced the surprising result that there was no increase in loudness beyond the value to be expected from the law of the addition of energies, i.e., 3 phon.

We tried all kinds of changes: alteration of the delay difference below the threshold of masking, variation of the directions of the incidence of direct sound and echo; the timbre of the echo was modified while the impression of equal loudness of both loudspeakers was maintained. The experiments were made in enclosed and open spaces and with different observers. Invariably the increase due to echo of the same intensity was 3 phon. The greatest deviations under different conditions and with different observers were $\pm$ 1 phon, i.e., such as are to be expected in subjective measurements.

A possible explanation of this discrepancy between the results of our and earlier investigations may be found in the method of measuring the increase in loudness. A direct comparison of the sound impressions in our case should be more reliable than the comparison of the single impression with a noise the intensity of which, according to Aigner–Strutt, could be changed only in steps of 5 phon. It is also surprising that none of the authors mentioned that the echo loudspeaker could not be located acoustically. The earlier findings may also to some extent be due to imperfections of the electroacoustic transmitting equipment of the time.

Summing up the results of the foregoing discussions we can say that single echoes with short delay differences are not perceived as echoes because of the inertia of our hearing mechanism. All we feel is an increase in loudness in concord with the law of the addition of energies, and a pleasant modification of the quality of the sound, an apparent enlargement of the sound source.

Before passing on to phenomena created by longer delay differences, we must point out an effect observed in connection with the study of the small differences, since the observation has a bearing on the technique of telecommunications. Let primary sound and echo be emitted from the same loudspeaker, e.g., by series connection of the output terminals of the playback channels I and II (Fig. 1), consequently operating only one loudspeaker amplifier and loudspeaker. This causes intense distortions of speech for delay differences up to approximately 20 ms, but they disappear almost entirely with greater differences.

This phenomenon is due to interference between primary sound and echo in electrical superposition, resulting in large linear distortions of the tone. According to the particular delay difference of the echo, certain frequency ranges are very much reinforced whereas others are completely wiped out, assuming equal intensities of sound and echo.

Strongly pronounced interference effects are less liable to occur when the delay difference exceeds the average duration of the single sounds in speech; practically no disturbance is experienced then. The said phenomenon has to be borne in mind, for instance, when a performance is recorded by several microphones and then played back on a single loudspeaker. As a rule, the sound will reach the microphones with a certain path difference. Distortions of the reproduced sound can be avoided by placing the microphones so that the value of the differences is either zero or lies between 20 and 40 ms, corresponding to a microphone distance of about 6–12 m. Another method of avoiding interference effects makes use of directional microphones, arranged so that in each case the impression is recorded mainly by a single microphone.

Interferences are not perceived in binaural hearing of sound and echo from two loudspeakers, i.e., with acoustical superposition, because the distance between the ears makes us always hear simultaneously at two different points of the sound field. Also there is shielding from head and body, and if reflections occur, as in practice they always do, such pronounced interference phenomena as in electric transmissions cannot arise.

Only consequence of small delay differences in acoustic superposition is the described directional impression.

EXPERIMENTS WITH LARGE DELAY DIFFERENCES

We now pass on to the investigation of the influence of single echoes with greater delay differences on the audibility of speech. Beyond a certain difference value we perceive primary sound and echo separate. Further extension of the interval impairs the clearness of the acoustical impression or, in the case of speech, the intelligibility. Although our capability of concentration enables us to overcome to a degree this disturbance, it will certainly render prolonged listening very tiring. Still further increase of the delay difference leads to rapid deterioration of the intelligibility.

The appropriate method for the study of the influence of an echo on the intelligibility of speech appeared to be the use of syllables, a method which has stood the test in judging telephone communications, and in some investigations on the acoustics of rooms [17].

Such measurements are based on the transmission of a number of logatoms (syllables devoid of meaning in ordinary language). They are read by a speaker and heard and written down by a number of observers. The percentage of correctly understood logatoms determines the intelligibility of syllables or logatoms. Since the lists of logatoms edited by CCIF [18] were not available, new lists have been prepared according to the frequency of the various sounds in the German language [19].

They are composed of one or two consonants as initial sounds, followed by a vowel, with one or two consonants at the end (e.g., gan, pris, murt, schlerm). A test text always has 50 logatoms. Numerous lists have been compiled, in order to prevent the observers from getting accustomed to the syllables and causing errors.

Our measurements took place on the roof of the Institute, so that room influences were avoided. The logatoms were always carefully spoken by the same speaker, and were written down by five observers. Every syllable undisturbed by echo was followed by one with echo, in order to make provision for fatigue or for alterations in the conditions of the team.

The observers themselves evaluated the measurements immediately. The procedure yielded two results for each measurement, viz., the intelligibility of the syllables for undisturbed and disturbed reproduction.

However, the very first preliminary experiments revealed that the conventional monosyllable logatoms are not suitable for the study of echo effects. The single echo is briefly delayed, so that the first echo sound will coincide in the most favorable case, i.e., at a certain speed of the speech, with the second direct sound when a logatom consists of three sounds. Thus the first direct sound will be heard clearly at any rate. The second sound, a vowel, will always be understood easily because its sound energy is much greater than that of the superimposed consonant. On the third sound of the direct speech the vowel of the echo is superimposed; it is therefore quite often not understood, but it is heard again, this time undisturbed, after the decay of the primary syllable. This conclusion has been confirmed by elaborate tests, that is to say, it is not possible to impair the intelligibility of monosyllable logatoms by single echos of any delay difference.

A superimposed noise merely causes equal deterioration of the intelligibility of undisturbed and disturbed texts.

Intelligibility tests in large (reverberant) rooms are sometimes made by speaking several words before and after the logatom, e.g., "Schreiben Sie bitte—klas—sorgfältig auf" (please write—klas—down carefully). This cannot be done in our case because the time interval between preceding and closing text cannot be kept sufficiently short owing to the rather difficult pronunciation of the logatoms. Besides, the capability of concentration enables the observers to suppress in their perception the framing text and to focus their attention on the logatom, a reaction that is bound to counteract the effect to be measured. Another procedure could be to engage the observer in some other distracting occupation, e.g., summing up columns of numbers, but it would hardly be possible to determine the degree of that distraction.

The next attempt was made with multisyllable logatoms. Some effects of the echoes were observed, but the reduction of the intelligibility appeared only in the case of delay differences far beyond the threshold of masking. This is due to the different speeds of speech which essentially determine the maximum of the admissible echo difference, as will be demonstrated later on. About 5 syllables per second are spoken in normal conversation, but the speed has to be reduced to a mere 1.5 syllables per second when logatom texts are to be spoken in such a manner that single sounds can be understood. Such measurements with a speed differing so considerably from normal speech would yield results whose conversion into data valid for normal speaking would necessitate special comprehensive investigations. We therefore definitely abandoned the method of the intelligibility of syllables for the present investigation.

Measurements yielding results suitable for direct practical application must be based on the transmission of a text spoken at the normal speed. A first attempt was made with numbers, still with a view to copying the logatom method in so far that the numbers were to be written by the observers, an easy task if the speed of the transmission is 4 syllables per second. This experiment became another failure due to the restricted quantity of numbers (zero to ten) and because generally they are easily recognized from their vowel. This leads to practically 100-percent intelligibility, even for extremely long echo delay differences.

The only possible expedient was therefore to give a continuous text true to practical conditions and with various echo disturbances, and then to leave to the observers the decision whether they felt disturbed or not. An echo was characterized as disturbing when listening became unpleasantly strained, although the text might still be fully understood.

The experiments were made by altering the delay difference in steps. A very long or very short interval allowed a practically instantaneous judgment, while near to the threshold of masking it took the observers about a minute to make up their mind.

The objection could be made that such a time is too short, and that listening for a longer period even to a slightly disturbed text would be far more straining and very tiring, so that measurements of this kind would show admissible echo delay differences which were too high. In practice, however, other circumstances exert an opposite and nuisance-reducing influence, for instance, we receive and integrate as a rule a visual beside an acoustical impression and we divide our attention between both impressions. A disturbance is felt less in such a case than when only one type of sensation is to be integrated.

This effect may be illustrated by the following example. The loudspeaker for the reproduction of sounds in cinemas is usually placed behind the middle of the screen. When an actor speaks from one side of the screen the

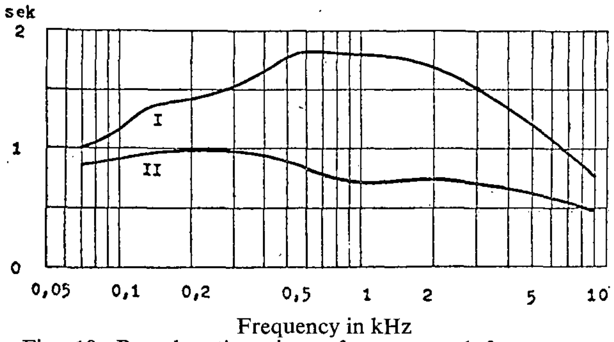

Fig. 10. Reverberation time of room used for measurements plotted against frequency. Curve I—empty; curve II—fully occupied.

sound is easily localized on the speaker. We do not realize at all that it arrives in fact from a quite different direction.

The assumption appeared therefore justified that useful results could be obtained in the manner described above, the more so as for statistical reasons the number of observers was raised to 50–100. Possible consequences of fatigue were reduced by restricting the time of intense attention to 30 minutes at the longest. The observers were mainly university students or scholars of higher schools.

The success of these measurements depends essentially on the order of the delay time alterations, particularly in the critical range. For instance, let a certain delay difference cause 50% of the observers to feel disturbed. If this particular difference had been preceded by one less effective, the number of the disturbed observers would have been greater than if the previous difference had produced a more effective disturbance. We therefore always made two straddling measurements in the critical range for each degree of the disturbance, defined by the number of observers who felt disturbed. The mean of those two measurements gave a fairly good criterion for the real condition. It was not always possible to determine in a single series of measurements the influence of a particular parameter. In the case of a further series it was important that the other parameters remain unchanged. To that end texts with a neutral meaning were recorded on magnetophone tape in order to eliminate influences of differences in the manner of speaking. During recording in the open air we took special care to maintain the same loudness and a constant speed of speech. Before every measurement the loudness of the primary loudspeaker and the echo loudspeaker were adjusted.

The results obtained with such precautions for the same acoustical conditions from different observers did not diverge by more than 10%. This applies even to small groups of six persons, not only to physicists who had taken part several times already in echo tests but also, after a short time of training, to members of the industrial staff of the Institute who for the first time took part in such tests.

Such limits are relatively narrow for subjective measurements, and it is surprising that they could be maintained by introducing the definition of "feeling disturbed," considering that the degree of a disturbance might have been judged very differently by individual observers, and that also the character of the disturbance varied considerably, corresponding to the particular circumstances. However, the method proved successful in all its applications, and it can therefore be definitely accepted as suitable for the estimation of echo effects.

Technical reasons (extension and loadbearing capacity of the roof of the Institute building) made it necessary to transfer the tests to a closed space. We are nevertheless confident that the results obtained there can be applied in practice, because in most cases where echos with long time differences occur in rooms, they will be accompanied by reverberation. The influence of the reverberation time will be discussed in detail at a later stage.

The room used for these measurements was an auditorium, 12.5 m long, 6.0 m wide, and 4.5 m high, with 80 built-in folding seats of plywood, rising in ten stepped rows. The volume is about 290 m^3. Fig. 10 gives the reverberation time as a function of frequency for the empty and fully occupied room. The average reverberation time for speech is 0.8 second for the full room.

The two loudspeakers had to be placed at a small distance in order to obtain, as far as possible, equal energy conditions and the same transit time differences all over the room. They were placed in the middle of one of the small sides in front of the blackboard, 4 m distant from the first row of seats, 160 cm above the floor, and were focused on the hearers.

It seemed advisable to get an idea of the effect the particular position of an observer might have on the results of the measurements. This was achieved by separating for one set of measurements the results obtained from observers in the front rows from those from observers in the rear. The deviations were found much smaller than the scattering between the measured points in general, i.e., they were negligible.

A recorded text spoken at a speed of 5.3 syllables per second, that is, similar to normal speech, was used for all investigations, except those on the influence of the speed.

For practical purposes the loudness of the primary loudspeaker could be considered constant over the whole lecture room because of the considerable distance from the first row of hearers. The loudness was maintained for the occupied room at 55 phon, except when the influence of the loudness itself was tested. It was controlled, as before, by comparison with the reference tone.

The foregoing conditions provided the basis for studying the disturbances caused by echoes as a function of various parameters.

Influence of the Speed of Speaking

The influence of the speed of speech was investigated for equal intensity and timbre of direct sound and echo.

Three sets of measurements were taken with speeds of 3.5, 5.3 and 7.4 syllables per second, i.e., within the limits found for normal speech. The continuous text was disturbed by altering in steps either transit time difference between direct sound and echo, or any other parameter.

The hearer had to decide whether he felt disturbed by the echo by answering "yes" or "no." The percentage of disturbed observers plotted against the echo delay differences is seen in Fig. 11. An inspection of the figure indicates a rapid rise of the percentage above 40 ms, to reach 100% at 100 ms. The measured points for curve II, i.e., 5.3 syllables per second, are given separated for two

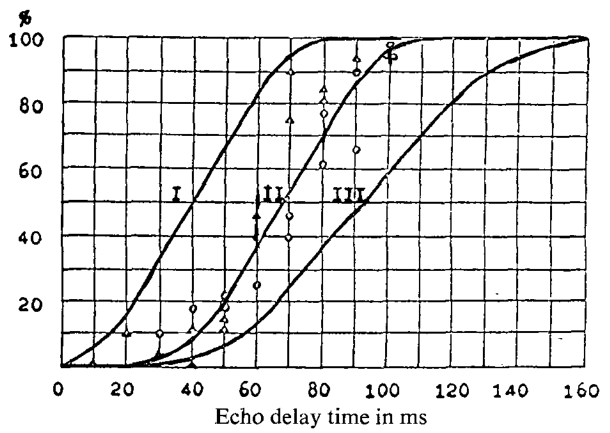

Fig. 11. Echo disturbance as a function of speed of speech. Curve I—7.4 syllables per second; curve II—5.3 syllables per second; curve III—3.5 syllables per second.

measurements under equal conditions, but with different observers, as is indicated by the round and triangular dots.

The evaluation of these results might be on the following lines. A critical difference would be assumed where 10–20% of the hearers felt disturbed, that is, 40 or 50 ms. Then follows a zone of relatively quick rise of the number of disturbed persons, due to differences in the individual judgment or to doubts about the decision. Differences above 80 ms to 90 ms cause 80 to 90% of the hearers to indicate disturbance, and it is a fair assumption that such echo delay times are certainly too long for satisfactory intelligibility.

The statement derived from those results would then be an admissible delay difference of 45 ms and a disturbing difference of 85 ms. However, the value of 45 ms would certainly be too low in practice, when the circumstances mentioned in the preceding are taken into account. We dealt here with wholly acoustical observations where the attention is concentrated on the hearing organ, whereas in the normal case optical impressions compensate to a certain degree for acoustical disturbances.

We considered this factor by drawing a curve with the closest possible approach to the points of the measurements, being a straight line in its middle section. The critical ratio of the delay differences is where 50% of the observers feel disturbed, that is, for curve II, 68 ms. This quantity is fairly close to the arithmetical mean between 45 and 85 ms and should serve as a good characteristic for the prevailing conditions.

The scattering of the measurements is rather small, considering their origins from two sets with different observers, and also the abrupt alterations of the degree of the disturbance. Greater deviations, e.g., in Fig. 11 those for 70 ms, are probably due to disturbing noises in the recorded text, erroneously taken as echo effects by some observers. Such incidental noises could not entirely be excluded, although the recording took place during the night hours (in the open air, in order to avoid room influences), e.g., sounds from aeroplanes, motor vehicles, and railways. On the whole they did not affect the results unduly.

The critical difference of 68 ms would not be affected appreciably when two separate curves would be drawn for the circular and triangular dots. The statistical demands are complied largely by the number of 2000 judgments derived from 80 observers in 25 measurements.

One item is worth mentioning. During all measurements made in the auditorium the first question was "when does the echo disturb you?" and a second question ran, "when do you hear an echo?" Again the answer was to be "yes" or "no," indicating whether the hearer was able to discriminate a difference at all from the anechoic transmission which started each set of measurements.

The answers to the second question showed no clear tendency, especially for small differences, and the results could not be evaluated. This failure is due probably partly to reverberation and partly to a psychological factor, i.e., the observer thought that almost always an echo was transmitted, as it was indeed, and he endeavored therefore to avoid the blame of being inattentive.

Fig. 11 makes clear how the critical delay difference depends on the speed of speaking. The value is 40 ms for high speed (7.4 syllables per second), 68 ms for normal speed (5.3 syllables per second), and 92 ms for slow (solemn) speech (3.5 syllables per second). That is to say, the critical difference is nearly inversely proportional to the speed of speaking. Of course, these statements should not be taken too rigidly, because the structure of speech is very involved and the above values are valid only for a certain range of speeds.

It is evident from these measurements that hearers may well follow without difficulty a slow speech in a room, whereas disturbances will arise with increasing speed of speaking.

Influence of the Intensity of the Echo

The next parameter to be varied was the intensity of the echo. The speed of speech for these tests was 5.3 syllables per second, the timbre of direct sound and echo was the same.

The intensity was altered in steps, the difference between the intensities of direct sound and echo being in turn + 10 dB, 0 dB, − 3 dB, − 6 dB, and − 10 dB. This investigation was divided into two series, the first with fluctuations between + 10 dB, 0 dB, and − 10 dB, and the second with fluctuations between 0 dB, − 3 dB, and − 6 dB.

Again the echo disturbance was altered abruptly, but its measure was now conditioned by differences of time as well as of intensity. The three sets of measurements were, as it was, interlocked. For instance, an echo disturbance defined as a delay difference of 60 ms and an intensity difference of − 6 dB were followed by one with 80 ms and 0 dB, these by 30 ms and − 3 dB, and so on. Deviations were not greater than in previous measurements, a further proof of the usefulness of the method.

The results are plotted in Fig. 12, indicating quantitatively how the nuisance effect of an echo can be reduced by lower intensity, as it is indeed normally in practice, viz., by covering reflecting surfaces with sound-absorbing materials.

The critical delay time, being again 68 ms for equal intensities of both loudspeakers, rises rapidly to 108 ms when the intensity of the echo is reduced by 3 dB and to 175 ms for 6 dB, while almost no disturbance at all

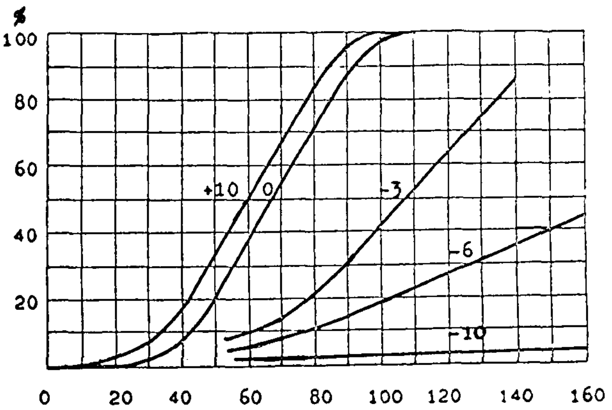

Fig. 12. Echo disturbance as a function of intensity of reflected sound. The numbers show the difference between the intensities of sound and echo in dB.

is felt when the reduction amounts to 10 dB. This value of 10 dB obtained for the complete stopping of the disturbance by reduction of the echo intensity is exactly the same as the magnitude of the echo-suppression effect for small differences (see Fig. 7).

An increase of the echo intensity by 10 dB occurs, for instance, through concentration of sound from curved surfaces, as in domes and planetariums of an out-of-date design. In such a case the critical delay time is reduced only from 68 to 60 ms.

Influence of the Timbre of the Echo

We now propose to show the effect of the timbre of an echo and that its influence on the subjectively felt disturbance cannot be neglected. We begin with the description of an experiment that is very suitable for demonstration and can easily be arranged by using a commercial magnetophone apparatus.

Two loudspeakers L_I and L_{II} are placed at a distance of several meters opposite each other and set at equal loudness (Fig. 13). Let L_I be the primary source and L_{II} the echo, with a delay difference of 60–80 ms. This can be produced with normal magnetophone equipment by feeding L_I with the level furnished to the recording channel, whereas L_{II} receives the level induced during the recording operation in the playback head. The delay difference is determined by the distance between recording head and playback head and its value usually reaches the required amount at a tape speed of 77 cm/s.

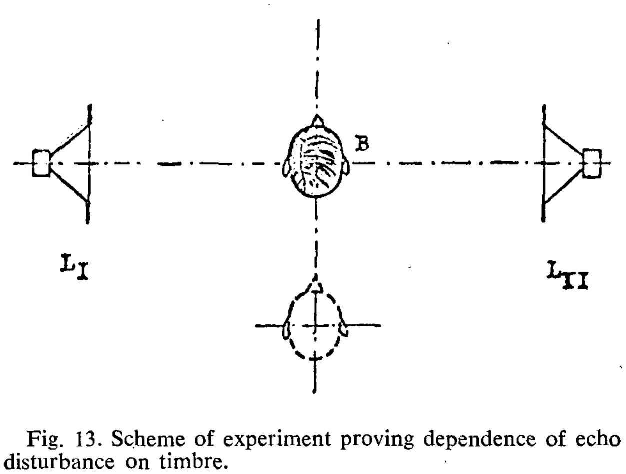

Fig. 13. Scheme of experiment proving dependence of echo disturbance on timbre.

An observer on the connecting line between the two loudspeakers will distinctly hear an echo, but this impression will decay gradually and disappear at last, when the hearer moves away on the median line. The delay difference is still the same, the ratio of the energies at the observer's position remains constant. The loudness is slightly reduced because of the larger distance between hearer and loudspeaker and the directional characteristics of the loudspeakers, but this is not relevant for the present investigation.

Only the timbre of the sound is considerably affected because the intensity of the higher frequencies decreases more and more when the distance from the connecting line grows, owing to their directional emission. The reduction of the echo nuisance must therefore be due to the absence of the high frequencies in the sound. The disturbance returns at once when the echo source, or both loudspeakers, are focused on the observer. The echo nuisance can be reduced also for an observer between both loudspeakers when the high frequencies are attenuated in the transmission channel to the echo loudspeaker.

This experiment reveals the importance of a high-class transmitting equipment for all previous investigations. The results would have been very much impaired by using disc reproduction with an upper limiting frequency of 5000 Hz.

The above experimental arrangement is not suitable for quantitative evaluation, because the directional characteristics of a cone loudspeaker as functions of frequency show a very complex and irregular pattern.

The difficulties of a quantitative evaluation, due to the appearance of numerous secondary maxima in case of the emitter being large in relation to the wavelength, had to be expected already from the directional characteristics computed for a piston diaphragm [20]. Nevertheless, we investigated how far those computed curves would agree with measured curves for a cone loudspeaker. The directional characteristics of one of the loudspeakers used in the above tests (diameter of diaphragm 30 cm) were recorded as a function of frequency. There is very poor agreement, especially for higher frequencies.

While the expected figure of an 8 was indeed measured for the directional characteristics at frequencies up to about 1000 Hz (diameter of emitter $\approx$ 1 wavelength), the curves measured for higher frequencies differed very much from the theoretical curves. This is due to unequal vibrations of different sections of the diaphragm, and to deflections of the baffle board and parts of the loudspeaker. The conditions become so involved that only the maximum fluctuations of the intensity in the region of a certain angle can be given.

In our particular case the fluctuation in intensity due to the directional effect, for instance for frequencies between 80 and 1000 Hz, kept below 10 dB only in the region of an angle of $\pm$ 20°. In the region between 0° and 180° we found fluctuations up to 40 dB.

Echoes for the quantitative investigation of timbre modifications were produced electrically by distortion of the transmission characteristic (Fig. 4) of the loudspeaker amplifier in the echo channel. This measurement was likewise made in the auditorium and with numerous ob-

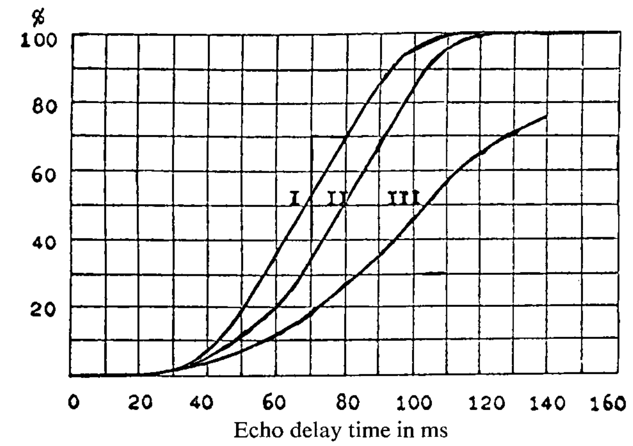

Fig. 14. Echo disturbance as a function of timbre of echo. Curve I—equal timbre of sound and echo; curve II—low frequencies of echo attenuated; curve III—high frequencies of echo attenuated.

servers. The speed of speech was 5.3 syllables per second; both loudspeakers had the same intensity.

The timbre of the echo from curve I in Fig. 4 was modified according to curves II and III, and again timbre differences and delay time differences were interlocked. The attenuation characteristic was intentionally chosen with a gradual rise at either the high or the low frequencies, as it agrees with practical conditions. In practice, attenuation of low frequencies similar to curve II in Fig. 4 would ensue from the use of resonance absorbers, an attenuation of the high frequencies according to curve III in Fig. 4 from porous absorbing materials.

Fig. 14 gives the results of the measurements. Curve I applies to equal timbre of sound and echo, the critical difference being again 68 ms. Curve II, obtained by attenuation of the low frequencies, gave a critical difference of 105 ms.

The objection could be put forward to the shape of the curves drawn in Fig. 14 that they indicate not so much alterations of timbre but rather the decrease in loudness connected with it, since this determines the critical difference (see Fig. 12). We checked this by comparing the subjectively felt loudness of a text transmitted according to curve I of Fig. 4 with a transmission according to curves II and III. The finding was that in both cases the decrease in loudness, caused by attenuating part of the transmitted frequency range, was less than 1 phon. An exacter result cannot be obtained by subjective measurements because of the physiological properties of our ear, but a quantitative estimate of the reduction in loudness can be derived from the following deliberation.

The spectral distribution of the energy in speech is highly frequency conditioned [7, Fig. 4]. The energy content per cycle in the frequency range from 100 to 1000 Hz is almost constant, but it decreases rapidly at higher frequencies, to become about 80 dB lower at 10 000 Hz than for 1000 Hz. An inspection of the transmission characteristics we used for effecting the timbre modulations (Fig. 4) indicates that curve II attenuates mainly regions of a high energy content and is apt to have a greater influence on the overall intensity than curve III which produces attenuation only in the frequency ranges of lower speech intensity.

Based on this argument we assume, in agreement with our loudness comparison, a reduction in loudness by 1 phon through damping of the low frequencies and we correct curve II in Fig. 14 according to the results plotted in Fig. 12. The result is that the curve now becomes almost identical with curve I of Fig. 14.

Let the decrease in loudness for curve III (Fig. 14) be assumed to be 0.5 phon. The correction according to Fig. 12 gives likewise a reduction of the critical delay difference, viz., from 105 to 95 ms approximately, but this is still much higher than the value of 68 ms for curves I and II. A conclusion from the foregoing consideration is therefore that the high frequencies of the echo certainly cause greater subjective disturbance than the low frequencies. However, the results shown in Fig. 14 have a direct practical bearing because a timbre modulation is almost always accompanied by a reduction in energy (frequency-dependent sound absorption).

Fig. 14 provides the quantitative foundation for the familiar experience that disturbances from echoes are reduced by covering the disturbing reflecting surface with porous absorbers in order to absorb mainly the high frequencies.

The reason for the dependence on frequency of the susceptibility of the human ear to echo disturbances is not yet known. The phenomenon might be due to a frequency-dependent time factor of the build-up and decay process.

A noticeable observation in echo experiments, particularly in the open air, was the impression of hearing from the echo loudspeaker, after it had not been perceived at all at small delay differences, at first only high frequencies, that is, in speech chiefly sibilants.

Influence of Loudness

So far the loudness of the primary loudspeaker had been maintained at approximately 55 phon, and only parameters of the echo had been altered. We now investigated the disturbing effect of the echo as a function of the general loudness. In this measurement the speed of speech was again 5.3 syllables per second; both loudspeakers had the same intensity and timbre.

The intensities of the two loudspeakers were varied in steps simultaneously, viz., toward lower degrees of loudness, between 55 phon, 45 phon, and 35 phon, to keep close to practical conditions. This was done by diminishing the amplifier output that produced 55 phon by 10 and 20 dB. It was found that the modification of the loudness in the said range exerted no influence on the subjectively felt disturbance.

Influence of the Angle of Incidence of the Echo

Stumpp [3] found an influence of the direction of the origin of the echo on the subjective disturbance value. His investigation dealt only with lateral incidence of the primary sound, while the echo arrived either from the same or from the opposite direction. The common experience, however, is that the sound comes from the front, while the echo may come from any direction. In our tests this was taken into account.

The measurements described so far had been made with numerous observers in a closed space, with primary sound and echo coming from the same direction, direct in front. This was necessary in order to create, as far as possible, uniform experimental conditions, independent of the position of the individual observer. Such a state would not obtain if for instance the loudspeaker was placed at a side wall or rear wall, because the intensities and delay differences would vary very much for different parts of the room, so that in such a case the observers ought to be concentrated as near as possible to a single point. Instead of 100 persons taking part we had to cut down the number of observers to a few. On the other hand, a small number makes it possible to operate on the roof of the Institute, where conditions are very favorable, an important factor, especially when the influence of direction comes into play.

We first tried direct comparisons between two different sound impressions, the method which had been applied so successfully in our previous tests. A single observer was asked to compare directly a constant reference disturbance from an echo in front of him with the sensations caused by echoes incident from various directions. The two degrees of disturbance had to be balanced by adjusting the delay difference of the random echo.

The experiment proved that such a direct comparison is not feasible because these sound impressions differ too much. To give a somewhat exaggerated illustration, let a frontal echo with a certain delay difference convey the impression of a stammering speaker. The lateral echo would, for example, produce the effect of two speakers talking at the same time. Monaural hearing would have made the two sound impressions comparable, but would not agree with practical conditions.

We tackled the problem in the following way. A group of six observers was concentrated on the smallest possible area, their eyes focused on the primary loudspeaker standing at a distance of two meters, level with their heads. Intensity and timbre were equal to those of the echo loudspeaker, placed anywhere on a circle of a 2-m radius, focused always on the hearers.

The delay difference of the echo was changed abruptly, within the interesting range, for different positions of the echo loudspeaker, as was done before in the auditorium tests, in order to obtain curves of the disturbance values. All measurements were made with two groups formed as before; the second group was again composed of members of the industrial staff to whom these measurements were entirely new.

The results obtained by group I showed no more scatter than with numerous observers used earlier in the lecture room, in spite of the small number of observers and the great difference in the impressions of the sounds. The deviations observed in the results obtained from group II decreased rapidly after a short exercise, for practical purposes reaching the level of group I. At that stage the values derived from both groups agreed to an astonishing measure, a proof of the correctness of results obtained even with so small a number of observers. Table I gives the results, viz., the critical delay difference as it corresponds to different angles of incidence, first separated for both groups, then combined. These values indicate a relatively small influence of the direction of the echo. The rise of the critical delay difference for lateral incidence is probably due to the decrease of the intensity at the remote ear compared with that of the direct sound.

Influence of Room Reverberation

The difference between the values of the critical delay difference obtained in the open air and in the fully occupied auditorium is considerable, i.e., 44 ms and 68 ms, respectively, for frontal incidence of both sounds and equal conditions. This leads to the conclusion that the reverberation time influences the critical difference.

The question arose whether further increase of the room reverberation time would be accompanied by higher critical delay differences. Measurements were taken with different groups, each of six observers, in the empty auditorium, having an average reverberation time of about 1.6 s (see Fig. 10).

Primary loudspeaker and echo loudspeaker emitted the same intensity and timbre, so that the loudness at the place of the observer was about 55 phon, the speaking speed was 5.3 syllables per second, the listeners were seated in two middle rows.

In these measurements the decision on the degree of disturbance was complicated because the reverberation time itself exerted an unfavorable effect on the reproduction of the speech. This was indicated by widely scattered results from less experienced observers. We selected therefore a number of discriminating listeners who were capable of judging with certainty the deterioration in audibility purely due to the echo, in spite of the disturbance caused by reverberation. In this way a critical transit time difference of 78 ms was obtained.

Table I

Direction and Angle of Incidence		Critical Transit Time Difference (ms) Group I	Group II	Groups I and II
Horizontal (frontal)	0°	46	42	44
	45°	53	50	52
Horizontal (lateral)	90°	52	51	52
	135°	50	52	52
Horizontal (from the rear)	180°	53	57	55
At 45° inclined from above	0°	49	45	47
	45°	59	53	55
	90°	60	56	56
	135°	58	59	58
	180°	62	57	59
Vertical from overhead		55	50	50

The effect of the reverberation time on the magnitude of the subjectively felt echo disturbance is plotted in Fig. 15. The critical delay difference increases with the reverberation time. Small delay differences seem to be veiled by the reverberation, so that their recognition becomes difficult.

CONCLUSION

A thorough investigation has been made of the influence of a single echo as a function of different parameters on the audibility of speech. An apparatus was built for the artificial generation of echoes with a time difference to the direct sound which could be varied. The

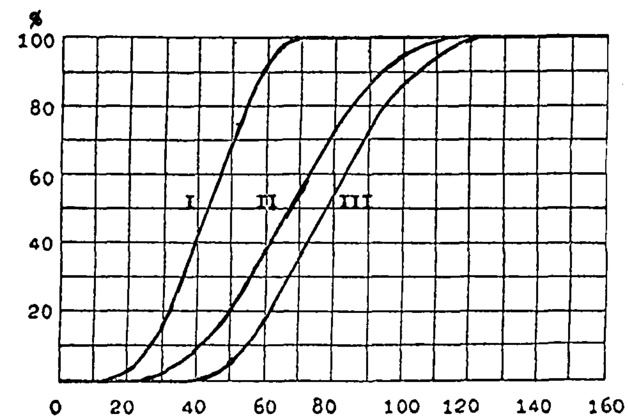

Fig. 15. Echo disturbance as a function of room reverberation time. Curve I—reverberation time 0 seconds; curve II —reverberation time approximately 0.8 second; curve III—reverberation time approximately 1.6 seconds.

apparatus is composed of a high-quality electroacoustic transmission installation and an endless-tape magnetophone equipment, from which the echo is obtained by a delayed pickup.

The measurements were made with a large number of observers under precisely specified conditions. Several suitable methods have been developed for the quantitative determination of the echo effects.

1) Tests with small echo delay differences between 1 and 30 ms showed an increase in loudness in agreement with the law of the addition of energies and a pleasant modification of the sound impression in the sense of a broadening of the primary sound source, while the echo source is not perceived acoustically. The magnitude of the auditory "suppression effect" for echoes with 1- to 30-ms difference was found to be 10 dB, i.e., the intensity of the echo must exceed that of the primary sound by 10 dB in order to make the echo separately perceptible in the said range of delay differences.

2) Echoes with greater delay differences above a rather marked threshold value cause the sound impression of speech to be disturbed, even to complete unintelligibility.

The conception of a "critical delay difference" is introduced as a mean of gauging the amount of disturbance, and of comparing measurements made under different circumstances.

a) Its value is approximately inversely proportional to the speed of speech in the range of 3.5 to 7.4 syllables per second.

b) The intensity of an echo exerts an important influence on the critical delay difference. An attenuation of the echo intensity by only 5 dB doubles the critical difference. Echo intensities more than 10 dB below that of the direct sound do not disturb at all the reproduction of continuous speech.

c) The high frequencies of the echoes determine the amount of the subjective disturbance. Their attenuation makes possible a considerable raising of the critical difference, with almost no noticeable reduction of the loudness of the echo.

d) The quantity of the echo disturbance does not depend on the loudness in the range belonging to speech.

e) The direction of incidence of the echo does not affect essentially the critical difference, provided that the direct sound is incident from the front.

f) A longer reverberation time in a room produces a greater critical delay difference.

REFERENCES

[1] Petzold, *Elementare Raumakustik* (Berlin, 1927), p. 8.

[2] H. Decker, "Eine Verzögerungsleitung für Messung und Vorführung von Laufzeitwirkungen in Fernmeldesystemen," (A delay conduit for measuring and demonstrating transit time effects in telecommunication systems), *Elek. Nachr. Tech.*, vol. 8, p. 516 (1931).

[3] H. Stumpp, "Experimentalbeitrag zur Raumakustik" (Experimental contribution to the acoustics of rooms), *Beihefte z. Ges. Ing.*, vol. 2, p. 17 (1936).

[4] W. Burk and H. Lichte, "Über die Schallfortpflanzung in Röhren" (The propagation of sound in tubes), *Akust. Z.*, vol. 3, p. 259 (1938).

[5] v. Bekesy, "Über das Richtungshören bei einer Zeitdifferenz oder Lautstärkeungleichheit" (On directional hearing in the case of a time difference or of unequal loudness), *Phys. Z.*, vol. 31, pp. 824, 857 (1930).

[6] M. Reich and H. Behrens, "Das Richtungsempfinden bei Tönen und Klängen" (Directional hearing of tones and sounds), *Z. Tech. Phys.*, vol. 14, p. 6 (1933).

[7] K. de Boer, "Plastische Klangwiedergabe" (Stereophonic sound reproduction), *Philips Tech. Rev.*, vol. 5, p. 108 (1940).

[8] H. Warncke, "Die Grundlagen der raumbezüglichen stereophonischen Übertragung im Tonfilm" (The fundamentals of stereophonic transmissions in the sound film with reference to rooms), *Akust. Z.*, vol. 6, p. 174 (1941).

[9] A. Lion, "Das Küchenmeister-Intervall und das Ultraphon," *Z. VDI*, vol. 70, p. 33 (1926).

[10] L. Cremer, *Geometrische Raumakustik* (Geometrical room acoustics), vol. 1 (1948), p. 126.

[11] v. Bekesy, "Zur Theorie des Hörens" (On the theory of hearing), *Phys. Z.*, vol. 30, p. 118 (1929).

[12] U. Steudel, "Über Empfindung und Messung der Lautstärke" (On the sensation of loudness and its measurements), *Z. Hochfreq. Tech. Elektrotech.*, vol. 41, p. 116 (1933).

[13] Bürk, Kotowski, and Lichte, "Die Lautstärke von Knacken, Geräuschen, und Tönen" (The loudness of clicks, noises, and tones), *Elek. Nachr. Tech.*, vol. 12, p. 278 (1935).

[14] F. Aigner and M. J. O. Strutt, "Über eine physiologische Wirkung mehrerer Schallquellen auf das Ohr und ihre Anwendung auf die Raumakustik" (On a physiological effect on the ear from several sound sources and its application to room acoustics), *Z. Tech. Phys.*, vol. 15, p. 355 (1934).

[15] E. Lübcke, "Über die Zunahme der Lautstärke bei mehreren Schallquellen" (On the increase in loudness obtained from several sound sources), *Z. Tech. Phys.*, vol. 15, p. 77 (1934).

[16] H. Barkhausen, "Ein neuer Schallmesser für die Praxis" (A new sound meter for practical use), *Z. Tech. Phys.*, vol. 7, p. 599 (1926).

[17] H. Panzerbieter and A. Rechten, "Subjective Bestimmung der Güte von Fernsprechverbindungen" (Subjective determination of the quality of telephone connections), *Arch. Tech. Messen*, vol. 3, p. 3719 (1942).

[18] CCI, "Recueil de Listes de Lagotoms," p. 2; also CCIF, White Book I, pp. 1-289.

[19] H. F. Mayer, "Verständlichkeitsmessungen an Telefonie-Übertragungssystemen" (Intelligibility measurements on telephone transmission systems), *Elek. Nachr. Tech.*, vol. 4, p. 184 (1927).

[20] H. Stenzel, "Über die Richtwirkung von Schallstrahlern" (On the directional effect of sound emitters), *Elek. Nachr. Tech.*, vol. 4, p. 239 (1927).