13. INTERNATIONAL CONGRESS ON ACOUSTICS

13. CONGRES INTERNATIONAL D'ACOUSTIQUE

13. INTERNATIONALER KONGRESS FÜR AKUSTIK

VOLUME 2
THEMES/
THEMES/
THEMEN: 2, 3, 4

Belgrade, 1989
13.
INTERNATIONAL CONGRESS
ON ACOUSTICS

13.
CONGRES INTERNATIONAL
D'ACOUSTIQUE

13.
INTERNATIONALER KONGRESS
FÜR AKUSTIK

Volume II

THEMES
THEMES
THEMEN: 2, 3, 4
# Table of contents

## Sommaire

## Inhaltverzeichnis

2. **AEROOACOUSTICS AND ATMOSPHERIC SOUND**  
AEROACOUSTIQUE, ACOUSTIQUE ATMOSPHERIQUE  
SCHALLAUSBREITUNG IM FREIEN

### 2.A

1. **PREDICTED PERFORMANCE OF PARALLEL NOISE BARRIERS USING BOUNDARY ELEMENT METHODS**  
HOTHERSALL D. C., CHANDLER-WILDE S. N., HAJMIRZAE M. N. (GB) ........................................... 3

2. **WIND EFFECT ON OUTDOOR SOUND PROPAGATION: EXPERIMENTAL STUDY IN WIND TUNNEL**  
ROSEN M., GABILLET Y. (F) ................................................................. 7

3. **COMPUTER SIMULATION OF OUTDOOR SOUND PROPAGATION: RASMUSSEN K. B. (DK)........................................................................................................11

4. **IMPULSE SOUND MEASUREMENTS FOR CHECKING THE METEO WINDOW**  
KURZE U. J. (D) ............................................................................. 15

5. **THE INFLUENCE OF METEOROLOGY ON GRAZING INCIDENT SOUND**  
WEMPE N. J. (D) ........................................................................ 19

6. **IMPULSIVE SOUND SOURCE OF HIGH INTENSITY FOR OUTDOOR SOUND PROPAGATION MEASUREMENTS**  
RADEK U., KLUG H., MELLERT V. (D) ............................................................. 23

7. **ACOUSTIC PROBING OF METEOROLOGICAL AND ACoustical PARAMETERS IN OUTDOOR SOUND PROPAGATION**  
MELLERT V., KLUG H., RADEK U. (D) ............................................................. 27

### 2.B

1. **CONTROL OF INFRASONIC AND LOW FREQUENCY NOISE**  
FUCHS G. L. (AR) ........................................................................... 33

2. **THE INFRA-SOUND FROM A VORTEX LOCALIZED NEAR THE OCEAN SURFACE**  
SUKHOREVUK A. I., PAVLOV V. I. (RU) ...................................................... 37

3. **MULTIPOLE IDENTIFICATION FOR INFRASONIC SOUNDS**  
PAVLOV V. I., ZHUkov A., NESTANIKOV M. (RU) ................................................... 41

### 2.C

1. **SCATTERING OF NON-ACOUSTIC ENERGY BY GEOMATERIAL INHOMOGENEITIES ON A FLAT PLANET**  
POWELL A. (USA) ........................................................................... 47

2. **L'AOUOUSTIQUE DES REVÊTEMENTS ROUTIERS**  
BERENGIER M., HAMET J. F., LETHuin P., DABURON P. (F) ................... 51

3. **ADVANCES IN THE REDUCTION OF SUPersonic JET NOISE THROUGH RADIAL DIFFUSION**  
DE KRASINSKI J., SUN S. (CDN) ................................................... 55

4. **PLANNED REVISION TO THE AMERICAN NATIONAL STANDARD FOR EXPLOSION AIRBLAST**  
REED J. W. (USA) ........................................................................... 59

5. **ON THE SOUND FIELD OF ORGANIZED VORTICITY IN JET FLOWS**  
MUSAFIR R. E. (BR) ........................................................................... 63
6. SOUND GENERATION IN UNSTEADY VERTICAL FLOWS  
FEDORCHENKO A.T. (SU) .......................... 67

3. ARCHITECTURAL AND BUILDING ACOUSTICS  
ACOUSTIQUE ARCHITECTURALE ET DU BATIMENT  
RAUM- UND BAUKUSTIK

3.S NEW TRENDS ON SINGLE NUMBER INDICES AND SHORT METHODS FOR ACOUSTICAL INSULATION

1. REFERENCE CURVE METHOD VERSUS WEIGHTING NETWORK METHOD  
MORENO A. (E) ................................................................. 73

2. PREDICTION OF SINGLE NUMBER RATINGS FOR AIRBORNE AND IMPACT SOUND INSULATION  
GERRETSEN E. (NL) .......................................................... 77

3. THE DEVELOPMENT OF I. S. O. SHORT TEST METHOD FOR THE MEASUREMENT OF AIRBORNE SOUND INSULATION IN BUILDINGS  
MACKENZIE R. (GB) ......................................................... 81

4. DOIT-ON CHANGER DE COURBE DE REFERENCE POUR EVALUER LA QUALITE DE LA PROTECTION CONTRE LES BRUITS D'IMPACT  
JOSSE R. (F) ................................................................. 85

5. LABORATORY SOUND INSULATION INDICES NEEDED TO ACHIEVE THE REQUIREMENTS IN THE FIELD  
HERINGA P. H. (NL) ......................................................... 87

6. THE FUTURE OF SINGLE NUMBER INDICES FOR ASSESSING SOUND INSULATION AS SEEN BY THE CONSULTANT  
YERGES J. F. (USA) ........................................................... 91

7. COMPARISON BETWEEN EUROPEAN SOUND INSULATION RATING METHODS FOR DWELLINGS  
VIAH J.-P. (F) ................................................................. 95

8. FIELD EXPERIENCES AND RELIABILITY OF SHORT TEST PROPOSALS FOR AIRBORNE SOUND INSULATION MEASUREMENTS  
TAIBO L. N. (AR) ............................................................. 99

9. ANALYTICAL CONNECTIONS BETWEEN THE REFERENCE CURVE AND THE WEIGHT METHODS  
PARMANEN J. (SF) ........................................................... 103

10. CORRELATION BETWEEN OBJECTIVE RATINGS AND HUMAN REACTION TO IMPACT NOISE  
MCKELL B. (GB) ............................................................. 107

11. FREQUENCY RANGE CONSIDERATIONS ON SINGLE-NUMBER INDICES FOR ASSESSING ACOUSTICAL INSULATION  
WALKER K. W. (USA) .......................................................... 111

12. COMPARISON OF SINGLE-NUMBER INDICES FOR ASSESSMENT OF SOUND INSULATION OF BUILDING WALLS BASED ON PSYCHOACOUSTIC EXPERIMENT  
KOYASU M., TACHIBANA H. (J) ........................................... 115

3. A

1. ZUR ABSCHAETZUNG DES FREQUENZGANGES DER REFLEXIONSFAKTOREN BEI REchteckigen PLATTEN  
CREMER L. (D) ................................................................. 121

2. THE INFLUENCE OF DIFFUSE SURFACE REFLECTIONS ON ROOM SOUND FIELDS  
HODGSON M. (CDN) .......................................................... 125

3. ESTIMATION OF IMPULSE RESPONSE IN A SOUND FIELD BY THE FINITE ELEMENT METHOD  
CHOI S., TACHIBANA H. (J) .................................................. 129
4. MEASUREMENT OF IMPULSE RESPONSES IN A REAL AND SCALE MODEL AUSTORIUMS  
HIDAKA Y., TACHIBANA H. (J) .......................................................... 133

5. MEASUREMENT OF VERY SHORT REVERBERATION TIMES  
RINDEL J. H., RASMUSSEN B. (DK) .................................................. 137

6. LATE DECAY TIME IN CONCERT HALL ACOUSTICS  
GUZINA B. (YU) ............................................................................. 141

7. ACOUSTIC ENERGY PROPAGATION IN ROOMS OF DIFFERENT FORM  
TOPALOVA V. (BG) ........................................................................ 143

3. B

1. RELATION BETWEEN ABSORPTION COEFFICIENT aSTAT DERIVED FROM α (∂)  
AND aMEASUREMENTS IN THE REVERBERATION ROOM  
KAKU J., YOSHIMURA J. (J) ............................................................ 149

2. EDGE EFFECT INFLUENCE ON THE STATISTICAL ABSORPTION COEFFICIENT  
COPS A., LAURIKS W. (B) .............................................................. 153

3. ON THE SOUND REFLECTION OF THE AUDITORIUM SEATS  
ISHIDA K., SUGINO K., MASUDA I. (J) .............................................. 157

4. PREFERENCE ANALYSIS DERIVED FROM PERCEPTIVE TESTS IN ROOM ACOUSTICS  
LAVANDIER C., WARUSFEL O., JULLIEN J. P. (F) ............................... 161

5. LA CARACTERISATION PERCEPTIVE DES SALLES DE CONCERT: INSUFFISANCE DES CRITÈRES TRADITIONNELS ET PROPOSITION D’UN MODÈLE BASE SUR LA DISTRIBUTION DES PREMIÈRES REFLEXIONS  
FISCHETTI A., JOUHAINEAU J., HEMIM Y. (F) .................................... 165

6. COMPARED ANALYSIS OF THE CLARITY INDEX AND REVERBERATION TIME  
JULLIEN J. P., LAVANDIER J., WARUSFEL O. (F) ................................. 169

7. PERCEPTION OF COLORATION AND SPATIAL EFFECTS IN ROOM ACOUSTICS  
WARUSFEL O., LAVANDIER C., JULLIEN J. P. (F) ............................... 173

8. PHYSICAL AND PSYCHOACOUSTIC EVALUATION OF FM SIGNALS PROPAGATING IN A ROOM  
OZIMEK E. (PL) .............................................................................. 177

3. C

1. ORCHESTRA ENCLOSURES AND STAGE DESIGN IN MULTIPURPOSE HALLS  
USED FOR CONCERTS  
STROM S. (N) ............................................................................. 183

2. ON THE ACOUSTIC FEEDBACK MODEL FOR SOLO SINGERS IN CONCERT HALL STAGES  
SHIRASUNA S., NAKAMURA S. (JP) .................................................. 187

3. THE ACOUSTICS OF THE NEW "OPERA DE LA BASTILLE" IN PARIS  
MUeller H. A. (D), VIEN J.-PI (F) .................................................... 191

4. ASYMETRIE ACOUSTIQUE DU STADE OLIMPIC DE MONTREAL  
MIGNERON J.-G. (CDN) ................................................................... 199

5. HALL AVERAGE CHARACTERISTICS OF 10 HALLS  
BRADLEY J. S. (CDN) ..................................................................... 195

6. MONITOR-ROOM ACOUSTICS AND ITS SIMULATION USING HEAD AND TORSO SIMULATOR  
OKABE K., MUZUUCHI T., HAMADA K., HONDA T. (J) ..................... 203

7. ACOUSTIC IMPROVEMENT OF CLASSROOMS MEASURED WITH THE RASTI-METHOD  
PEKKARINEN E., MINNI E., VIJANEN V., KOSKELA H., HONKA J. (SF) 207

8. MEßUNG DER SCHALLAUSBREITUNG IN FLÄCHEN RÄUHMEN  
ELFFERT G. (D) ........................................................................... 209
3.D

1. AUDITORY SPACIOUSNESS AND ENVELOPEMENT
   MORIMOTO M., MAEKAWA Y. (J) ................................................................. 215

2. AN EXPERIMENTAL SYSTEM FOR KNOWLEDGE-BASED ARCHITECTURAL
   ACOUSTICS
   YANAGIDA M., HIDAKA T., YAMAGUCHI T. (J) ........................................ 219

3. RAUMAKUSTISCHE VERBESSERUNG BEIM UMBAU DER PAULSKIRCHE IN
   FRANKFURT AM MAIN
   WÄSTPHAL W., KUTTRUFF H. (D) ............................................................... 223

4. ACOUSTIC DESIGN OF TOKYO DOME
   SYO D., KAGEYAMA K., KIMURA H. (J) .................................................... 227

5. ACOUSTICAL DESIGN OF THE "KARLOS KOUN" THEATRE IN ATHENS, GREECE
   TSINAKIS G., TZERAKIS E. (GR) .............................................................. 231

6. ACOUSTIC DESIGN AND FINAL RESULTS OF RECONSTRUCTED THEATRES
   KOTSCHY A. (H) ......................................................................................... 235

7. VARIABLE ACOUSTICS. THE DESIGN OF DIFFERENT HALLS
   HERINGA P. H., PEUTZ V. M. A. (NL) ...................................................... 239

8. ACOUSTIC DESIGN OF VARIABLE REVERBERATION UNITS FOR MULTIPURPOSE
   HALLS
   FURUKAWA S. (J) ....................................................................................... 243

3.E

1. AN AUTOMATIC SUPPRESSION OF MICROPHONE-LOUDSPEAKER COUPLING
   - CONCEPT AND PRELIMINARY SIMULATION
   RIKO Y., MAKINO T. (J) ............................................................................... 249

2. BINAURAL SOUND SIMULATION OF CONCERT HALLS BY A BEAM TRACING
   METHOD
   MARTIN J., VIAN J.-P. (F) ......................................................................... 253

3. GEOMETRICAL ACOUSTICS DESIGN SOFTWARE
   TZERAKIS E. (GR) ....................................................................................... 257

4. EXPERIENCE WITH A NEW SYSTEM OF ACOUSTIC CONTROL
   DE VRIES D., BERKHOUT A. J. (NL) ........................................................ 261

5. A HOLOGRAPHIC APPROACH TO SPEECH REINFORCEMENT
   VOGEL P., BERKHOUT A. J., (NL) DE VRIES D. (NL) ................................ 261

6. THE DSP 610 - A COMPACT PROCESSOR TO UTILIZE A NEW DIRECTIONAL
   SOUND REINFORCEMENT SYSTEM, THE DELTA STEREOPHONY SYSTEM
   NADLER W. (A) .......................................................................................... 271

7. QUALITE ACUSTIQUE, ISOLATION ET INTIME DES IMMEUBLES
   D'HABITATION FN COPROPRIETES
   MARTEL CH., MIGNERON J.-G. (CDN) .................................................... 275

8. SOME EFFECTS ON THE LABORATORY MEASURED SOUND TRANSMISSION
   LOSS
   DONNER U. (D) .......................................................................................... 279

3.F

1. "STC" RATING VERSUS dB (A) - LABORATORY AND FIELD STUDIES
   VASS T. (AUS) ............................................................................................ 285

2. MEASUREMENTS ON THE SOUND REDUCTION INDEX OF WINDOWS AND
   GLAZED EXTERIOR DOORS
   PANAYOTOPoulos N., TZERAKIS E. (GR) ................................................... 289

3. AN INVESTIGATION OF ROOM-TO-ROOM SUSPENDED CEILING INSULATION
   MEASUREMENT BY ONE-TENTH SCALE MODELLING
   DODD G., MEYNIAL X. (NZ) ...................................................................... 293
4. STRUCTURAL TRANSMISSION - SOUND INSULATION OF TWO LAYER WALLS
REIS F. (H).................................................................................................................297

5. SOUND INSULATING PROPERTIES OF POROUS MATERIALS
NAKAMURA S. (J)......................................................................................................301

6. INFLUENCE ON THE AIRBORNE SOUND INSULATION OF BUILDING ELEMENTS
BY SLITS; DowELS AND PARTIAL COVERINGS
RUECKWARD W. (D)...................................................................................................305

7. THE PROPAGATION OF STRUCTURE-BORNE NOISE THROUGH MODELS OF
MODERN BUILDINGS
OSIPOV A. (SU)........................................................................................................309

3.G

1. AN ORIGINAL METHOD FOR IN SITU IMPEDANCE MEASUREMENT
IULZATI F., ADORF S., ROSSMUT D. (F)......................................................................315

2. RESTORATION AND RECONSTRUCTION OF THE LIBRARY BUILDING IN THE CITY
OF BELGRADE
GAVRILOVIC M. (YU)...............................................................................................319

3. A RAY-TRACING STUDY OF THE SOUND FIELD IN A FLEXIBLE MEMBRANE
STRUCTURE
IANNIELLO C., MAFFEI L. (I)....................................................................................323

4. DESIGNING TSS ROOMS
STANOJEVIC T., SAKOTA G. (YU)................................................................................327

5. CALCULATION OF THE TRANSMISSION LOSS OF PANELS WITH
NON-HOMOGENEOUS CROSS-SECTIONS
KURTIVIČ H. (YU)......................................................................................................331

6. MODEL EVALUATION IN BELGRADE'S NATIONAL THEATRE RECONSTRUCTION
MIĐIĆ M., KURTIVIČ H., PEROLO Y., ONICIN V. (YU)....................................................335

7. ACOUSTICS OF THE BELGRADE NATIONAL THEATRE
PEROLO Z. (YU)........................................................................................................339

3.H

1. IMPROVEMENT METHOD FOR FLOOR IMPACT SOUND INSULATION ON WOODEN
FRAME STRUCTURE (CONTROL OF FLOOR IMPACT SOUND ON WOODEN
FRAME STRUCTURE: PART 1)
INOUE K., KIMURA S. (J)..........................................................................................345

2. PRACTICAL APPLICATION OF METHOD FOR REDUCTION OF FLOOR IMPACT
SOUND TO WOODEN HOUSE. (CONTROL OF FLOOR IMPACT SOUND ON
WOODEN FRAME STRUCTURE: PART 2)
KIMURA S., INOUE K. (J)..........................................................................................349

3. THE PREDICTION OF NOISE FROM MACHINE VIBRATION
WALKER R., MATHERS C. D., FLETCHER J. A. (GB)....................................................353

4. EQUIVALENT FORCE AND ITS APPLICATION IN STRUCTURE-BORNE NOISE
BETWEEN MACHINE AND BUILDING
STENIČKA J. (C).........................................................................................................357

5. DESIGN OF PARTITION WALLS WITH HIGH STANDING INSULATION VALUES IN
THE RECONSTRUCTED ARCHITECTURAL STRUCTURE
ŠMIDHEN E. (YU).......................................................................................................361

6. SOUND INSULATION OF MULTIPLE GLAZING FILLED WITH LIQUID
GALAKTIONOVA T. I., ROZOVÁ K. B., SOLOJOV S. P. (SU)...........................................365
4. SPEECH COMMUNICATION
COMMUNICATION PARLE
SPRACHKOMMUNIKATION

4.A

1. REPRESENTATION OF THE SPEECH SIGNAL WITH ELEMENTARY WAVEFORMS:
A PRELIMINARY PERCEPTIVE STUDY
GROZEL S., LIENARD J. S., D'ALESANDRO C. (F).............................................. 369

2. UN NOUVEAU MODELE ACOUSTIQUE DE PRODUCTION DE PAROLE
MRAYATI M. (SYR), CARRE R., GUERIN B. (F).................................................. 373

3. ON THE EXTENT OF BITE-BLOCK COMPENSATION ON VOWEL ARTICULATION
RAFAEL L. J., FABER A. (USA).............................................................................. 377

4. ISOLATION OF MICROEVENTS IN SPEECH
STRONG W. J., BERRY D. A. (USA).................................................................. 381

5. FATIGUE EFFECTS ON VOWEL ARTICULATION
FABER A., RAFAEL L. J., LEHMAN H. H. (USA)................................................. 385

6. DIE ERKENNUNGZEIT DER SPRACHLAUTE
GREGURICI M., VUJNOVIĆ M. (YU).................................................................. 389

7. APPLICATION OF DIGITAL SIGNAL PROCESSING ON THE ARABIC LANGUAGE
AIHAM O. A., FL. BOGDADLY N. (ET)................................................................. 393

4.B

1. MACHINE RECOGNITION OF SPEECH USING TIED MIXTURE CONTINUOUS
MODELS
BELLEGARDA J. R., NAHAMOO D. (USA)............................................................. 399

2. DEEP UNDERSTANDING OF SPEECH
TAKIZAWA O., YANAGIDA M. (I)....................................................................... 403

3. MULTISPEAKER VOWEL-CLUSTER COMPRESSION BY FORMANT FREQUENCIES
NORMALIZATION
CHRISTOV PH. (BG).......................................................................................... 407

4. MACHINE SPEECH RECOGNITION BY BSDP, A KNOWLEDGE BASED-TEMPLATE
MATCHING ALGORITHM
KIŚ M., MATUŠEK M. R. (YU).............................................................................. 411

5. A LABORATORY SYSTEM FOR MACHINE SPEECH RECOGNITION EXPERIMENTS
BASED ON BSDP ALGORITHM
MITIĆ D. S., VIGNJEVIĆ V. Ž., JOVANOVIĆ M. R., KIŚ M. L. (YU)......................... 415

6. RECOGNITION OF THE PHONEMS BASED ON THE NETWORK MODEL
CHUCHUPAL V. J. (SU)..................................................................................... 419

7. VERWENDUNGSMÖGLICHKEIT DER BINÄREN DARSTELLUNG VON
SPEKTRALMERKMALEN IN WORTERKENNUNG
KUBZDELA II. (PL)............................................................................................. 423

8. SPEAKER IDENTIFICATION BY MEANS OF SPEECH SPECTRA FROM KEY WORDS
MAJEWSKI W. (PL)............................................................................................. 427

4.C

1. LA DUREE COMME INDICE DE RECONNAISSANCE ACOUSTIQUE DE DIVERS
PROCEDES DE VALORISATION EN FRANCAIS
FLAMENT B. (F)................................................................................................. 433

2. A FORMANT ANALYSIS OF ITALIAN DIPHTONGS
MATURI P. (I).................................................................................................. 437

3. TEMPORAL VARIABLES AND FUNDAMENTAL FREQUENCY FOLLOWING
UNILATERAL LEFT ANTERIOR OR POSTERIOR LESION
BHATT P. (CDN)................................................................................................. 441
4. D

1. LONG-TERM SPECTRAL DIFFERENCES OF CONTINUOUS SPEECH BETWEEN SPEAKERS OF ENGLISH AND FRENCH PRODUCED NORMALLY OR WITH MASKING NOISE WEISS W. (CDN)................................................................. 463

2. A CHINESE SPEECH SYNTHESIS SYSTEM CHONGZHI Y., FENG G., LONGMING Z. (PRC)........................................ 467

3. EXPERIMENTAL DETERMINATION OF ARRAY MINIMAL SIZE FOR HULL MILLER-FIELD-COMPENSATION FROM NEARFIELD MEASUREMENTS JORNEAU CH., ROHAN T., PASQUALI N. (F)........................................... 471

4. A COMPARISON OF A MAGNITUDE ESTIMATION TECHNIQUE AND A PAIR COMPARISON TECHNIQUE FOR ASSESSING TEXT-TO-SPEECH SYNTHESIS PAVLOVIĆ Č. (USA), SORIN CH., ROUMIGUIERE J. P., LUCAS J. P. (F).......................... 475

5. MAGNITUDE ESTIMATION TECHNIQUE IN EVALUATING TEXT-TO-SPEECH SYNTHESIS ROSSI M. (F), PAVLOVIĆ Č. (USA), ESPRESS R. (F)................................................. 479

6. ACOUSTIC STRUCTURE OF HINDI WORD RHYTHMS (ANALYSIS, SYNTHESIS, PERCEPTION) ROUMYNTSEVA I. M. (SU).......................................................... 483

7. APPLICATION OF KNOWLEDGE-BASED TECHNIQUES AND STOCHASTIC MODELLING FOR RUSSIAN SPEECH RECOGNITION KANITCHEVA E. B., GREBCHENKO A. L. (SU).................................................. 487

8. RUSSIAN SPEECH RECOGNITION AND TRANSLATION INTO OTHER LANGUAGES GREBCHENKO A. L. (SU)............................................... 487

4. E

1. EFFECTS OF SPEECH INTELLIGIBILITY UPON PERFORMANCE GARANTHER G. R., PETERS L. J. (USA)..................................................... 497

2. SOME PSYCHOLINGUISTIC ASPECTS OF TEXT PERCEPTION AND UNDERSTANDING IVANOVA G. V., SHAKHANOVICH A. M. (SU).............................................. 501

3. SPEECH COMMUNICATION IN NOISE AND REVERBERATION: AN EXPERIMENTAL APPROACH LEGROS C., GAMB A. R. (F)......................................................... 503

4. MODELLING OF THE SYSTEM "ATTENTION-MEMORY-PERCEPTIVE PHONETICS" MRCHEV S. J. (BG)................................................................. 505

5. NASOPHARYNGEAL TRACT TRANSFER FUNCTIONS: MEASUREMENTS WITH WHITE NOISE EXCITATION CASTI E., BADIN P. (F)................................................. 511

6. ECHO ACCEPTANCE CONVERSATION TESTS PASCAL D., CORVAISIER D. (F).......................................................... 515
2. A
PREDICTED PERFORMANCE OF PARALLEL NOISE BARRIERS USING BOUNDARY ELEMENT METHODS

D. C. Hothersall†, S. N. Chandler-Wilde* and M. N. Hajmirzae†

†Department of Civil Engineering, University of Bradford, Bradford, BD7 1DP, U.K.

*Department of Mathematics, Coventry Polytechnic, Priory Street, Coventry, CV1 5FB, U.K.

Introduction

In this paper predictions are reported of the Insertion Losses of single and parallel barriers calculated using a boundary element solution of a variation of the two dimensional Kirchoff-Helmholtz equation. One advantage of the use of this method is that many different configurations of cross sectional shape and admittance distribution can be considered with equal accuracy. Also the method does not rely on the semi-empirical approach of ray tracing, images and edge diffraction calculations common in many methods of predicting barrier performance.

Drawbacks to the method are that at present the calculations are limited to two dimensions and that the propagating medium is stationary and homogeneous so that atmospheric effects are ignored.

Method of Calculation

A two dimensional system is considered, as shown in Figure 1. S is the source. The cross sectional shape of the barrier is defined by the line \( \gamma \) and an admittance distribution is defined along the line. It is also necessary to assume that the ground is flat and of constant admittance. The wave field in the air and in the surface of the barrier can be expressed as the solution of a developed form of the Kirchoff-Helmholtz integral equation. The solution of the equation to produce the acoustic pressure in the barrier surface, \( \gamma \) can be carried out using boundary element techniques. The surface pressure can then be used to calculate the pressure at any position in the air [1, 2, 3].

![Figure 1. Two dimensional model](image)

The shape of \( \gamma \) and the admittance distribution on \( \gamma \) are provided as data for the computer program as a series of line segments which can describe any form of barrier. By introducing a null segment barriers with apertures or parallel barriers can also be considered. The solution method involves the determination of the surface pressure in \( \gamma \) at intervals of
0.2λ, where λ is the wavelength, so that for long boundaries and high frequencies the computing resources required to solve the problem are considerable.

The method enables the Insertion Losses of single or parallel barriers to be calculated for a monofrequency source. Calculations have also been carried out for a broad band source spectrum characteristic of road traffic noise. Insertion Losses are determined at 1/3 octave band centre frequencies from which the total Insertion Loss for the designated spectrum is then derived approximately.

Results and Discussion

A fundamental check of the application of the theory and the software was carried out by calculating Insertion Losses for a barrier of semi-circular cross section on ground of zero admittance. An accurate analytical series solution is available for scattering from an obstacle of this nature [4]. There was very good agreement between the results.

Results of the present method were also compared with calculations of Insertion Loss for road traffic noise for single and parallel barriers. The published results of Hajek [5] were used. These were calculated using the STAMINA 2.0 traffic noise prediction model, with the introduction of image sources due to reflections at parallel barriers. The source, receiver and barrier geometry is indicated in Figure 2. When the Insertion Losses calculated using the two methods were compared agreement was poor. Insertion Losses calculated using the boundary element method were systematically greater than those predicted by Hajek for both the single barrier and parallel barrier cases. The maximum difference was 7.3dB. Closer agreement was observed when the degradation of Insertion Loss due to the introduction of the second barrier was considered. The results for parallel barrier degradation using the two methods are shown in Figure 2. The boundary element method systematically produces greater values of degradation, but the trends with distance from one of the barriers are similar and the maximum difference between the results is 3dB.

Many fundamental differences exist between the two methods of calculation. The boundary element calculation is two dimensional in the plane of the barrier cross section, so that the source translates into three dimensions as a coherent line source which is parallel to the barrier. A more appropriate description of road traffic is as a line of incoherent point sources. It is expected that the absolute values of Insertion Loss produced by the present model will overestimate results for road traffic noise. However it is likely that the model will be useful in examining the relative performance of different shapes, surface covers and positions of single and parallel barriers.

In Figure 3 the Insertion Losses of three different configurations of parallel noise barriers are compared. The geometries are indicated on Figure 3. To allow comparison the maximum height is 3m in each case and the separation of the highest points is 25m. The cases are a) vertical walls with surfaces of zero admittance; b) as in a) but the inward facing surfaces having finite admittance. c) sloping inner faces at 45° with zero admittance. The admittance of the ground was zero and the source was
Figure 2: Parallel barrier degradation as a function of distance from a barrier (D). The solid line is from Hajek [5]. The points are calculated using the boundary element method.

Figure 3: Insertion Loss for a traffic noise source as a function of distance from one barrier. Source and barrier geometries are indicated. All the surfaces have zero admittance, except the inward facing surfaces in case 2.
positioned in the ground surface halfway between the barriers. The finite admittance used was that predicted by the Delany Bazley formula [6] with a flow resistance of 20,000 (Kg/sm*) and depth of 0.1m. The insertion losses calculated for a road traffic noise spectrum are plotted as a function of distance from one of the barriers at a receiver height of 1.5m. Values of insertion loss peak at a distance of 20m behind the barrier for each case. At 100m the insertion loss for case a) is about 3.5dB, but for cases b) and c) a similar improvement in insertion loss of ≈8.5dB is observed. At 20m the improvements are lower, but still significant at 5dB for case b) and 3dB for case c). Similar plots can be made of the insertion loss for a monofrequency source. For a 100Hz source the values of insertion loss are all considerably reduced from those in Figure 3 for each case, with case c) producing lower insertion losses than case a). At 500Hz the relative positions of the results look similar to those in Figure 3 but at 100m the insertion loss for cases a), b) and c) are 5.0, 9.8 and 9.5 dB respectively. Insertion losses of barriers for traffic noise sources have commonly been estimated using results obtained for monofrequencies of 500 or 550 Hz. At both 500 Hz and 1KHz strong interference effects were observed at receiver positions close to the barrier. The results at 1KHz bore no resemblance to those in Figure 3. At 100m the insertion loss for cases a), b) and c) were 16.1, 14.5 and 19.1dB. At this frequency case b) producing the lowest insertion loss.

Conclusion

The boundary element method has been used to calculate insertion loss for single and parallel barriers. The present two dimensional model is probably not useful for determining absolute values of insertion loss in noise from road traffic, but could be used to examine the relative efficiency of different types of barrier configuration.

References


WIND EFFECT ON OUTDOOR SOUND PROPAGATION:
EXPERIMENTAL STUDY IN WIND TUNNEL

by M. Rosen and Y. Gabillet

Centre Scientifique et Technique du Bâtiment
24 rue Joseph Fourier  38400 Saint Martin d'Hères, France.

1. INTRODUCTION

Outdoor sound propagation is strongly influenced by several factors: meteorological
effects, ground effects, natural or man-made barrier effects. Today, wind effects on acoustic
propagation are well-known qualitatively, but theoretical models are not accurate enough.
The acoustic detection in an inhomogeneous and moving medium or the noise level
calculation behind a barrier built in a perturbed atmosphere make the knowledge of
meteorological effects very important. Some outdoor propagation experiment have been
already carried out [1] [2] [3], but often, the unexpected fluctuations of the wind in amplitude
and direction give some results difficult to understand.

An experiment with a 1/20 scale model in a wind tunnel where all meteorological
parameters are fully determined and controlled, is described. Experimental results concerning
upwind propagation situations are presented in this paper. The influence of the wind
turbulences on the sound level in the shadow zone is shown.

2. EXPERIMENTAL CONDITION AND INSTRUMENTATION

2.1. Wind tunnel description

The great advantage of the CSTB wind tunnel is the closed circuit working. The physical
characteristics of the air especially the flow temperature are perfectly controlled and
regulated within ± 0.5°C. The important dimensions of the tunnel (4 m x 20 m) make possible
the simulation of natural wind at great scales. The height of the ceiling can varied from 1.5 m
to 2.5 m. Low positions are required to obtain high flow speeds (> 30 m/s). An axial engine
(600 tours/min) of 200 kW sets in motion one twelve blade propeller. The flow speed in the
tunnel varies continuously from 0 to 30 m/s within 0.1 m/s. This continuous acceleration is
due to the variable blade stop. This regulation method associated with an acoustic treatment
shifts to 120 Hz, the parasite frequency attached to the propeller siren effect.

2.2. Acoustic measurement method

It is difficult to carry out acoustic measurements in a wind tunnel because of the noise of
the propeller and the parasite rays coming from top and lateral walls of the tunnel.

The experimental method TDS (Time Delay Spectrometry) was used to solve these
problems [4]. T.D.S. has a very good background noise rejection and can identify and isolate
two sound rays arriving from the same source. T.D.S. gives the narrow band frequency
response in amplitude and phase of each identified path.
The scale 1/20 has been chosen for the following reasons:

- the wind tunnel can simulate correctly winds at scales from 1/20 to 1/1000.
- the measurement system T.D.S. is limited to 20 kHz. Then at scale 1/20, the frequency range is 0 Hz to 1000 Hz and representative of the acoustic phenomena to be studied.
- the geometric dimensions of the wind tunnel impose to the source-receivers distances to be smaller than 5.5 m (110 m full scale) to be able to separate rays coming from top and lateral walls with T.D.S.

The sound source used is a small tweeter which has a flat response from 2 kHz to 20 kHz corresponding to 100 Hz – 1000 Hz full scale. This source is small enough and careen spherically to avoid aerodynamics perturbations. This tweeter gives 105 dBA at 40 cm to be able to drown the propeller noise. Measurements were made with three kinds of absorbive grounds having a flow resistivity of $100.10^3$ MKS (sand), $300.10^3$ MKS (grass), $10000.10^3$ MKS (concrete) respectively. Grass and sand were simulated by thin pieces of cloth material placed on a reflected support (0.6 mm for grass, 1.4 mm for sand). Different kinds of representative winds were simulated in the wind tunnel: winds of -5 m/s and -10 m/s with different vertical distributions of turbulence. The measurements of wind gradients, turbulence profiles were made with hot wire anemometers.

3. EXPERIMENTAL RESULTS

For each configuration, several measurements were carried out (between 20 and 50) to get a mean value which is characteristic. The source receiver distance is 80, 95 or 110 meters. The limite of the shadow zone is about 40 m from the source. This limit has been calculated with a ray path computer program. Two different turbulence profiles have been simulated in the wind tunnel. One of these profile consists in a wind with 5% of turbulence at 10 m height, the other in a wind with 15% of turbulence at 10 m height. All presented results are normalized to free field at 3 meters (model scale). This reference is chosen to be able to isolate the direct path from rays coming from the ceiling, the ground, and the lateral walls. All figures illustrate full scale situations.

Figure 1 show the fluctuations of the sound pressure level obtained with a high turbulent wind of -10 m/s at a distance of 95 m from the source above grass. In this case, 30 consecutive measurements have been carried out. The variations of the pressure level are about ± 10 dB around the mean value and can affect strongly the acoustic level.

Figure 2 show the comparison between upwind propagation with a wind of -5 m/s and a wind of -10 m/s.

Rays are bended strongly to the tunnel ceiling with the increasing of the wind speed. With the wind of -10 m/s, the pressure level is decreased at the receiver from 5 dB in the low frequencies to 12 dB in the high frequencies compared with the wind of -5 m/s. The profiles of turbulence are the same for the wind of -5 m/s and the wind of -10 m/s.

Figure 3 show a comparison between upwind propagation with winds of -10 m/s differentiated by their profiles of turbulence. The increasing of the pressure level at the receiver in the case of high turbulences is important (5 dB between 200 and 400 Hz) particular in the frequency range 800 to 1000 Hz.

This increasing effect can be explained by a best penetration of the shadow zone due to the turbulences.
4. CONCLUSION

The knowledge of meteorological effects is important in problems of noise detection or noise estimation. From this experimental work, a better understanding of acoustic propagation in inhomogeneous and moving medium is gained. This study points out the great influence of the turbulences in upwind propagation cases. A computer model is actually developed to predict these effects. As the ray theory fails, for calculation in the shadow zone, an hybrid method of superposition of gaussian beams will be used.

5. ACKNOWLEDGEMENTS

This study has been supported by Direction des Recherches Etudes et Techniques (D.R.E.T.).

6. BIBLIOGRAPHY

(1) CANARD (S.), LEWY (S.), VERMOREL (J.), PARMENTIER (G.)
"Long range sound propagation near the ground."
Proceedings INTER-NOISE 88.

(2) OSWALD (G.P.)
"Propagation acoustique extérieure – Validation du programme PAPE"
Rapport EDF-DER HE/22/4827 (1965)

(3) PARKIN (P.H.), SCHOLES (W.E.)
"The horizontal propagation of sound from a jet engine close to the ground at HATFIELD"
J.S.V., 2(1965) pp. 353-374

(4) VILLOT (M.), DEGRYSE (D.)
"Acoustical measurement in a wind tunnel using Time Delay Spectrometry"
Proceedings INTER-NOISE 88

![Diagram](image)

Figure 1: Pressure level fluctuations normalized to free field at 60 m for a wind of -10 m/s.
Figure 2: Comparison between measured pressure levels normalized to free field at 60 m for
a wind of −5 m/s and a wind of −10 m/s.

Figure 3: Comparison between measured pressure levels normalized to free field for a low
turbulent wind and a high turbulent wind.
COMPUTER SIMULATION OF OUTDOOR SOUND PROPAGATION

Karsten B. Rasmussen

The Acoustics Laboratory
Technical University of Denmark
DK-2800 Lyngby, Denmark

INTRODUCTION
In connection with community noise problems the use of practical calculation procedures are becoming increasingly popular. The development of complex and accurate calculation procedures is useful as a basis for the development of practical calculation procedures. The present work is aimed at simulating outdoor sound propagation over ground taking the influence from wind and temperature gradients into account. The mathematical approach is a direct implementation of the solution of the wave equation in the Hankel-transformed domain. The atmosphere is treated as a horizontally stratified medium. The atmospheric layers may each have a sound speed gradient. The problem formulation and computer implementation is based on a direct matrix formulation of the boundary conditions between all layers [1] and numerical stability problems present in other implementations are avoided. The results are compared with measured data.

THEORY

\[ p(r, z) = \int_0^\infty K_0(s, z) J_0(sr) s ds \]  \hspace{1cm} (1)

where the kernel \( K_0 \) is the solution to the Hankel transformed homogeneous wave equation, since equation (1) is in fact an inverse Hankel transformation. The kernel is given by,

\[ K_0(s, z) = A^-(s) e^{-s \alpha(s)} + A^+(s) e^{+s \alpha(s)} \]  \hspace{1cm} (2)

where \( A^+ \) and \( A^- \) are determined according to the boundary conditions between the layers. \( \alpha \) is given by

\[ \alpha(s) = (s^2 - k^2)^{1/2} \]  \hspace{1cm} (3)

Figure 1. Horizontally stratified medium. Layer one is ground.

For a horizontally layered structure as shown in Figure 1 the acoustic pressure \( p \) may in a layer without sound speed gradient be shown to be given by [1],
where $k$ is the wavenumber for the layer in question. The wavenumber is complex due to the losses in the medium. It should be specified that the square root in equation (3) is understood to have positive real part. From the Hankel transform it is seen that $s$ is the horizontal wavenumber component and that in the transformed domain the coefficients $A'$ and $A''$ do not depend on $z$. The vertical displacement component $w$ is found from differentiation of pressure with respect to the $z$-coordinate.

Equation (1) is for the source-free field. Including the source field (source located on $z$-axis) is done by means of the relation for a point source,

$$ p(r,z) = \int_0^{\infty} K_p J_0(sr) s ds $$

where the kernel is given by,

$$ K_p = \frac{e^{-|z-z_s|/\alpha}}{\alpha} $$

where $z_s$ is the source height. The star denotes source contribution. The field given by equation (4) is equal to $e^{ikR}/R$ which is the appropriate expression for a point source when the $e^{-ikt}$ time convention is used.

In each layer the kernels given in equation (1) are defined separately, and $s$ as well as $A'$ and $A''$ are layer dependant parameters, which are more accurately written $\sigma_n$, $\alpha_n^+$ and $\alpha_n^-$, where $n$ is the layer number.

Finding $\alpha_n^+$ and $\alpha_n^-$ as a function of $s$ is the central part of the field calculation. The integration from zero to infinity is done numerically and the necessary integration interval is limited, since the kernel approaches zero for large and small values of the horizontal wavenumber. For each $s$-value used in the numerical integration, the value of the $A$-parameters are found by means of solving $2N$-2 linear equations, where $N$ is the number of layers, see Figure 1. The linear equations are determined from continuity of $w$ and $p$ at the boundaries. Equations are formulated for each layer-boundary on the basis of the kernel (2) and the corresponding kernel for vertical displacement. For the boundaries of the layer containing the source, contributions from the kernel given by the equation (5) and the equivalent kernel for vertical displacement are added.

Hence, each layer $n$ gives rise to equations for pressure $p$ and vertical displacement $w$,

$$ K_{w,n,u} - K_{w,n+1,u} = K_{w,n,u} - K_{w,n+1,1} $$

$$ K_{p,n,u} - K_{p,n+1,u} = K_{p,n,u} - K_{p,n+1,1} $$

where the subscripts $n$ and $n+1$ denote the layer number, whereas the subscripts 1 and $u$, denote lower and upper interface respectively for the layer. The source contributions on the right hand side will be zero except in the source layer.

It is an advantage to scale the equations (6) and (7) in order to obtain coefficients of the same order of magnitude [1].
A local z-coordinate is introduced in each layer, in such a way that z is zero on the lower interface. This choice saves computation time, since exponential functions having zero argument may be left out. Furthermore, the choice of local z-coordinate is a part of the precautions ensuring numerical stability of the equations system. If the equations are arranged in a matrix structure where each interface has two rows (one for displacement and one for pressure) and each layer has two columns (corresponding to \(A^-\) and \(A^+\), respectively) then a band matrix emerges, which may readily be solved by means of Gaussian elimination. The \(A^-\) and \(A^+\) values for the layer containing the observation point are used to determine the contribution to the pressure integral, eqs. (1) and (2).

Up until this point the stratification consists solely of homogeneous layers. It is possible to introduce layers having a linear sound speed variation as a function of the vertical z-coordinate. The solution is based upon the Airy function.

The wavenumber in a layer having a velocity gradient is given by

\[
k(z) = \frac{\omega}{c_0(1+i\tau z)}
\]

leading to (valid for \(|r_z|<<1\))

\[
k^2(z) = k_0^2(1-2r_z), \quad k_0 = \frac{\omega}{c_0}
\]

Inserting the approximation used in equation (8) in the wave equation leads to Airy's differential equation, and hence to Airy function solutions. The solution may be stated [2],

\[
p(x,z) = \int_0^\infty K_0 J_0(\pi t) \text{d}t
\]

where

\[
K_0 = A^-(z) \cdot V(\eta-y) + A^+(z) \cdot \bar{W}(\eta+y),
\]

\[
\eta = \frac{z}{2} \frac{c_0}{\omega}, \quad \text{and} \quad V(\eta) = \frac{\sin(\eta)}{\eta}, \quad \bar{W}(\eta) = 2\eta^{1/3} \cdot \text{Ai}(\eta e^{2\pi/3})
\]

Calculating the field for a problem involving one or more gradient layers means setting up the linear equations as previously described, but in gradient layers the kernels are expressed by eq. (10) and the equivalent equation for vertical displacement. Due to the assumption that the source is always located in a homogeneous layer, the source contribution is always determined from equation (4). When gradient layers are present, the problem should always be formulated so that the source height is greater than the receiver height. When this condition is not met, the principle of reciprocity is employed, which is equivalent to an interchange of source and receiver heights for the situation given in Figure 1.

The ground layer is described as a fluid layer of infinite depth and having heavy losses. The complex density and wavenumber is given by the Delany and Bazley model [6] for porous materials.

Figure 2 shows calculated and measured data for downwind propagation (positive sound speed gradient) over plane grass, and Figure 3 shows similar results after having interchanged source and receiver locations (upwind/negative sound speed gradient). The grass is specified by a flow resistance of \(200 \cdot 10^3 \text{Ns}^{-1} \text{m}^{-2}\).
Fig. 2. Sound pressure level relative to free field. Downwind propagation. Distance 80 m. Source height 1.42 m. Observation point 0.5 m. ————, measured; ·······, predicted for linear sound speed variation; ————, predicted for two-segment linear variation. Wind speed 2-2.5 m/s at a height of 10 m.

Fig. 3. Equivalent to Fig. 2 except for upwind propagation.

REFERENCES
2. K.B. Rasmussen and B. Vistisen 1986 Report 96.149, Ørsted & Danneskiold-Samsøe Kroghsgade 1, DK-2100 Copenhagen, Denmark (in Danish).
IMPULSE SOUND MEASUREMENTS FOR CHECKING THE METEO WINDOW

U.J. Kurze
Müller-BBM GmbH, Robert-Koch-Str. 11, D-8033 Planegg

1. Introduction

The assessment of industrial noise impact is often related to weather conditions favourable for sound propagation from the industry to the sensitive neighbourhood. Sound measurements must be carried out in a meteo window which may be rather restrictive with respect to wind direction and velocity in order to make sure that the essential wind and temperature gradients stay within certain limits (s. Fig. 1).

![Diagram of wind component and source-receiver setup.](image)

**Fig. 1** Meteo windows proposed in VDI 3745 (a) and by Nordtest (b) for source and receiver heights $h_s + h_R < 0.1 d$ /1/.

Instead of setting up meteorological equipment for monitoring the weather conditions at a few points in the field, an impulsive sound source has been used in the past in connection with field measurements of railroad barrier performance /2/. The calibrated source positioned aside from the

![Graph showing receiver height and zero wind.](image)

**Fig. 2** 10-min average SPL received at a distance of 408 m; calculated from DIN 16005 /3/: 80.3 dB.
end of the barrier allowed for determining the free field sound propagation conditions all along the propagation path from the railroad line to the receiver up to a distance of 400 m. As shown by the example in Fig 2, the maximum A-weighted impulse sound pressure level received under early morning inversion conditions may exceed the level predicted from standard calculation procedures by a few dB. But when the sun rises, the observed level drops unless there is some downwind as it happened at 7:45. Based on this type of experience, a research project was supported by the Bavarian Environmental Protection Agency (LJU), aiming at more detailed investigation of sources, procedures and field applications /4/.

2. Sources

The special impulse sound source already used in previous investigations consists of a tube with a plastic membrane at its top end. The tube is filled with pressurized air until the membrane bursts. The diameter of 100 mm and the membrane material and thickness of 0.2 mm determine a maximum sound pressure in the frequency range of 1 kHz, while the tube length of 500 mm provides for an enhanced low frequency radiation around 160 Hz. In an anechoic room, the sound energy output was found rather uniformly distributed in the frequency range from 63 to 4000 Hz and in the angular range from -30° to 30° relative to the plane of the membrane.

When used outdoors 2 m above a grass field, the frequency characteristic of the SPL $L_{F_{\text{max}}}$ received at a distance of 25 m and 5 m above the ground already exhibits some effects of ground interaction in the frequency bands of 250 and 630 Hz (see Fig. 3). The levels are roughly 6 dB higher than those from a 9 mm pistol fired overhead and about 20 dB higher than the energy mean levels received from a 100 W vented loudspeaker box (a cubicle with a single speaker of 0.3 m diameter).
3. Measurements

A measuring line was set up on a former airfield which was partly covered with grass and gravel and partly cultivated with cereal plants. Two dual-channel realtime analyzers were used for simultaneous measurements at different distances up to 800 m and heights of 5 and 10 m. Results for the level \( L_{F_{\text{max}}} \) from pistol shots are shown in Fig. 4 in terms of average values \( L_{F_{\text{max}}} \) from ten repeated shots fired in a 5 min interval, energy mean values \( L_{\text{eq}} \), minimum and maximum values, and standard deviation \( s_\text{L} \) of SPL \( L_{F_{\text{max}}} \). Also plotted is the background noise level due to remote traffic and wind at the microphone, which was protected by a foam ball only.

During a period of half a year, the measuring line was set up on seven days always in the direction of the wind. Two series of measurements have been carried out each for zero wind during early morning inversions, light wind (0 - 2 m/s) under overcast sky, moderate wind (3 - 6 m/s) with bright skies, and stronger or variable winds on cloudy days.

![Fig. 4](image)

4. Evaluation

From the various aspects for evaluation of the large amount of data, the comparison of impulse and continuous sound measurements and the frequency characteristic of ground effects is highlighted in Fig. 5. Obviously, the impulse method - here applied in terms of \( L_{F_{\text{max}}} \) - in order to integrate over an effective time interval of 125 ms - is consistent with standard procedures for transmission loss measurements close to the ground. In addition, Fig. 5 shows typical differences between a simplified prediction scheme /5/ which just roughly accounts for the source and receiver height in terms of an excess ground attenuation in the range from 0 to 5 dB, but not for its typical frequency dependence with a maximum value between 200 and 500 Hz.
The suitability of a reference sound source for long range sound transmission measurements is limited at low frequencies (below 250 Hz) by environmental noise and (below 500 Hz) due to strong effects of ground interaction. At high frequencies (above 2500 Hz), air absorption is substantial. Consequently checking of the existing meteor condition is preferably done in the frequency band around 1 kHz or by evaluating A-weighted overall levels.

The standard deviation of 1/3-octave band levels $L_{F_{\text{max}}}$ increases with increasing distance and turbulence of the atmosphere and typically reaches 6 dB. It is just slightly lower for A-weighted overall levels from pistol shots. Consequently, a number (about 10) of repeated measurements is required for deriving a reasonable accuracy for the average level (error margin about ±2.5 dB at a confidence level of 80%). From measurements on different days, the standard deviation of averaged octave-band levels was determined for various conditions within the usual meteor window. The values generally do not exceed 4 dB for the 1-kHz band or for the A-weighted overall levels of the impulsive sound sources employed.

5. Conclusions

Pistol shots with a reference level $L_{AF_{\text{max}}} \approx 100$ dB at 25 m are particularly useful for checking doubtful meteor conditions over distances up to 800 m during light wind and sunrise. In cases of high background noise or for safety reasons, the impulse source with the bursting membrane may be preferred. Repeated measurements of the level $L_{AF_{\text{max}}}$ allow for reliable discrimination of situations inside and outside of the meteor window.

References:

/1/ H.G. Jonasson, Report on Nurdtest Workshop in Boras, 8 - 9 June, 1988
/2/ U.J. Kurze, Inter-Noise 87, 379 - 382
/3/ DIN 18005, Part 1, May 1987
/5/ VDI 2714, Outdoor sound propagation, January 1988
THE INFLUENCE OF METEOROLOGY ON GRAZING INCIDENT SOUND

J. Wempen, Fachbereich Physik, Universität Oldenburg, D-2900 Oldenburg, FR Germany

The prediction of sound propagation outdoors often has to deal with the case of nearly grazing incidence. The ground influence for this case is described with the theory of spherical wave propagation [1] by taking into account the impedance of the specific grounds. The impedance can sensitively be determined by measurements of grazing sound propagation along those grounds [2], [3] where the grazing incidence is approached by very low source and receiver positions. The influence of meteorology on grazing incident sound is a phenomenon not clarified until now. Stable sound speed gradients above outdoor grounds, i.e. gradients of wind speed and in particular temperature, are well known to establish in opposite ways during day and night [6]. It is shown, that the determination of ground impedances is very sensitive to meteorological effects. Sound speed gradients above the ground have a strong influence on sound propagation, even over small distances of a few decameters.

Measurements

The measured functions of excess attenuation shown in the figures were gained with a correlation technique [5]. It delivers confident measurements of magnitude and phase of the excess attenuation relative spherical spreading in short averaging time.

A sunny winter day was chosen, at which a strong variation of the temperature gradient during the day was expected. The transmission of sound from a source, height above ground 40cm, to a receiver in the same elevation was measured in different distances twice that day, the first time at noon, the second time short before sunset. The air temperature was monitored to assure a positive and negative temperature gradient during the corresponding measurements.

Influence of varying wind

Slow winds, varying within seconds from zero up to about 2 m/s were present during the measurements. Even in a calm atmosphere, there are always slightly varying wind and temperature fields, which scatter high frequent sound, yielding well audible changes in the sounding of a ground near source.

Fig. 1 shows the influence of alternating winds on the sound transmission over 100 m. Three measurements are depicted, taken consecutively for a constant geometry. The time between the measurements is of the order 10 seconds, a single measurement is an average over a few (about 4) seconds. A mean value of the excess attenuation is simply found by a longer averaging time.

Low frequencies are hardly affected by the wind, the excess attenuation is constant up to the ground dip. The main effect of the wind is an increase or decrease of the path length difference between direct and reflected waves, which is responsible for the reincrease of the excess attenuation towards high frequencies. These variations can easily be averaged out, but a much stronger effect is shown to arise from the different temperature gradients during daytime and nighttime.

Influence of temperature gradients

Fig. 2 shows the influence of the different temperature gradients during the two measurements at noon (A) and towards evening (E) for four distances. The measurements over 12.5 m and 25 m show a similar behaviour of low frequencies, the excess attenuation approaches +6 dB, indicating, that the waves are transmitted to the receiver without losses. The different behaviour of the measured magnitudes above 1 kHz is again due to variations in the path length difference...
between direct and ground reflected waves, originating from constant temperature gradients in this case.

The differences in excess attenuation between day- and nighttime increase dramatically with increasing distance. During day conditions (Curves A), the ground gets more and more as a high order low-pass filter with increasing distance, no waves of high frequency are transferred to the low receiver position at 100 m at all, the effective path length difference rapidly decreases.

During evening conditions (Curves η), things have widely changed. The path length difference increases with increasing distance (instead of decreasing linearly), yielding a strong reincrease of the excess attenuation functions towards high frequencies. The sound pressure level of low frequencies exceeds +6 dB relative free propagation of a spherical wave for distances of 50 m and more.

A qualitative explanation of the measured effects can be given with the aid of the image of sound rays, being bended downward through positive gradients, i.e. temperature or wind speed increasing with height (nighttime, downwind), and upward during the day, forming a shadow region [4]. Positive gradients yield an increase, negative ones a decrease of the path length differences, as it is confirmed by the measurements [5]. The shadow region is penetrated by waves of low frequency, as long as their wavelength is large compared to the spatial dimensions of the gradients. The increase of the sound pressure level to more than +6 dB is explained by a focusing effect of the gradients. These two last statements can not be described in a quantitatively correct manner with a ray tracing method, because it does not account for the wave nature of sound.

The measured behaviour of low frequencies is more correctly described with the theory of spherical wave propagation, predicting a ground wave coupled to the ground surface and travelling with a phase speed smaller than free waves. The phases of the measured excess
Fig. 2 Measured functions of excess attenuation during noon (N) and towards evening (E). The source and receiver height was 0.40 m, the horizontal separation is noted.
attenuation functions show a strong decrease with frequency up to the ground dip, exceeding 720° at 100 m. This decrease is not due to a simple time delay, as can be seen from the nearly constant behaviour in the high frequency domain. The phase functions are normalized to free spherical propagation, this means, that high frequencies behave like spherical waves while low frequencies are relatively delayed. The decrease of the phase lag of the low frequent ground wave part at all distances is constant, if it is depicted versus the square root of the distance in wavelengths [7]. The phase term of the ground wave is determined to roughly $e^{\frac{1}{2} \frac{\sqrt{r}}{r}}$ for the noon measurement and $e^{\frac{1}{2} \frac{\sqrt{r}}{r}}$ in the evening. This result is confirmed by the theory of spherical wave propagation only if the numerical distance is small [2], predicting in this case a propagation constant of the ground wave, which depends on the square root of the distance in wavelengths, on the effective admittance of the surface and on the grazing angle. The measurements indicate no angle dependence while the differences between the day- and evening measurement indicate an effective surface admittance, which obviously depends on meteorology.

As a consequence, an attempt to fit the measurements for both meteorological situations with the same ground admittance, like it is easily possible for any single measurement, fails. The reincrease of the measured magnitudes allows for a calculation of the effective path length difference, but the whole excess attenuation function, especially the cut-off of the ground wave part, can not be described with a variation of geometrical parameters alone.

Conclusion

The meteorological circumstances during propagation of sound outdoors are capable of strongly increasing or decreasing the immission of a source at a given receiver position. In terms of the simple interference model of spherical wave propagation, which does not include meteorology, the effective impedance of the grass surface is always affected by the meteorological circumstances. Measurements of ground impedance close to outdoor surfaces, where, of course, the sound speed gradients are strongest, have to take them into consideration.

IMPULSIVE SOUND SOURCE OF HIGH INTENSITY FOR OUTDOOR SOUND PROPAGATION MEASUREMENTS

Radek, U.; Klug, H.; Meller, V.
University of Oldenburg, P.O.B. 2503, FB 8, D-2900 Oldenburg

1. Impulsive sound source
The impulsive sound source described here consists of a capacitor (100 μF, 3.5 kV, max. 612 Ws), which is discharged over a spark gap. A short reproducible pulse is produced (duration 0.5 ms, fluctuations 0.5 dB) and spherically radiated with a maximal sound pressure level of 150 dB at a distance of 1 m. The spectrum has its maximum about 1.4 kHz.

Some essential details concerning the mechanism of this type of sound source are described in the following chapters.

1.1 Ignition
A special configuration of the ignition system is necessary for maximal source output. An additional electrode is positioned between the main electrodes of the spark gap (fig. 1). This auxiliary electrode and one of the main electrodes (zero potential) are connected to the secondary winding of the ignition coil. The ignition spark ionizes one half of the spark gap (i.e., the part between auxiliary and negative main electrode). This part of the spark gap becomes conductive originating a plasma. Free electrons of this plasma are accelerated by the electric field to the other main electrode (zero potential) and ionize by impact (avalanche breakthrough) the second part of the spark gap. The primary winding of the coil has to be connected to the proper terminals. In this way a spark gap with a length of 12 mm can be ignited by a capacitor voltage of 3.5 kV. Otherwise configurations only allow shortenings over a gap of a few millimeters.

1.2 Dimensioning

![Fig 1: Equivalent circuit diagram of the sound source including ignition system](image)

Capacitance, inductance and resistance within the circuit form an electrical resonant circuit, additionally influenced by the time dependent resistance of the plasma. After ignition the plasma resistance and therefore the total resistance of the circuit will become small within a very short period of time and the resonant circuit will execute a damped oscillation. The occurring high currents (a few kA!) require an appropriate type of capacitor. The capacitor has to be resistant to impulsive discharging.

For maximal energy dissipation in the plasma and a good reproducibility the total resistance within the circuit has to be small compared to the resistance of the plasma. In first order the plasma resistance is proportional to the length of the spark channel (i.e., distance between the main electrodes) which can be maximally chosen in case of optimized ignition as described above. Melting loss and the shape of the electrodes have no influence on reproducibility.

Capacitor and electrical connections have inductances. Although inductances have losses, it can be advantageous (see below) to decrease the resonant frequency by increasing inductances, e.g., lengthening the connections between capacitor and electrodes.
1.3 Spectrum

In case of high resonant frequency (Rayleigh's estimation) the total capacitor energy will be dissipated instantaneously, i.e., before the air volume heated by the spark expands. Thus the following expansion is a reactive process. The radiation of a sound wave takes place, if the pressure at the surface of the volume drops below a well defined value $P_0$, which is independent of the energy dissipated in the heated volume. $P_0$ and the pressure distribution within the volume determine the radius $R_0$ of the sphere at which the pulse is radiated.

The resulting time function is a "N"-shaped wave. At low frequencies the spectrum increases and at high frequencies the spectrum decreases linearly with frequency (6 dB per octave). An interference pattern is caused by the first and second steps in the time signal.

The interference minima give rise to problems if the signal is used as input for the calculation of transfer functions. They can be diminished by extending the time of energy dissipation, i.e., by reducing the resonant frequency. The low pressure domain of the pulse is smoothed, the case of an "explosive wave" is approximated.

At low resonant frequency the energy is not dissipated immediately. The time function of the electric power is in first order proportional to the squared sinus function of the resonant oscillation. i.e., energy is dissipated in several pulses with decaying amplitude. Thus the spatial distribution inside the expanding volume is not constant. The resulting sound wave (fig. 2) consists of a series of overlapping pulses nearly forming an explosion wave. The first two steps have the highest amplitudes. At the beginning they propagate with an accelerated sound speed caused by nonlinearity resulting in a time delay of the following steps. Nonlinearity vanishes at a distance of some meters. The interference of the single steps always results in an irregular pattern of more or less distinct dents in the decaying part of the spectrum.

An increase of the dissipated energy yields a parallel shift of the slope at lower frequencies, while the high frequencies are hardly effected (fig. 3).

![Fig. 2: Time signal of a sound pulse in a distance of 3 m, 3.5 kV.](image1)

![Fig. 3: Spectra of sound pulses at different capacitor voltages: 1.75 kV, 2.47 kV and 3.5 kV.](image2)

![Fig. 4: Time signal and spectrum of a pulse steepened in a tube, duration 0.8 ms, electric power 1.4 Ws.](image3)
A nearly flat spectrum is achieved, if the spark gap is situated in a small tube. The produced sound wave steepens when propagating through the tube and alters to a very short pulse (Fig. 4). Otherwise the acoustic power is lowered by the absorption in the tube.\textsuperscript{1/}

The source has an omnidirectional pattern up to 3 kHz (Fig. 5).

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{sound_pressure_level.png}
\caption{Spectra of a sound pulse recorded at three distinct directions.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{time_signal.png}
\caption{Time signal of the direct and the reflected pulse with time windows added.}
\end{figure}

2. Applications of an impulsive sound source outdoors

In sound propagation, sound source and receiver can be regarded as input and output of a transfer channel. This system is influenced by the following parameters:

- **Geometry:** given by the height of source and receiver, horizontal separation and contour of the surface (barriers).
- **Reflection coefficient of the surface:** given by the physical properties of ground (structure, humidity etc.) and vegetation.
- **Distribution of the acoustical refraction index:** given by the field of temperature and wind velocity.
- **Sound absorption and extinction:** influenced by humidity, temperature and turbulence.

In a given situation all of these parameters can influence the sound pressure level. They fluctuate in time and space.

The transfer channel is completely characterized by its transfer function or its impulse response. Transfer function and impulse response are Fourier transforms. Thus measurements with impulses are equivalent to measurements with continuous signals (e.g. correlation measurements with pseudorandom noise).

In the simplest case, the impulse response consists of the direct and ground reflected impulse. It can be detected by using a sufficiently short sound pulse with a known spectrum and sufficient intensity in the spectral area of interest.

Two different sound paths exist. A sound source with omnidirectional pattern provides the same input for each path.

Under normal environmental conditions the signal to noise ratio is sufficient for broadband measurements of sound propagation up to a distance of 1 km.
2.1 Measurements of channel parameters analysing particular sound paths

If geometry allows to separate these two pulses (fig. 6), it is possible to measure the following parameters:

**Air absorption** (in a non-turbulent medium with vanishing sound speed gradient): the signal is multiplied by a time window which suppresses the reflection from the ground. The frequency dependent absorption coefficient can be calculated from transfer function between a reference signal and the direct signal.

**Reflection coefficient of the surface**: the reflection coefficient can be derived from the transfer function between direct and reflected pulse.

**Acoustical refraction index**: The acoustical refraction index is mainly dependent on height over ground. The time difference between direct and (first) reflected pulse yields the sound speed gradients without measuring temperature and wind speed gradient. Each pulse allows for detection of the instantaneous sound propagation condition in the channel. /2/

Examples for application in more complex situations:

**Measurement of shielding effect behind a barrier** (fig. 7) The sound pressure level is dominated by reflections at surrounding buildings. The signal diffracted by the barrier can be analysed separately.

![Fig. 7: Time signal behind a barrier. The first two pulses are diffracted by the barrier. (The second is additionally reflected at the ground). A lot of reflections appear a few hundred milliseconds later. They dominate the sound pressure level.](image)

**Measurement of spatial correlation length**: correlation length of spatial structures formed by the diffraction pattern in the transfer channel could be obtained by means of a microphone array, because turbulent structures are "frozen" for the individual pulse propagation. No time averaging is necessary.


ACOUSTIC PROBING OF METEOROLOGICAL AND ACOUSTICAL PARAMETERS IN OUTDOOR SOUND PROPAGATION

Mellert, V.; Klug, H.; Radek, U.
University of Oldenburg, P.O.B. 2503, FB 8, D-2900 Oldenburg

I. INTRODUCTION

The propagation of sound over long distances is influenced by meteorological and ground effects. Sound speed gradients due to wind and temperature gradients cause refraction of the sound waves yielding variations of measured sound pressure levels. These variations are measured outdoors using an impulse sound source /1/. Wind speed and temperature are monitored in several heights. It is shown, that sound speed gradients resulting from the meteorological measurements are equal to sound speed gradients directly calculated from the acoustical measurements in the way described in the present paper. The measured meteorological effects on sound propagation are illustrated by ray tracing methods.

II. PROFILES OF THE SOUND SPEED IN THE SURFACE LAYER

For the nearly neutral case (small or negligible temperature gradients) the sound speed profile can be described by:

\[ c(z) = c_0 + a \ln(z/z_0) \]

where \( c_0 \) is the sound speed at the small height \( z_0 \) (roughness length) above the ground and \( a = (u^*/k)\cos f \) (\( u^* \) friction velocity, \( k \) von Karmans constant \( \approx 0.4 \), \( f \) angle between wind direction and measuring direction).

The sound speed gradient at the height \( z \) is then determined only by the parameter \( a \):

\[ \frac{dc}{dz} = \frac{a}{z} \]

In the case of temperature gradients the profiles can be described by the Monin-Obukhov-similarity functions /2/ and we get for the stable case /3,4/ (positive temperature gradient):

\[ \frac{dc}{dz} = \frac{(a^*/z)(1 + 5z/L)}{1} \]

where \( a^* = (u^*/k)\cos f, 0.6(T^*/k) \), \( T^* \) Temperature scale, \( L \) Monin-Obukhov-Length \( L = (u^*/T^*)(T^*/gk) \), \( g \) acceleration due to gravity (9.81 m/s²).

III. DETERMINATION OF THE SOUND SPEED GRADIENT FROM ACOUSTICAL MEASUREMENTS

Fig. 1 shows the measured time signal in 50 m distance from the impulse sound source, which was mounted 30 m above the ground. The direct pulse is followed by the ground reflected pulse with a delay of about 8 ms. The time difference \( \Delta t \) resulting from geometry for the homogenous case (no sound speed gradient, \( c(z) = const. \)) is:

\[ \Delta t = \frac{\sqrt{(h_s + h_r)^2 + D^2} - D}{c} \]

\( c \): sound speed, \( D \): distance, \( h_s \): source height, \( h_r \):receiver height

Fig. 2. shows a further measurement, where the time difference \( \Delta t \) between direct and ground reflected pulse is 5.2 ms. The time difference expected from geometry would only be 3.7 ms. The additional time difference \( \Delta t = 1.5 \) ms is due to the different sound speeds for the direct ray path and the reflected ray path. A sound speed increasing with height leads to an additional time difference, which allows the separation of both signals in the time domain also for geometries, for which it would not be possible for the homogenous case.

Assuming a logarithmic sound speed profile, it is possible to determine the parameter \( a \) from the time difference \( \Delta t \). At the height \( z' \) (source and receiver height) the direct pulse travels with the sound speed \( c(z') = c_0 + \ln(z'/z_0) \).
The ground reflected pulse has an average sound speed which can be determined by

\[
\frac{c_{Rf1}}{c_0} = \frac{1}{\frac{c(x')dx}{x' - z_0}} = c_0 + \frac{a \sin(z(x') - z_0) x'}{x' - z_0} - c_0 + \frac{z(x')}{z_0} = c(x') - a \quad (5)
\]

Therefore the average sound speed difference \(\Delta c\) between both paths is equal to the parameter \(a\). It is assumed that the actual ray paths due to refraction are close to the geometrical paths. The sound speed difference expressed by the time difference \(dt\) is

\[
\Delta c = a = \frac{c_{Dir} - c_{Ref1}}{t_{Ref1}} = \frac{D}{T_{Dir} + dt} - \frac{D + dt - T_{Dir} + dt}{T_{Dir} + dt} - \frac{c_{Dir}(dt + dt)}{T_{Dir}} = \frac{c_{Dir} dt}{D} \quad (6)
\]

The sound speed \(c_{Dir}\) can be estimated or put 340 m/s for the most cases. Calculating the average sound speed of the reflected pulse (eq. 5) in order to compare it with travel times (eq. 6) is possible in this simple manner, because the variation in sound speed is small compared to \(c_0\). For the measurement in Fig 2 (\(D=230\) m) the time difference \(dt=1.5\) ms leads (eq. 6) to a parameter \(a=0.69\) m/s yielding a sound speed gradient over height \(dc/dz=0.69\) m/s, plotted in Fig 3.

The sound speed gradient determined by simultaneous measurements of temperature and wind speed at several heights, estimating \(u^*\) and \(T^*\) by a method described in /4,5/, leads almost to the same profile (see Fig.3).

---

\(Fig. 1:\) Measured time signal (\(h_c, 30\) m, \(h_r, 5\) m, distance: 30 m)  
\(Fig. 2:\) Measured time signal (\(h_c, h_r, 12.5\) m, distance: 230 m)

\(Fig. 3:\) Sound speed gradient (1/s) over height  
a) derived from wind and temperature measurements (\(u^*\): 0.19 m/s, \(T^*\): 0.64 °C, \(L\): 88 m)  
b) derived from acoustical measurements (\(a\): 0.69 m/s)
**Fig. 4a:** Measured time signal
\( h_0, h_1: 6 \text{ m, distance: 500 m} \)

**Fig. 4b:** Measured time signal
\( h_0, h_1: 6 \text{ m, distance: 500 m} \)

**Fig. 4c:** Magnitude of the autocorrelation function of signals 4a and 4b allows the determination of the time difference excluding the influence of phase changes due to ground reflection.

**Fig. 5:** Ray pattern plots

**Fig. 6:** SPL re free field
Solid line: measurement, see Fig. 4a
Broken line: measurement, see Fig. 4b
IV. THE INFLUENCE OF SOUND SPEED GRADIENTS ON SOUND PROPAGATION

For a geometry of $h_s = h_r = 6m$, $D = 500m$ time differences of $dt^H = 0.7ms$ and $1.5ms$ were determined from acoustic measurements at the same site (flat grassland), but at different times in the late afternoon on a November day (see Figs. 4 a,b,c). This leads to parameters $a = 0.15$ and $0.27$, thus the sound speed gradient (eq. (2)) almost being doubled for the second measurement (one hour later, at 5 pm). Fig. 5 shows ray pattern plots for these measurements (both sound speed gradients compared to the homogenous case). It can be seen that for a logarithmic sound speed profile the density of the reflected rays is increasing with increasing sound speed gradients, while the density of the direct rays is decreasing. The direct sound rays are converging (focusing) and the reflected rays are diverging compared to the homogenous case. That means that the intensity of the ground reflected pulse should be higher than the intensity of the direct pulse (for hard grounds). Even measurements over grassland show that for a sound speed gradient of $dc/dz = (0.27 m/s)/z$ the intensity of the reflected pulse is higher than the direct pulse (see Fig. 4b). This leads to negative "Excess attenuations" as shown in Fig. 5. The interference pattern is shifted to lower frequencies with higher sound speed gradients as to be expected. The presented time signals are representative single pulse measurements, while spectrum and autocorrelation function are averaged over 30 pulses. For increasing sound speed gradients or lower source heights (where the gradients are stronger) we measured multiple ground reflections (see Figs. 7, 8). The first pulse always is the direct one (Fermat's principle).

Fig. 7: Measured time signal 
(h_s, h_r; 6 m, distance: 500 m)

Fig. 8: Measured time signal 
(h_s, h_r; 2 m, distance: 500 m)

V. CONCLUSIONS

Positive sound speed gradients lead to an increase of the sound pressure level by focussing effects or multiple ground reflections. The direct sound then often is negligible compared to the reflected sound (dependent on the ground).

The sound speed gradient can easily be determined for the nearly neutral case by the time difference between direct and ground reflected signal.

For stronger positive temperature gradients we determine the sound speed gradient from eq. (3) by two measurements at different heights.

/References/


2. B
**13th INTERNATIONAL CONGRESS ON ACOUSTICS**

* YUGOSLAVIA * 1989 *

CONTROL OF INFRASONIC AND LOW FREQUENCY NOISE

PROF. EM. ING. G. L. FUCHS

CIAL- Universidad Nacional de Córdoba, Argentina

Infrasounds and Low Frequency Noise

The audio range is assumed to extend from 20Hz to 20KHz as standardized by ISO-R 226. Research by von Békésy and others has proved that the audible threshold reaches 2Hz at high enough levels. Loudness level (LL) curves crowd closer together below 50 Hz according to H. Møller. Low frequency (LP) and infrasound (IS) thresholds are not an exact extension of ISO-R 226.

Infrasounds below 2Hz exist in Nature. Volcanic explosions and other phenomena (aurorae, waves, etc.) have periods of 500 sec. or more. Systematic research only started around 1950 at NBS by Dr. R.K. Cook in Washington, D.C. (Ref. 1). We set up an infrasound detection station in Córdoba, Argentina 1972 (Refs. 2 and 3).

Infrasounds generated by industry, transport, gun fire, etc. were studied by W. Tempest and co-workers in the U.K. and the von Gierke group in the USA around 1960. Annoyance and non-auditory effects were studied both for IS and low frequencies.

In building acoustics we met with complaints of annoyance at acceptable dBA level generated by industrial and gunfire sources containing intense LP&IS components. It is not easy to absorb or isolate them especially in the 3-1500Hz range. We studied these sources, but conventional equipment does not reveal the fine structure of the spectra. Aided by an FM recorder of recent design and adequate accelerometers, microphones and level meters we have been able to analyze LP&IS with a two channel digital module based on a Z-80 microprocessor storing 4096 points in a 0-375Hz passband with 0.2Hz resolution per line. The module is connected to a PC via an IEEE488 bus interface. It was designed at our institute (Ref. 4). Sharp tones and narrow band components are obvious in the 2Hz-200Hz range (see Figs 1 to 12).

Some Practical Cases

1. A gun factory testing ground disturbed people in a town 3Km away. We recorded shots from a 155mm cannon with 137 dB at 250m. Its FFT spectra disclose a strong infrasonic peak of 19.5 Hz almost unattenuated by distance, as opposed to high frequencies. We designed a silencer of some 50dB attenuation, reported elsewhere (Fig. 1) (Ref. 5).

2. A local automotive diesel engine factory used Renard prop brakes to power test new engines in the range 2400-4500 r.p.m. with two bladed propellers. Noise levels at 1m from source was 121dB at 114Hz in the open. Spectrum shows a tone in exact coincidence with rotation speed (Fig. 3)

\[ f_r = \frac{3420 \times 2}{60} = 114 \text{ Hz} \]
We designed a selective absorber reported elsewhere (Ref.6) based on a tuned multiresonance absorber (Patent pending).

3. A gas-fired boiler in a hospital (150000 Kcal/h) created a disturbing low-frequency noise, with flame roar IS components at 12 Hz & 33 Hz and level of 101 dBLin at 1 m from the burner. Boilermaker geometry enhanced other LF components with peaks 1.25 Hz, 17 Hz, 260 Hz (Fig.2) which check with clusterings of normal modes computed from: \[ f = \sqrt{\left(p/\rho \right) \left( q/w \right)^2 + \left( r/\rho \right)^2} \]

Nearby wardrooms where sound levels did not exceed 48 dBA, strong IS & LF contributed to a measured 75 dBLin, replicating the 1.2 Hz and 3.3 Hz from the boiler. Audible components were also in the boiler’s range of 120-140 Hz (Fig.4).

H. Möller & Y. Inukai (Ref.7) agree that dBA weighting is not realistic when strong IS & LF are present.

Other Types of Noise Sources Analyzed

In search of confirmation of marked IS 7 LF tones which cause annoyance but are not obvious in conventional band spectra, we analyzed several vehicles, bridges and turbines.

- Cabin noise in a popular airliner (Boeing 777) during take-off (Fig.5). The moderate 82 dBA contrast with the 108 dBLin. A 48 Hz tone emerges nearly above the 90 dBLin general level. Level flight displays a steady 4.2 Hz IS. Boeing’s 727 are quite similar.

- A diesel-engined Ford Falcon sedan at 3500 rpm & 110 km/h has a sharp peak of 17 Hz and a cabin level of 110 dBLin (Fig.7). Piezoelectric has confirmed that engine vibration generates tones of 10 Hz while driving. Cabin resonates at 101 and 17 Hz at 90 dB, while general background is below 80 dB. The 17 Hz tone is 10 dB above threshold or 70 Phon LF and therefore disturbing (Fig.17).

- A Mercedes Benz bus at 80 km/h on smooth pavement (Fig.6) has a 13.2 Hz tone in center of cabin with levels 112 dBLin and 82 dBA or 15 dB above threshold according to Yewart. Vibration acceleration spectra repeat the 13.2 Hz in center of windows (Fig.8). Similar results were reported by Sandberg in Sweden (Ref8).

- One of four Francis type hydraulic power turbines (190 MW) in cavern 180 m below water level generates 102 dBA and 112 dBLin. Noise and velocity spectra seen in Figs. 9 & 11. Backstream predictive formula: \[ L = L_0 + 10 \log (A) + 10 \log (A) + 10 \log (A^2) \log 2 \] yields 110 dB for our turbine, within 2% of the actual measurements.

- Fig.10 is a steel girder bridge (40 m span) that resonates at 2.2 Hz generating IS of 112 dBLin below the bridge (Ref.9).

Conclusions

Most of the annoying sources measured contain strong IS & LF tones and the difference dBLin-dBA = 10 dB confirms the advisability of using weightings such as C1 and LF (Möller and Inukai) instead of dBA for evaluation of annoyance.

At low dB levels IS & LF noises annoy people and disturb sleep.

The neurophysiological mechanisms involved are not well understood. We have planned evoked response audiometry experiment in search of subcortical processing of IS and LF signals which might explain its cortical interpretation as annoyance.
REFERENCES

4. Ramos, Oscar: Two channel module (digital) based on a 2-80 microprocessor (Unpublished) CONICET, Argentina.
THE INFRASOUND FROM A VORTEX LOCALIZED NEAR THE OCEAN SURFACE

A.I. Sukhorukov

Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation, USSR Academy of Sciences, 142092 Troitsk, Moscow Region, USSR

V.I. Pavlov

Moscow State University, Department of Physics, 117234 Moscow, USSR.

In the air flow above the disturbed ocean surface there can occur local disturbances of the flow velocity. The flow can be inhomogeneous by the horizontal coordinates or can transfer "frozen" turbulent pulsations. Local inhomogeneities can be also formed in the inhomogeneous flow, having limited dimensions and appreciable local velocities. The examples of such formations can be structures of the type of vortex threads rings, Hill vortices [1], single toroidal vortices [2]. Being near the disturbed ocean surface, such vortices can cause infrasound radiation. The fact is that a local vortex forms a high hydrodynamic voltage zone above the moving ocean surface creating a certain additional distribution of pressures at the ocean surface. This results in the transformation of surface waves into volume ones, i.e., acoustic waves are emitted into the atmosphere and ocean. Since in the spectrum of a developed ocean disturbance the low frequencies are most energy-carrying ones, we should expect the radiation of infrasound waves in the case under consideration. The said effect belongs to the circle of problems concerned with the transition radiation theory [3].

The purpose of the present communication is to examine the effect of an infrasound emitted into the atmosphere by a vortex inhomogeneity localized near the ocean surface, and to estimate the energetic characteristics of the emission.

To describe the process of our interest we use hydrodynamics equations.

We shall neglect the stratification of the medium which occurs due to the presence of the gravitational force. That is possible if the characteristic frequencies of the process under investigation exceed considerably the Brunt-Väisälä frequency \( N \sim 0.02 \text{ s}^{-1} \) for terrestrial conditions.

We shall assume that in the region of the localization of a vortex the values of hydrodynamic (i.e., not connected with the density disturbances) velocity \( \mathbf{v}^{(0)} \) and pressure \( p^{(0)} \) appreciably exceed acoustic values \( \mathbf{v} \) and \( p \), i.e.,

\[
p = p^{(0)}(\mathbf{x}) + p(\mathbf{x},t), \quad |\mathbf{v}| \ll |\mathbf{v}^{(0)}|, \quad |p'| \ll |p^{(0)}|
\]

Outside the region of vortex localization the relation of
pressure $p'$ and density $\rho' = \rho - \rho_0$ is put linear, i.e. $p' = c^2 \rho'$. Here $\rho_0$ is an equilibrium density of the medium, $c$ is sound velocity.

For pressure $\rho = \rho_0(a) + \rho'$ we obtain the equation

$$\frac{\partial^2 \rho}{\partial t^2} - c^2 \frac{\partial^2 \rho}{\partial x_i^2} = c^2 \frac{\partial^2 \mathbf{T}_{ij}}{\partial x_i \partial x_j}$$   (1)$$

where tensor $T_{ij} = \rho \bar{u}_i \bar{u}_j \approx \rho \bar{u}_i^{(a)} \bar{u}_j^{(a)}$ expresses the surplus value of the flux momentum. We choose a coordinate frame so that plane $xy$ coincides with the nondisturbed ocean surface, and direct the $Oz$ axis upwards (Fig. 1). To describe region $\Sigma$ of the vortex localization we shall use the following model which seems simple enough. We shall believe that the distribution of local hydrodynamic disturbances in the $\Sigma$ region is described by a certain function $f(\bar{x}) - f(|\bar{x}|/\sigma, |\bar{x} - H|/\ell)$, where $\bar{x} = \{\bar{x}, \bar{z}\}$, $\bar{z} = \{x, y\}$, $\sigma$ and $\ell$ are characteristic dimensions of the $\Sigma$ region. The $f(\bar{x})$ function reaches the extreme value in point $x_0 = (0, 0, H)$ and decreases rather fast while moving away from it. Assume also that the $\Sigma$ region is fairly stretched along the vertical, i.e., $\sigma \ll \ell$; and the particles of the medium are moving inside it so that the vertical components of their velocities do not exceed the horizontal ones, i.e., $\bar{u}_z^{(a)} \ll \bar{u}_x^{(a)}$.

Under the assumptions made, the term in the equation (1), which characterizes the effect of a source like that, can be rewritten as

$$\rho_0 c^2 \frac{\partial \bar{u}_i^{(a)}}{\partial \bar{x}_j} \frac{\partial \bar{u}_j^{(a)}}{\partial \bar{x}_i} = \rho_0 c^2 \frac{U^2}{\sigma^2} f\left(\frac{|\bar{x}|}{\sigma}, \frac{|\bar{x} - H|}{\ell}\right),$$   (2)$$

where $U$ is local hydrodynamic velocity characteristic for the $\Sigma$ region.

The disturbance of the surface will be described by function $\nabla \times (\bar{z}, \ell)$ which determines the displacement of the points of the surface with respect to its mean level, and $\langle \bar{z} \rangle = 0$.

The equation (1) is supplemented by a boundary condition which, assuming the infinitesimal of the inclinations to the surface $(|\bar{z} - H|/\ell \ll 1)$, can be represented in the form of the power expansion $[4]$

$$\frac{\partial^2 \rho}{\partial \bar{z}^2} \bigg|_{\bar{z} = 0} = \rho_0 \frac{\partial^2 \bar{z}}{\partial \bar{z}^2} \bigg|_{\bar{z} = 0} + \bar{\nabla} \bar{z} \cdot \bar{\nabla} \rho \bigg|_{\bar{z} = 0} + O(\eta^2)$$   (3)$$

38
and by the condition of radiation. Then it is only natural to seek the solution of the problem stated in the form of the series

\[ \rho = \rho^{(0)} + \rho^{(1)} + \ldots \]  

(4)

where \( \rho^{(n)} \sim \eta^n \).

The first term in this expansion describes a static field of the pressure disturbances, caused by the presence of stationary flows in semi-space \( z > 0 \); the \( \rho^{(0)} \sim \eta \) term determines the radiation field. The subsequent terms in (4) are responsible for the nonlinear effects and will not be treated here.

We shall assume, that the displacements \( \eta(\vec{x}, t) \) of the points of the ocean surface are distributed according to the normal law, and Fourier-expansion of their correlation function looks as \[ G_\eta(\vec{x}, \omega) = \frac{2a_t}{a_q} \frac{\omega}{2\pi} e^{-\omega^2} \frac{g^2}{2a_0} \delta(\omega - g) \cdot F(\gamma) \]  

(5)

where \( V \) is wind velocity above the surface; constant \( \beta \approx 6.5 \cdot 10^{-3} \); function \( F(\gamma) \) sets the angular distribution of the spectrum. Besides, in order to obtain an analytical solution of the problem, we shall choose function \( f(\vec{x}) \) in the form

\[ f(\vec{x}) = a^{-3/2} \exp \left\{-\frac{1}{2} \left( \frac{a}{\beta} \right)^2 \right\} \]  

(6)

Then, for the spectral emission power we find

\[ \hat{E}_\omega = \frac{a_0 g^2 u^4}{2\omega^2} \frac{g^2}{2a_0} \omega e^{-\omega^2(\gamma + \omega)^2/2} \]  

(7)

As follows from (7), the region of the optimal infrasonic generation for \( V^2 \gg g^2 \) corresponds to frequencies \( \omega \sim g/V \); for \( V^2 \ll g^2 \) - to frequencies \( \omega \sim (g/V) \cdot (V^2/g^2)^{1/4} \).

The radiation is practically isotropic. This shows a principle difference from the mechanisms of infrasonic radiation in the atmosphere-ocean system treated earlier (see, e.g., 6 - 8). Those works give a diagram of the direction of the infrasonic radiation which is characterized by factor \( \cos^2 \beta \), \( \beta > 2 \). In our case the radiation along the ocean surface is practically the same as upwards from it. This result is important, for it permits to explain the super-remote propagation of infrasound from the region of its origin.

References

2. Petviashvili, V.I., Pokhotlov, O.A., and Chudin N.V.
MULTIFOCAL IDENTIFICATION OF SOUND SOURCES

V.I. FAVLUN, A.N. ZHUKOV, A.N. DVANNIKOV.

119039, PHYSICAL DEPARTMENT, MOSCOW STATE UNIVERSITY,
LENINGRAD GORAY, MOSCOW, USSR.

The subject of this paper is the investigation of the space distribution of energy characteristics of multifocal sources field, the obtaining of analytical equations in energy terms, as well as the illustration on a number of typical examples of the connection of these values with actual parameters of source [1-5].

The complex structure of near source field allows to use the vector of energy density flow: \( \mathbf{S} = \rho(\omega) \mathbf{R}(\omega) = \mathbf{R} + i \mathbf{J} \).

There is actual physical interpretation of active and reactive components of vector \( \mathbf{S} \): \( \mathbf{R} \) describes the energy radiated from the source; \( \mathbf{J} \) describes hydrodynamic motion of medium near the source.

Let the motion of medium near compact source be singularly described by the velocity potential. In accordance with the wave principle the sound source forms in the medium a separate spherical wave of velocity potential, and, consequently, pressure since the potential satisfies the wave equation. Having this circumstance in mind it is not difficult to show that there exists the correlation between radial partial velocity and pressure \( \mathbf{P} \):

\[
\mathbf{\nabla}(r,t) = \frac{\mathbf{P}(r,t)}{\rho c} + \int_0^r \mathbf{\nabla} \mathbf{P}(r',t) / \rho c \, dr',
\]

(1)

where: \( \rho \) is the density of air,
\( c \) is the velocity of sound,
\( r \) is the distance between point source and point of observation.

In eq. (1) the first member is wave component, and decreasing in accordance with the condition of radiation and equal to powerflow of energy radiated from the source.

In eq. (1) the second term describes hydrodynamic medium motion not connected with the wave process of energy transference, and decreasing is comparable to \( r^{-2} \). It describes a "fur-coat" round the source. To be sure of this one should put \( c \rightarrow \infty \) in eq. (1) (i.e. hydrodynamic approximation). In this approximation the second member in eq. (1) is non-zero, unlike the first one.

By using Fourier-transform of eq. (1), the Fourier components of velocity and pressure are connected as

\[
\mathbf{\nabla}(k) = \rho(k) (\mathbf{1} + i / k c) / \rho c,
\]

(2)

where: \( k = \omega / c \) is the wave number. Equation (2) shows the physical sense of phase mismatch between Fourier components of velocity and pressure. The phase mismatch also shows that there also exists non-wave component of medium motion the source, which does not radiate energy.
The introduced sound intensity vector with the account of $P$ and $\vec{v}$, following from the linear equation of hydrodynamics, can be presented in the form: $\vec{S} = i \vec{P} \vec{V} P^* / \varphi \omega$. Omitting the calculations, we'll demonstrate the final result. For the noise source multipoles, the following form is obtained:

$$\vec{S} = \kappa^2 |M_{l,m}|^2 \{ (e^{(1)}_l) (z_2) Y_{l,m}^*(z_2) (e^{(2)}_l) (z_1) Y_{l,m}(z_1) \} / \varphi \omega [(2l+1)!!]^2 }$$

For the object of this work, the expression is fundamental. Let's make use of the convenient form of the $\nabla$-operator ($\nabla \approx \vec{e}_x \partial_1 - i [\vec{e}_y, L] / z^2$), where $L = - i [\vec{e}_y, \vec{v}]$ — the operator momentum of quantity movement, affecting only the angle variable.

The radial component vector $\vec{S}$ according to eq. (3) can be expressed in the form

$$\vec{S}_{(r)} = \vec{e}_r \cdot \vec{S} = -i \kappa^2 |M_{l,m}|^2 Y_{l,m} \Omega_L(z_2) / \varphi \omega [(2l+1)!!]^2$$

Here was used the correlation, received from the recurrent expansion for the multipole sphere of Bessel and Neumann functions

$$\Omega_L(z) = \frac{1}{z} (1/2) (z/2)^l, \quad \text{(Neumann function)}$$

The expansion is correct for any meaning of the argument $z = k \rho$. To calculate the component $\vec{S}$, which is orthogonal to the vector $\vec{S}$, we propose summing by repeating index from 1 to 3. The second part $\nabla$-operator is convenient to write in components:

$$- \vec{e}_{\text{sym}} \cdot \vec{\nabla} = \frac{1}{z^2} L_{\text{sym}} / z^2$$

For complex fundamental $3$-vectors $\vec{e}_{\mu}$, which give definite advantage in calculation with the spherical functions and which are connected with the cartesian coordinate unit vectors from expansion:

$$\vec{e}_x = a_2 + i a_3 / \sqrt{2}, \quad \vec{e}_y = (a_3 - i a_2) / 2$$

with the observation of the condition: $\vec{e}_\mu \cdot \vec{e}_\nu = \delta_{\mu\nu}$. In this conditions the equality has place

$$L_{\text{sym}} Y_{l,m} = (-1)^m \sqrt{(2l+1)} (L, m) Y_{l,m}$$

Here $(L, m)$ are the vector addition coefficients or the Clebshe~Gordan coefficients, taking into consideration $Y_{l,m}^* = (-1)^m Y_{l,-m}$ and the correlation of symmetry for the Clebshe-Gordan coefficients: the expression (6) can be simplified.

Then, in view of general expression for the intensity vector (3), as stated above expressing angle variable $\nabla$-operator, and also that can be given in the form $\vec{S} = \sqrt{\varphi / \omega} \cdot \vec{V} P^*$, the tangential components vector $\vec{S}$ can be obtained in the following form.
\[ S_{(k)} = \left\{ k^2 (-1)^k \left[ 4 \pi / 3 \left( M_{lm} \right)^2 \right] \right\}^{1/2} \frac{1}{\rho \omega \rho} \left[ \left( \frac{2L+1}{2L+1} \right) \right] \]

In this description we propose summing by repeating indexes. The equations (4), (7) present precise expressions for the intensity vector components from multiple sound source order \( L \). In spite of unwieldy appearance they possess clear expressing properties of symmetry and are highly informative as a matter of fact. It is not difficult to see that intensity vector components can generally contain both real and imaginary parts, whose physical meaning was discussed earlier.

The analysis of the expressions (4), (7) shows that they contain the amount of information sufficient to determine the coordinates and the orientation of the multipole sound source in the space. First of all it is possible to solve this problem, because the real \( \overline{R} \) and imaginary \( \overline{I} \) parts of the intensity vector are not collinear in any point of space observation. For example, in the spherical system of coordinates for the dipole sound source the parts are defined by:

\[ \overline{R} = \left\{ M^2 k^2 \cos^2 \Theta / 16 \pi^2 \rho \omega^2 \right\} \overline{e}_e \]

\[ \overline{I} = M^2 \left[ (k^2 / \rho^3 + 2 / \rho^5) \cos \Theta \overline{e}_e + (1 / \rho^5 + k^2 / \rho^7) \sin \Theta \cos \Theta \overline{e}_e / 16 \pi^2 \rho \omega^2 \right] \]

where \( M = \text{Ad} \cdot d \) is the distance between the point sources of the same productivity \( A \). The non-triviality of this type of source sound-field structure results is non-equality of real \( \overline{R} \) to zero:

\[ \text{Re} \overline{R} = \left\{ M^2 k^2 \sin 2 \Theta / 16 \pi^2 \rho \omega^2 \right\} \overline{e}_e \]

It is important to note, that in such sound fields the motion of medium particles is realized by ellipse trajectories and, in other words, the velocity vector in the field presentation rotates along the ellipse plane [6].

It's easy to see from equations (8), (9) that by measuring dipole source sound field vectorial energy characteristics' magnitude and direction one can obtain the source location in full volume. In particular, to solve this problem we can use calculation formulae determined by (8), (9):

\[ \frac{1}{2} \frac{|R|}{|d|}, \frac{|R|}{|d|} = \frac{2}{2} \left[ \Theta \Theta \Theta \right], W = 4 \pi \frac{|R|}{3 \cos^2 \Theta} \]

where \( W \) is the power of a dipole source. The direction to the dipole source is defined by the direction of \( \overline{R} \) vector, the plane
in which the source is oriented is the plane of \( \mathbf{R} _{1} \), vectors, the distance \( r \) and the angle \( \Theta \) of the axis dipole to the direction \( \mathbf{R} \) vector can be considered from equation (10).

An interesting supplement of the conducted general consideration is the possibility to determine the type of source with the help of measuring vectorial energy characteristics. To clarify the essence of the question let us consider the case \( z = k z < 1 \). From the exact presentation (5) we can found

\[
4(2)(3/2k) \left\langle h_{l}^{(2)} \right\rangle^{2} \sim - \left\langle (l+1)[(2l+1)!]^{2} \right\rangle^{2} \left( k \right)^{2l+3}
\]

Then from (7) we obtain for the radial component intensity

\[
S_{(r)} \sim \left[ \frac{k^{2}}{l} M_{lm}^{2} |v_{lm}|^{2} \omega \right] \left[ \frac{1}{(k^{2})} \left\langle (l+1)[(2l+1)!]^{2} \right\rangle \right]^{2} \left( k^{2} \right)^{2l+3}
\]

Further it is not difficult to see that by the frequency dependence of the function \( \Phi(\omega) = |\mathbf{P}(\omega)|/|\mathbf{P}(\omega)| \) the information can be acquired about the order of multipole of the source

\[
\frac{\Phi(\omega)}{\Phi(\omega_{2})} = \left( \frac{\omega}{\omega_{2}} \right)^{2l+4}
\]

This obtained correlation is of sufficiently general character. In the case of axisymmetrical sources, it is possible to measure both real and imaginary parts from the radial intensity vector \( \mathbf{S}_{(r)} \) and accordingly to use the correlation (10) for the determination of the order of multipole source [7].

SCATTERING OF NON-ACOUSTIC ENERGY BY GEOMETRICAL INHOMOGENEITIES ON A FLAT PLATE

Alan Powell

Department of Mechanical Engineering, University of Houston
Houston, Texas 77204-1972, USA

Introduction

Rayleigh (1954) analyzed the scattering of sound waves by inhomogeneities in the medium by the use of the Inhomogeneous wave equation. The method can be adapted to any motion satisfying the wave equation, e.g., to the scattering of evanescent waves, while a simple physical argument hinges on the differential acceleration and compression (due to density and compressibility variations) of the inhomogeneities can be used for propagating or evanescent waves, (Powell 1969). This method also applies to fixed, rigid inhomogeneities.

Lighthill (1952) expressed the sources in an unsteady flow as volume distributions of lateral and longitudinal quadrupoles. Curle (1955) included the influence of boundaries, showing surface distributions of dipoles and monopoles, (from surface stresses and boundary motion respectively), and Powell (1960) showed by an image argument that for a plane and rigid surface the surface pressure equals to just the reflection of the quadrupole sources.

It was pointed out that the image argument fails at the edge of a semi-infinite plane surface where, for example, a boundary layer gives rise to 'edge noise' (Powell, 1969). Chandiramani (1974) considered the scattering at the edge of evanescent waves which represented the turbulent convecting wall pressures under a boundary layer.

Rabinovich et al. (1984) used Curle’s results in an original image argument to consider the scattering of the energy of a boundary layer caused by a single rather smooth bump on the wall, the technique was used again by Routov and Rybushkina, more generally (1986). Howe (1984) noted that the image argument, as presented, also fails when there is a local bump — a "roughness element" — on the plane surface, and estimated the sound generated by the consequent scattering of the wall pressure field of a boundary layer on a rough surface, (see also Howe, 1986).

Here we use a the image scheme together with the physical scattering argument to consider some of fundamental issues concerning the effect of a bump and many bumps on a plane rigid surface bounding an aerodynamic flow. The pertinent characteristics of the flow are assumed to be the same whether or not the bumps are present.

Modified Image Argument

Consider a region of unsteady vertical flow of a slightly compressible fluid which generates sound. There is a contiguous infinite rigid plane surface, on which there is a rigid bump, which is taken to be very small in extent compared to acoustic wavelengths of interest, see Fig. 1.

In the formalism of the Lighthill-Curle theory, the sound sources are volume distributions of lateral and longitudinal quadrupoles, and surface distributions of dipoles associated with the local fluctuating pressure on the plane and on the bump, as symbolised in Fig. 1. As they play no part in the scattering phenomenon of present interest, the dipoles associated with the viscous stresses are neglected.
We postulate the image flow in the plane, see Fig. 1. The pressure dipole on the plane of symmetry cancels exactly. On the body formed by the bumps and its image the components of the pressure dipole normal to the plane of symmetry cancel, with a negligible quadrupole remainder.

The components parallel to the plane fail to cancel in both the real flow and in the image flow in the presence of a pressure gradient parallel to the plane. Therefore, the surface pressures on the body result in a force, parallel to the plane and opposite to the direction of the pressure gradient. If the pressure gradient is taken as locally constant and unaltered by the body, the force is equal to the pressure gradient times the volume of the body.

We estimate the scattered acoustic radiation. The fluid at the position of the body but in its absence has velocity \( u \) given by \( \frac{3u}{2l} = -(1/\rho) \frac{\partial p}{\partial y} \) where \( \rho \) is the fluid density and \( p_0(y) \) is the local incident pressure. A sphere of radius \( a \) oscillating with velocity \( u \) with respect to the surrounding medium generates an acoustic far field pressure \( p(x) = \left( p/2 \right) (a/2R)^2 (a^2/c^2R^2) \cos \theta \), \( c \) = sound speed, \( \theta \) = angle between \( u \) and \( x \). Hence, taking the body to be spherical for present convenience, we find that

\[
p(x) = \left( 1/2 \right) (a/2R)^2 (a^2/c^2R^2) \cos \theta
\]

At this point, the vertical flow may be in contact with the plane, or not.

If it is far away, the body is in its far field, and classical Rayleigh-type scattering occurs, resulting in a dipole parallel to the plane and an additional monopole-like source of exactly the same order of magnitude due to the lack of compressibility of the body. The incident pressure from the quadrupole sources varies as the characteristic velocity to the fourth power, so the scattered sound pressure varies as the sixth power, a very weak effect at low Mach numbers.

But when it is very close to the plane, the body lies in the very near (hydrodynamic) field of the quadrupole sources. If fluctuations of the near field pressure are taken to vary as \( p_{ii} \), then the scattered sound pressure is seen to vary with \( a^4 \) and the sound power with \( a^6 \), of order (Mach number) \( D \) greater than that of the other (quadrupole) sources present.

The forces must depend on \( h^2 \) if the quadrupoles are distance \( h \) from the plane, giving strong relative emphasis to sources close to the plane.

Further discussion is postponed until combined with the result for many bumps.

Many small close bumps

Consider a large number of hemispherical bumps of uniform size, randomly but homogeneously spaced at average distances apart \( 8y_1 = 8y_2 \) in a narrow rectangle with sides \( 2l_1 \) and \( 8y_2 \). Throughout these dimensions will be considered small compared to the acoustic wavelength. Take the \( y_1 \) direction. Each bump has a scattering dipole source strength proportional to the local pressure gradient in the \( y_1 \) direction and the volume \( V_0 \). Replace each spherical volume \( V_0 \) by a layer of thickness \( 2w \) and area \( 8y_1 \) as in Fig. 2. Then together

\[
\sum \frac{\partial p(y_1,t)}{\partial y_1} V_0 = \sum \frac{\partial p(y_1,t)}{\partial y_1} \wedge \delta y_1 \rightarrow 2w8y_2 \int \left[ a(y_1) \frac{\partial y_1}{\partial y_1} \delta y_1 \right] = 2w8y_2 \left( p_{ii} + l_{ii} \right)
\]
Fig. 3

The limiting process is valid if the bump distribution is fine enough compared to the rate of change of the pressure gradient, i.e., if the \( \gamma_1 \)-wave numbers of the pressure field are much lower than that of the spacing between the bumps. For example, if the pressure field is due to quadrupoles at height \( h \) above the wall, it is seen from Fig. 4 that \( 8y < h \) is necessary.

The total dipole strength is then to the difference in force due to the pressures on the virtual steps of height \( 2w \) and width \( 8y \) at distance \( 2L \) apart at the edges of the scatterers.

This result is the same as for a single real rectangular bump (but note that the virtual one has \( 3/2 \) times the total volume). The hemispherical bumps, due to the virtual inertial of the fluid oscillatory flow about the bumps. The following considerations apply equally well, for the given circumstances, to a single flat bump or the the virtual one representing a bumpy patch.

For rectangular areas \( 2L \times 2L \) wide, the pressures have to be integrated over the faces of the steps to yield the total dipole strength in the \( y_1 \) direction. Similar considerations apply for the dipole strength in the \( y_2 \) direction.

Consider the situation when various types of quadrupole are centered above the bumpy patch at height \( h \); the pressure fields for two are sketched in Fig. 4.

Quadrupoles with one axis parallel to the \( y_1 \) direction, as the lateral one shown at the left, will result in relatively large forces in the \( y_1 \) direction, especially if \( 2L \), is about \( h \), and the others have zero forces by reasons of symmetry.

For \( L_1 \) and \( L_2 \) larger (a few times \( h \)) but still in the very near field, the force diminishes as \( 1/L^2 \) in each case. In the limit, \( L \to \infty \), there is simply a new plane at height \( h \) above the original one, and the integral of the \( \gamma_1 \)-derivative of the pressure over the plane must be zero (as the integral of the pressure is zero, by the reflection argument).

Now consider the situation when the various types of quadrupole are situated near a single step, either a real one or an virtual one at the edge of a large bumpy patch. The result is radically different. Maximum pressures will be exerted on the face of the step when any of the types of source are near by over, or over, the step, with the exception of those quadrupoles which have one axis parallel to the step (the result then being zero). We note in passing that the force on the step is the same as that exerted on a strip \( 2w \) wide on the plane at the position of the step.

While this picture is very simple for many real situations, it is highly suggestive that even though there is very little scattering to be expected from a very fine uniform roughness covering a large enough patch (but still compact, i.e. small compared to an acoustic wavelength), that edges of the patch will be relatively strong scatterers, like a single step.

By a simple extension of this reasoning, we can consider a smooth shallow excrescence, either real or representing a graded surface roughness: The in-plane dipole force then depends on the variation of the pressure gradient over its profile, instead of just the values at the edges.

**Distributed connected source field**

The discussion has concerned time-varying stationary sources. Now moving constant ("frozen") sources are considered, that is, the sources do not change significantly in phase during the time of interaction. Take the sources to be in a layer some distance \( h \) from the plane, moving at a constant subsonic velocity, see Fig. 2.

The previous type of argument carries over qualitatively to this case, but each quadrupole
now completely traverses the bump or bumps. Individual small bumps (a ≪ h) behave just as before. Larger areas of roughness do not radiate from the central parts (further than a few h from the edges), provided still that the roughness is fine enough, i.e. 8y ≪ h. There is relatively vigorous interaction at the edges (real or representing the edge of an area of roughness) from all quadrupoles, except for those with both axes parallel to the plane which have a weak directionality towards the plane. The resultant scattered sound power is proportional to the mean square pressure, the area of the step squared (and the cross-stream correlation length would be introduced if the quadrupoles were randomly distributed). The U^2 law applies.

At the opposite end of the scale, if 8y > 5h or so in this model, the pressure due to a particular quadrupole on each bump becomes uncorrelated with that on its neighbor, as is evident from Fig. 4, and for boundary layers the correlation between neighboring quadrupoles falls rather rapidly. In that case, there is scattering from each of the bumps, according to the formula given, which is uncorrelated with that of its neighbors. The sound power is due to the sum of them, i.e. proportional to the area of the roughness. The U^2 law applies of course.

An alternative approach is to consider the pressure field of randomly distributed "frozen" quadrupoles in the layer as evanescent waves carried between the layer and its image at subsonic speed U, Fig. 5. The pressure falls almost exponentially with distance from the layer (corresponding to the h^2 earlier).

The scattered sound pressure is given by the equivalent formula to that before if the Mach number is small, (Powell 1969), namely for a solitary hemispherical bump

\[ p(x) = \left(\frac{1}{2}\right) \frac{a^2}{\pi a} U k L p_{in}(k) \cos \theta \]

for wavenumber \( k \) of the pressure of given amplitude \( p_{in}(k) \). Here \( U k L = (k)c k_1 \), \( k = \) acoustic wavenumber, corresponding to \( 2\pi a/2\pi a \) before.

In this approach we have a variation on the "edge noise" due to the scattering of evanescent wave at the edge of a half-plane (Chandrasekhar, 1974), due to the displacement thickness of a real slot or the virtual one due to very fine roughness, see Fig. 6.

This short discussion may provide some insight to some aspects of scattering offlow energy occurring in more complex situations. For more comprehensive and detailed treatments of boundary layers see the appropriate references, and especially Howe 1988.


1 One result of the analysis is that when many bumps radiate individually (uncorrelatedly) there is also an edge effect dipole due to the displacement effect caused by the mean thickness of the bumps.
L'ACOUSTIQUE DES REVETEMENTS ROUTIERS POREUX.

M. Bérengier (*), J.F. Hamet (**), P. L'Hermite (*), P. Daburon (*)

(*) L.C.P.G. - Centre de Nantes - BP.19 - F44340 BOUGUENAISS
(**) I.N.R.E.T.S.- 109, Av. S. Allende - F69675 BRON Cedex

Le bruit routier est actuellement un des facteurs de gène majeur. L'amélioration acoustique importante apportée aux véhicules a mis en évidence la prédominance du bruit de roulement au delà de 60 km/h. Afin de réduire ce bruit, plusieurs moyens d'action sont envisageables : sur le pneumatique, sur la chaussée, sur l'environnement proche de la route. Etant donné les critères de sécurité, l'action sur le pneumatique est très limitée et l'implantation d'écarts acoustiques est quelquefois problématique notamment en milieu urbain. La seule action restante demeure dans le domaine de la chaussée et notamment du revêtement. Dans cet article, nous présentons des études tant théoriques qu'expérimentales réalisées sur de nouveaux types de structures poreuses qui allient des qualités de drainabilité (élimination de l'hydroplanage) à des qualités d'absorption acoustique.

I. RAPPELS THERORIQUES

1.1. Coefficient d'absorption acoustique

Sous incidence normale, le coefficient d'absorption acoustique d'un milieu quelconque, s'exprime en fonction du module $r$ du coefficient de réflexion par la relation : $\eta = 1 - r^2$ (1)

ou : $\eta = (Z - \rho_s c_s)/(Z + \rho_s c_s)$ ($\rho_s c_s$: impédance de l'air) (2)

Pour une couche, d'épaisseur $e$, de matériau poreux reposant sur un support d'impédance $Z_\Omega$, l'impédance $Z$ de la couche est :

$$Z = \frac{\rho_s c_s}{\Omega} \left\{ \frac{Z_\Omega \coth \gamma e + \frac{\rho_s c_s}{\Omega}}{(Z_\Omega + \frac{\rho_s c_s}{\Omega} \coth \gamma e)} \right\}$$ (3)

Où $\rho_s c_s$ est l'impédance spécifique du milieu et $\gamma$ la constante de propagation du milieu. Connaissant l'impédance de la couche inférieure, il est possible par un procédé itératif de déterminer l'impédance de chaque couche et donc l'impédance de la surface d'une structure multicouche. $\eta_n$ est ensuite évalué avec (2) et (1).

1.2. Coefficient d'absorption acoustique d'un revêtement poreux

Le milieu poreux est assimilé à un fluide homogène au sein d'une structure rigide immobile. À partir des équations d'état, de continuité et de la dynamique, il est possible d'évaluer $\rho_s$ et $\gamma$ en fonction de paramètres physiques caractéristiques du milieu ($\sigma$ : résistance spécifique au passage de l'air, $\Omega$ : porosité communicante et $\eta$ : facteur de forme lié à la tortuosité du milieu et identifié directement à partir d'une mesure au tube de Kundt) [1]. Dans ces conditions, on obtient pour $Z$ la relation suivante (dépendance temporelle en $\varepsilon$) :

$$Z = \epsilon_p c_p \left[ 1 + \frac{\rho_p}{\rho_s} \right] \coth \left\{ \frac{\varepsilon}{\Omega} (2 \pi \rho_s / \rho_p) \cdot \left[ 1 + \frac{\rho_p}{\rho_s} \right] \right\}$$ (4)
avec :  \( \rho_p = \rho_s / \Omega \) ;  \( c_p = c_s / \sqrt{\Omega} \) ;  \( \rho_s = \sigma / (2\pi \cdot \rho_p) \)

\( \rho_s \) : masse volumique de l’air ;  \( c_p \) : vitesse du son dans l’air

La relation (4) correspond à celle trouvée par Von MEIER [2]

1.3. Application aux chaussées monocouches de faible épaisseur

Quand le support de chaussée est rigide et imperméable (\( Z_f = \infty \)), la relation (3) devient :  
\( Z = \csc \gamma e \)  

(5)

Le coefficient d’absorption croît en un premier temps, puis oscille en fonction de la fréquence. Les valeurs extrêmes d’absorption sont obtenues pour des fréquences [1]  
\( \rho_m = m (c_s / \rho e) \cdot \sqrt{\Omega} \)  

m = 1, 2, 3...

(6)

Les maxima sont obtenus pour les valeurs impaires de m ; la vitesse des particules est alors maximum à la surface du revêtement. Connaissant la fréquence du premier maximum (\( f_1 \)), on peut alors calculer le facteur de forme \( f \) introduit dans la relation (4)

1.4 Application aux chaussées monocoaches de forte épaisseur

Lorsque l’épaisseur de la couche est suffisamment grande, \( \csc \gamma e \) tend vers 1. Dans ce cas \( Z = \xi \), la courbe d’absorption devient monotone et tend vers une valeur asymptotique indépendante de \( \sigma \) [3]. Dans ce cas :

\( \alpha_e(\text{maximum}) = 1 - \left\{ \left[ \sqrt{\Omega \xi} / \Omega - 1 \right] / \left[ \sqrt{\Omega \xi} / \Omega + 1 \right] \right\} \)

(7)

Quand \( \sigma e > 3. \rho_s c_s / \sqrt{\Omega} \) (avec \( s_f < 7 \) et \( \Omega < 0.5 \)), la différence \( |e_2 - e_1| \) où \( e_2 \) est donné par (1) et \( e_1 \) est la valeur de \( e \) pour \( Z = \xi \), devient négligeable. On dit alors que l’on a atteint la condition de super épaisseur.

2. RÉSULTATS DE MESURE-INFLUENCE DES DIFFÉRENTS PARAMÈTRES.

En un premier temps, les différents calculs ont été confrontés à des résultats d’essais sur des échantillons monocouches et multicouches de 10 cm de diamètre et d’épaisseur 5, 10 et 15 cm. Alors que la littérature [1, 2] permet de quantifier l’influence des variations de \( \sigma \), \( \Omega \) et \( s_f \) sur le facteur d’absorption ; ces essais effectués au tube à ondes stationnaires ont permis de mettre en évidence l’influence des autres facteurs tels que : l’épaisseur, le nombre de couches, l’humidité, la granulométrie et le liant. Pour l’ensemble des courbes, nous avons pris \( \sigma = 20 \) cgs rayons mesurée suivant la norme ISO [4], et des valeurs de \( \Omega \) et \( S_f \) respectivement égales à 0.25 et 3.5.

La figure 1 compare les valeurs de \( e_1 \) pour deux structures de même composition et d’épaisseur différente (5 et 10 cm). Nous observons, d’une part, le déplacement des maxima d’absorption et la bonne concordance entre le modèle et les expérimentations d’autre part. Lorsque nous comparons les résultats théoriques et expérimentaux pour un multicouche (figure 2), nous retrouvons un bon accord. Les mesures physiques de \( \sigma \) n’ont pas montré de différences importantes entre des granulométries voisines. Un multicouche se comporte donc dans ce cas, au niveau de l’absorption, comme un monocouche d’épaisseur équivalente.

52
En revanche, la présence d'eau en rétention dans le milieu peut modifier de façon sensible la courbe d'absorption (figure 3). Celle-ci baisse de façon globale et les maxima sont décalés vers les basses fréquences. Ce dernier facteur est à prendre en considération car les chaussées réelles sont rarement parfaitement sèches.

Nota : Le problème du liant a été abordé, les premiers résultats semblent intéressants mais demandent confirmation.

3. CAS D'UNE STRUCTURE DE FORTE ÉPAISSEUR

Sur une chaussée de forte épaisseur (51 cm) récemment construite sur la commune de Rezé-France, des essais ont été réalisés par une technique impulsionnelle [5] et ont été comparés avec la relation (6) avec les paramètres suivants : \( \sigma = 20 \) cgs rayls ; \( \varphi = 3.5 \); \( \Omega = 0.15 \) (figure 4-1) et 0.25 (figure 4-2).

Nota : Une seconde expérimentation de chaussée de forte épaisseur est actuellement en cours en France.
4. IMPACT D'UN REVÊTEMENT POREUX SUR L'ENVIRONNEMENT

Différents calculs et expérimentations ont montré l'impact d'un enrobé poreux sur l'environnement urbain et péri-urbain. Dans le premier cas, les récepteurs sont placés en façade d'habitation à 1m20 et 3 m de hauteur et dans le second cas le récepteur est placé à 1m20 de hauteur et à 7m50 de l'axe de roulement du véhicule.

Pour ces deux hypothèses, les résultats sont comparés à ceux obtenus sur un revêtement réfléchissant imperméable. Les calculs en milieu urbain ont été effectués par une technique d'éléments finis [6], avec une source de bruit identique pour tous les revêtements. Les mesures en milieu péri-urbain ont été effectuées par la méthode du "véhicule au passage" en conformité avec la norme ISO R 362. Dans ce cas, les résultats tiennent compte à la fois de l'absorption et de la modification du mécanisme générateur de bruit (bruit de contact). Les valeurs moyennes ont été portées dans le tableau suivant :

<table>
<thead>
<tr>
<th>Environnement poreux</th>
<th>Péri-urbain (*)</th>
<th>Environnement urbain (**)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e = 4 cm</td>
<td>- 5 dB (A)</td>
<td>e = 5 cm</td>
</tr>
<tr>
<td>e = 6 cm</td>
<td>- 6 dB (A)</td>
<td>e = 10 cm</td>
</tr>
<tr>
<td>e = 50 cm</td>
<td>- 7.5 dB (A)</td>
<td></td>
</tr>
</tbody>
</table>

(*) : Niveau maximum au passage à 80 km/h (Lp max.- moteur coupé)
(**) : Atténuation propre à l'absorption de la structure (roulement + moteur)

5. CONCLUSION

L'ensemble des calculs et mesures exposés dans ce papier permettent d'apporter les conclusions suivantes : Un revêtement routier poreux peut être modélisé de façon réaliste et son impact sur l'environnement peut être important. Les facteurs prépondérants sont l'épaisseur et le pourcentage de vides communicants qui conditionnent les valeurs σ et Ω. Les atténuations en milieu urbain ou péri-urbain peuvent paraître mal corrigées entre elles, mais cela est dû au bruit du moteur qui est important dans un cas (milieu urbain-faible vitesse) et négligeable dans l'autre (milieu péri-urbain-vitesse > 60 km/h). En plus de ces qualités absorbantes, un revêtement poreux présente l'avantage majeur d'éliminer le phénomène d'air pumping fortement générateur de bruit [7]. L'ensemble de ses propriétés acoustiques alliées à celles d'adhérence à haute vitesse et de drainabilité font de ce type de revêtement un procédé d'avenir.

6. BIBLIOGRAPHIE

ADVANCES IN THE REDUCTION OF SUPERSONIC JET NOISE THROUGH RADIAL DIFFUSION

J. de Krasinski and S. Sun

Dept. of Mech. Eng., University of Calgary, Calgary, AB Canada T2N 1N4

1. Introduction

A substantial decrease in supersonic jet noise was observed during research on a radial diffuser and was reported in /1/. A similar sudden drop in noise level was observed during a sudden expansion of a supercritical convergent duct into a larger diameter one /2/. A research program was designed under two headings: (i) a further study of supersonic noise attenuation on a modified radial diffuser in which attention was given to improve its acoustic performance, (ii) a study of acoustic attenuation due to a sudden expansion into supersonics. This study to be performed on very simple models was hoped to provide more basic information, particularly concerning common factors relating cases (i) and (ii) in apparently very different situations. Also noise attenuation in accelerating flows might prove useful for supersonic jets in aircraft.

2. The Models

Fig. 1 shows the geometry of the modified radial diffuser called "Silencer-Diffuser". As the flow proceeds the annuli get narrower in the radial direction and the expansion occurs along the widening circumference. A spike located on the axis, due to its geometry produces a stable curved conical shock wave which reflects from the opposite wall. The flow leaves the diffuser at low speed. This model was tested at M = 1.5 to 4.0. Later on a muffler composed of thin walled tubes was added. Fig. 2 shows the noise spectra and the performance of the silencer.

The expanding flows were studied on models consisting of converging ducts fitted with pipes of varying lengths and diameters to allow a Prandtl-Meyer expansion from the edge of the duct into the pipe (Fig. 3). Tests were also conducted using standard nozzles fitted at their end with pipes of varying lengths. Thus sound levels could be compared between the two cases differing by their expansion modes. Five models were measured.

3. Discussion of results

Considering the space allowed, it would be preposterous to attempt a discussion of the five cases studied, even to cite appropriate references. The essential question remains, what is the common factor leading to "quiet zones" in jet flow operation. One may try however to do it by a process of elimination in a tabular form. The five cases are: (1) The Silencer-Diffuser, (2) Sudden expansion into a short pipe, (3) Sudden expansion into a long pipe, (4) Standard nozzle with a short pipe, and (5) Standard nozzle with a long pipe. The likely causes contributing to noise attenuation suggested by various authors are: (a) stable shock wave structure, (b) spatial constriction reducing eddy size (i.e. compactness of noise source), (c) low velocity at the exit, (d) formation of vortex rings, (e) stretching of vortex rings, (f) entropy increase at an early stage of flow (i.e. total pressure drop), (g) orderly wave structure leading to an orderly flow, and (h) shortness of expansion.
5.1 Comparative Table

<table>
<thead>
<tr>
<th>Case</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>w</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>m</td>
<td>(a) oblique shock, spike, (b) constriction in anuli, (c) diffuser, (d) curved shock, (e) stretching in anuli, (f) very strong drop due to oblique shock, (g) stable, (h) short distance expansion.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>See Figs. 1 &amp; 2</td>
</tr>
<tr>
<td>2</td>
<td>m</td>
<td>w</td>
<td>w</td>
<td>s</td>
<td>m</td>
<td>s</td>
<td>m</td>
<td>s</td>
<td>(a) at exit, (b) due to pipe, (c) high subsonic, (d) blunt edge, (e) due to expansion, (f) at pipe entry, (g) at exit, (h) very short pipe.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>See Figs. 3, 4, 5, 6, 7, 8 &amp; 9</td>
</tr>
<tr>
<td>3</td>
<td>w</td>
<td>m</td>
<td>w</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>s</td>
<td>s</td>
<td>(a) at exit, (b) long pipe, (c) up to high subsonic, (d) blunt edge, (e) due to expansion, (f) at pipe entry, (g) unknown, (h) very short.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>See Figs. 4, 5 &amp; 9</td>
</tr>
<tr>
<td>4</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>(a) at exit, (b) short pipe, (c) up to high subsonic, (d) none, (e) none, (f) very little, (g) moderate, (h) moderate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>See Fig. 4</td>
</tr>
<tr>
<td>5</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>w</td>
<td>(a) at exit, (b) long pipe, (c) up to high subsonic, (d) none, (e) none, (f) very little, (g) at exit, (h) short.</td>
</tr>
</tbody>
</table>

A glance at this table indicates that most of the cases leading to sound attenuation are incorporated to a strong degree in the silencer. In the pipe flow, a large drop in total pressure (i.e. entropy increase) is due to the discontinuous expansion into the forming layer and formation of ring vortices contributing to the existence of the "quiet zone". A properly designed nozzle fitted into a pipe has no stable quiet zones. It appears that the energy spent against viscos forces producing vortex rings and stretching them reduces the energy available otherwise for sound productions. It may be noted that Bechert /3/ was one of the early researchers who indicated the important role played by ring vortex formation.

The abrupt occurrence of "silent regions" indicates that they are associated with the wave structures and formation of zones of silence.

References


/2/ W. Sieradzki Institute of Aerospace Mechanics, Warsaw Polytechnic (oral communication).

Fig. 1 A cross-section of the Silencer-Diffuser.

Fig. 2 Noise spectra of the Silencer-Diffuser. Note the sound level of 125 dBA, nozzle only.

Fig. 3 Minimum pipe length \( l_{\text{min}} \) against \( A/A^* \).

Fig. 4 A typical hysteresis loop (short pipe) & silent zone (long pipe).

Fig. 5 Noisy, silent, hysteresis loops & screech ranges for various area ratios \( A/A^* \), short & long pipes.
Fig. 6 Separated flow, sudden expansion, noisy operation.

Fig. 7 Shock re-enters the pipe from outside. Beginning of silent operation, short pipe expansion.

Fig. 8 Bubble-shaped shock after re-entry, short pipe expansion, silent operation.

Fig. 9 Long pipe expansion, silent operation, shock inside, orderly flow at exit.
PLANNED REVISION TO THE AMERICAN NATIONAL STANDARD FOR EXPLOSION AIRBLAST

Jack W. Reed
Sandia National Laboratories
Albuquerque, New Mexico 87185, USA

After five years of generally accepted usage, ANSI S2.20-1983, American National Standard "Estimating Air Blast Characteristics for Single Point Explosions in Air, with a Guide to Evaluation of Atmospheric Propagation and Effects," was reaffirmed in 1988. Recent experimental results now allow an improved and expanded revision, which has just begun. Our plan is to provide a slightly modified overpressure-distance function based on later hydrocode calculations, that reduces overpressures by approximately 4 percent. A TNT weight equivalence of 5/6 will be recommended for calculating overpressures from ANFO (Ammonium nitrate - fuel oil slurry), now widely used for economical blasting.

In a test program at Cape Canaveral, Florida, in 1979, about 300 small explosions were fired, mostly 45.4-kg TNT with some of 2.2-kg and 1134-kg TNT, with airblast pressure measurements at 200-m, 500-m, 1-km, 2-km, and 5-km distances near the four cardinal directions, as shown in Figure 1, and over flat terrain. A 152-m meteorological tower observed close-in weather conditions of temperatures and winds while upper air observations were made by radiosonde balloon. Directed sound velocities, V, temperature-dependent sound speeds plus directed wind components, versus height were calculated for each gage line direction. Propagated airblast amplifications and attenuations were correlated with sound velocity-height gradient or inversion strength, as shown in Figure 2.

With simple gradients or inversions in the observed 'boundary layer', there is a distinct relationship, beyond a yield-scaled 1-km range from 1-kg TNT, of airblast amplitudes that increase with increased inversion strength, δV, as shown by Figure 3, and decrease with gradient strength, -2V, as shown by Figure 4. No correlation was found with the height of the definitive velocity value. Geometric averages and standard deviations of amplitudes were calculated at five yield-scaled ranges, and statistically fitted to provide a fan function for overpressure versus distance curves in Figure 5 for selected sound velocity structures. It is assumed that peak-to-peak amplitudes are 35 percent greater than overpressures as determined by the Standard hydrodynamic explosion wave model. An empirical function fitting these data beyond the closest 200-m gages, where there was no clear weather correlation, was determined for amplitude, Pk, that

\[ P_k(ΔV,R) = P_k(Std,200m) \times (R/200)^{-α} \]  

where

\[ α = 0.02247 + 1.2066 \exp(-0.0356ΔV) \]
In complex weather conditions that can focus airblast, with decreasing, then increasing sound velocity with height above ground, focusing may hit or miss specific gages. Also, it was found that when there was no significant trend of sound velocity with height, as in near cross-wind propagations, small deviations from constant sound velocity could cause focusing but in unpredictable directions, right or left of the wind. In the opposite direction there would be strong attenuation. Cases with calculated focusing and cases with such zero-gradient uncertainty (50% probability of focusing) were added to other data on focused airblasts to provide a probability distribution for predictions. About 90 percent of measurements with possible focusing fall below the dog-leg upper curve in Figure 5.

Now consider the approximate acoustic ray path equation

\[ x_2 - x_1 = \left[ \left( z_2 - z_1 \right) / \left( \sqrt{v_2} - \sqrt{v_1} \right) \right] \left[ \left( v_p^2 - v_2^2 \right)^{1/2} - \left( v_p^2 - v_1^2 \right)^{1/2} \right] \]  (3)

where \( x \) and \( z \) are horizontal and vertical coordinates, respectively, \( V \) is directed sound velocity at subscripted levels, \( \Delta V = v_2 - v_1 \), and \( v_p \) is the ray characteristic velocity from Snell's Law that

\[ v_p = V / \cos \theta \]  (4)

for elevation angle \( \theta \). For yields, \( W \), other than 45.4-kg TNT, \( \Delta V \) must be defined for an atmospheric 'boundary layer' that is \( \left( z_2 - z_1 \right) / (W/45.4)^{1/3} \) deep, so that distance travelled by the refracted ray is also cube-root-of-yield scaled as required by similarity scaling laws that apply to single-point explosions. Thus the entire pattern of predicted overpressures in Figure 5 for 1-kt NE (nuclear explosion equivalent to 454-Mg TNT) may be shifted to yield-scaled distances so long as the depth of atmosphere considered for definition of \( V \) is also yield-scaled. For 1-kg TNT only 28 m of atmospheric depth is considered for propagations to 1400-m distance, while a 1-kt NE analysis requires 3300-m atmospheric depth for propagations to 108-km range.

Predictions for buried explosions are based on an overburden factor which represents the quantity of material blown away per weight unit of explosive, as shown in Figure 6. Measurements of buried test blasts have shown that overpressure transmissivity, defined as the ratio of observed to Standard overpressure, is small close to the surface zero point, but increases with distance and gradually approaches a nearly constant value beyond 916 m from 18-Mg HE bursts. For other yields, constancy is reached at cube-root-of-yield scaled distances.

This establishes an overpressure-distance point beyond which overpressure decreases inversely with distance as a quasi-acoustic wave rather than as a shock wave. This results from slow compression rates of blast emissions from buried shots with little high frequency attenuation. A similar prediction methodology applies to underwater explosions, but with a different source function based on yield-scaled burst depth.

Finally, following development of a method for calibrating the weakening of glass plates by selective abrasion, test specimens can be produced that fall at low overpressures of the order of a few hundred of pascals. These have been subjected to laboratory pressure chamber tests.
as well as extensive field exposure to large explosions. Test results will be used to refine window damage predictions as well as assessments of associated hazards from breaking and falling glass from atmospheric airblast propagations, where only fractions of percents of normal panes could be broken.

References


Fig. 1 Map of Blast Propogation Experiment at Cape Canaveral

Fig. 2 Example of Measured Explosion Airblast Amplifications in Four Directions
Fig. 3 Log-averaged Measurements, Explosion Overpressures Versus Distance Inversion Conditions

Fig. 4 Log-averaged Measurements, Explosion Overpressure Versus Distance, Gradient Conditions

Fig. 5 Overpressure Versus Distance Curved for a 1-kt Nuclear Free-Air-Burst Standard Explosion with Fitted Curves for Various Atmospheric Conditions

Fig. 6 Reference Overpressures (914 m from 18 Mg HE) Versus Overburden Factors for Buried Explosion Tests
ON THE SOUND FIELD OF ORGANIZED VORTICITY IN JET FLOWS

Ricardo E. Musafir

PIM/COPPE and DME/EE - Universidade Federal do Rio de Janeiro
C.P. 68503, 21945, Rio de Janeiro, BRAZIL

INTRODUCTION

In the past few years, it has been shown that situations exist where the sound field of turbulent jets, particularly when excited, shows a marked contribution from organized vortical structures [1-4]. The processes associated with the narrowband sound generation are both vortex pairing and the downstream development of the vortical structures between (eventual) successive pairings till they finally break down.

Theoretical models for these phenomena were developed [4-6], providing partial explanations for the measured far field directivity, whose dominant features, however, present significant variations. In some experiments a strong dependence on polar angle \( \theta \), that cannot be explained without the help of an exponential directivity factor, is seen [1] (also data of Kim reproduced in [6]); a dip for \( \theta \) between 50 and 80°, as shown in Fig.1, is sometimes observed [3,4,6]. In this work, some aspects of models concerning circular jets [4,5] are discussed, emphasis being placed on the azimuthal structure of the sound field. The initial stage of development of a model aiming to be consistent with both one and two microphone measurements is reported.

SOURCE MODELS

The approaches used in [4,5] are based on Lighthill's analogy, the adopted descriptions of the quadrupole source term leading to strong differences in the far field solution. We recall that the acoustic pressure in the far field of a point quadrupole \( T(t) \), located at \( x=0 \), in a medium at rest, is given by

\[
(4\pi xc^2) p(x,t) = n_0 r x_0^2 \frac{d^2 T_{ij}(t-x/c)}{dt^2}, \quad n = x/c \tag{1}
\]

To investigate the sound field due to the development of the vortical structures (or instability waves), Luerre and Crigthon [5] modelled the linear part of the momentum flux tensor \( T_{ij} = \rho v_i v_j \) for a nearly parallel flow, obtaining thus as dominant quadrupole components those with at least one axis in the flow direction \( x_k \), the model predicting no radiation at \( \theta = 90^\circ \). The source model is axisymmetric, a (second) dip being predicted due to azimuthal interference between the longitudinal \( T_{11} \) and the lateral source components. Its position depends on the Helmholtz number \( N_e \), equal to the ratio of jet diameter to acoustic wavelength. An exponential directivity factor, such as expected for an axially non-compact source appears, which is strongly dependent on the shape of the wave packet envelope [7].

Bridges and Hussain [4] modelled sound generation by vortex pairing, expressing \( T_{ij} \) in terms of vorticity as proposed by Moring. Their source is represented by a single point quadrupole \( T_{ij} \), which contains only longitudinal components, related by \( T_{22} = T_{33} = -2T_{11} \). This dependence leads to a zero in the directivity (at \( \theta = 55^\circ \)) providing thus a totally different explanation for its existence.

The dip at 90° predicted in [5] is not seen in the experimental data. Indeed, the existence of only one diagonal component in that model violates the incompressibility condition, expressed in the compact limit by
\[
\text{tr}(\mathbf{T}) = \int \text{tr} \left( \mathbf{T} \right) dV = 0
\] (2)

A significant feature of both models is that they predict an instantaneous axisymmetric sound field, something that cannot be verified by one microprobe measurements.

To the author's knowledge, the only available experimental data of far field azimuthal correlations for excited jets are those of Juvé et al. [3,4], for Mach number \( M = 0.4 \). They show clearly that for \( M = 0.14 \), the first subharmonic of the excitation frequency, for some angles \( \theta \) significant changes occur in the plots of \( r(\theta, \phi, M) \), the normalized real part of the cross-spectrum of pressure signals received by two observers \( x \) and \( x' \) equally distant of the jet exit, against azimuthal separation \( \phi \) (see Figs. 2 and 3).

For \( \theta = 90^\circ \), although the excited power spectral density shows a narrowband increase of approximately 3 dB (Fig. 1), the small changes observed in \( r(90, \phi) \) — the He dependence will be henceforth omitted — are not consistent with an axisymmetric source, which would produce a large positive increase in \( r(90, 90) \). Thus, at least for this case, the axisymmetric model is not appropriate.

**A More General Point Source Model**

Far field correlations due to a statistically axisymmetric (SA), but otherwise general, point quadrupole were discussed in [8]. It is known that the decomposition of the field of such a source in azimuthal modes \( A_m \cos(m \phi) \) yields only modes \( m = 0, 1, 2 \).

In the vortex pairing model, the dip in the directivity is due both to the phase relationships between the (longitudinal) components and to the absence of the lateral \( T_{11}, T_{13} \) components, which are alone responsible for the azimuthal mode \( m = 1 \). Since the main effect of the excitation in the correlation data in [2,3] can be roughly described as a reduction in importance of mode \( m = 1 \) for low values of \( \theta \) at \( M = 0.14 \) which, as discussed above, cannot be ascribed to an axisymmetric source, it is thought that a more adequate point source model for the vortical structures should, as in [4], have \( x_1 \) as a fixed principal emission direction, but allowing for any set of average phase relationships between the three longitudinal components consistent with the incompressibility hypothesis (2), besides being SA. These conditions imply that for the description of source average properties only diagonal components are needed, although the model requires, in general, \( T_{23}(t) \neq 0 \).

The needed phase relationships are the same that would be obtained for the monochromatic (non SA) diagonal point quadrupole \( T'(t) \), whose components have zero mean value and are expressed, in adimensional form, as

\[
T_{12}'(t) = \cos(\omega t), \quad T_{13}'(t) = \cos(\omega t + \alpha), \quad T_{11}'(t) = \sqrt{\beta + 1} \cos(\omega t + \beta)
\] (3)

where \( \alpha, \beta \) and \( B \) are constants to be determined.

Insertion of (3) in (2) yields, after squaring and averaging:

\[
\cos \alpha = \frac{(B-1)/2}{B, B \in [-1, 3]}
\]

The source adimensional directivity \( D_0(\theta) - n_i \eta_j (T_{ij})^2 \) is then expressed as

\[
D_0(\theta) = 1 + B \cos^2 \theta - (B + 3)/4 \sin^2 2\theta
\] (4)

The location of the dip is seen to depend on \( B \), the only adjustable parameter of the model, which is a particular case of the one discussed in [8].

For \( B = 3 \), the axisymmetric case studied in [4] is recovered.

Decomposing the signal into a random component (subscript 1) and one due to the vortical structure alone (subscript 0), the resulting \( r \) can be expressed as

64
\[ r(\theta, \phi) = \frac{\xi_1 + \xi_0}{1 + \xi} \]  

where \( \xi = c(\theta, He) \) is the ratio of the power spectral densities of the two components, \( P_x(x, He)/P_y(x, He) \).

Although the data show the additional source to be stationary, there is no reason not to consider the embedding in the jet flow, which can be approximated by a plug flow with Mach number equal to 0.6M [8]. This introduces the additional dependence of \( r_0 \) on He and He, through the mean flow transmission coefficients \( F_{ij} \), which replace \( n_i n_j \) in (1).

**COMPARISON WITH EXPERIMENT**

**Correlations** For calculating \( r \) it was assumed that excitation did not alter \( r_0 \), unexcited jet correlation data being then used as input in (5). That this hypothesis is not strictly true is evidenced by the cross-spectral density data for \( \theta = 0^\circ, \phi = 180^\circ \) shown in [2].

The values of \( \xi \) were taken from the plot of \( \Delta(\theta) = 10 \log(1 + \xi) x \theta \) in [3] (reproduced in Fig. 1) as 3, 1, 1.2 for \( \theta = 30^\circ, 45^\circ \) and \( 90^\circ \) respectively. \( B \) was chosen as 0.7 to match the measured value of \( r(90, 90) \). The allowance of in uncertainty of \( \pm 1 \) dB in \( \Delta(90) \) affected \( B \) by less than 5%.

Due to the low value of \( B \), 0.14, the low frequency \( F_{ij} \) [9] were used to model the embedding in the jet flow.

The predicted curves are shown in Fig. 3, together with the measured \( r \) for clean and excited conditions. For \( \theta = 90^\circ \) the \( B = 3 \) case was also included.

The qualitative agreement is very good although the model overestimates \( r \) as \( \phi = 180^\circ \), this effect increasing as \( \theta \) is reduced.

Including only longitudinal components, the model predicts \( r_0(90, 180) = 1 \), so that while the discrepancies observed for \( \theta = 90^\circ \) and \( 45^\circ \) could be credited to the use of clean jet data for \( r_0 \), this cannot be done for \( \theta = 30^\circ \); even if \( r_1 = 0 \) and \( c(30) \) where reduced from 3 to 2, \( r(30, 180) = 0.67 \) would be obtained, still higher than the measured 0.54.

![Fig.1 - Difference in narrowband directivity due to vortical structure. Experiment [3]: 0; Theory: —.](image1)

![Fig.2 - Geometry for azimuthal correlations.](image2)

![Fig.3 - Azimuthal correlations, He = 0.14, M = 0.4. Experiment [2,3]: — clean jet; — excited. Theory —, B = 0.7, unless otherwise indicated.](image3)
It should be noted that the present model stands for the compact limit of a ring source with a strong azimuthal coherence that could be expressed by a complex-valued function. The neglect of the ring effect may be responsible for the overprediction. Another point to be mentioned is the total omission of the lateral $T_{12}$, $T_{13}$ components, which needs further discussion. They would be, however, more important around $\theta = 45^\circ$ than at $\theta = 30^\circ$.

**Directivity** If the dimensional far field directivities, denoted by $D^1$, $D^1$, are known, $\Delta$ can be calculated as

$$\Delta = 10 \log \left\{ 1 + \epsilon (\phi) \right\}$$

where $\epsilon = 1 - 0.64 \cos \phi$.

This form fits well the data extracted from the excited jet power spectral densities shown in [2]. The calculated $D^1$ was obtained from correlation modeling for $\phi = 0$.

The theoretical curves for $B = 0.7$ and 3 are shown in Fig. 1 with the measured values. The position of the dip is well predicted, but for low $\theta$, a strong disagreement is observed, what suggests the existence of a $\theta$-dependent amplification factor, related to axial non-compactness as modelled in [5]. This factor, modifying only amplitude, would not affect $\Delta$. Due to non-zero $M$, the zero in $D^1$ for $B = 3$ ($\theta = 55^\circ$) is seen in $D^1$ at $\theta = 39^\circ$.

**CONCLUSION**

A point source model for vortex noise, depending on a single adjustable parameter, was developed, which is expected to be relevant to sound generation both by vortex pairing and by instability waves development. Unlike previous ones, this model predicts reasonably well phase characteristics that can only be verified by correlation measurements. The disagreement with the experimental directivity is supposed to be due to the axial structure of the wave packet, neglected herein. It is expected that extension to a ring source will improve the comparison with correlation data. This extension and the relationship of the present model to vorticity dynamics and to Lighthill's quadrupole source term will be discussed elsewhere.

**REFERENCES**

SOUND GENERATION IN UNSTEADY VORTICAL FLOWS

A. T. Fedorchenko

Moscow Physical-Technical Institute,
141700 Dolgoprudnyi, Moscow Region, USSR

1. The investigation of sound generation phenomena in unsteady subsonic flows is of great significance for gasdynamics wave problems. A detailed analysis of the role of vortex sound generation and upstream acoustic feedback is needed when investigating shear-flow instabilities.

In considering such problems it is often necessary to use complex nonlinear systems of differential equations governing spatial wave processes in compressible media (the Navier-Stokes equations or corresponding equations for ideal-gas flows).

At the present time, despite of the wide spread of gasdynamics numerical methods, the volume of results in the above-mentioned problems is rather limited. This fact can be accounted for by some methodical handicaps. For instance, the difficulty of specifying the set of conditions on a permeable boundary with locally-subsonic shear flow. Besides, some difficulties arise because of the necessity of accurate resolution of acoustic processes by a numerical scheme.

We have considered some general problems of the numerical investigation of wave processes in viscous and ideal gases by applying finite-difference methods. A number of nonlinear permeable-boundary models has been proposed for the numerical solution of the wide class of unsteady gas-dynamics problems. Different ways of active controle of the normal acoustic modes by these boundaries (e. g. for sound abscorption) are discussed. A special case when intense vortices cross the exit of computational domain is investigated. Some nonlocal boundary conditions are proposed to minimize the effects of sound generation on this boundary element [1].

The suggested models and algorithms are realized in a number of important problems where sound generation by subsonic and transonic flows in ducts of variable crosssections are studied on the basis of the Navier-Stokes equations.

In particular, the effects of acoustic self-excitation in viscous subsonic flows in a plane duct with symmetric sudden expansion (in the ratio 1:3) are detected [2, 3], when the flow instability develops with increasing Reynolds number. A special attention is paid to sound-vortex interactions near the sharp wall-step edges, to the formation and further evolution of coherent vortex structures downstream the expansion section (with sound generation). Some formulae are derived for the resonant self-excitation conditions, primarily depending on duct geometry, Mach number of centreline flow, and thickness of boundary layers in inlet section. These
formulae in a good agreement with the results of calculations.

2. It is instructive to supplement the methods of the numerical integration of Navier-Stokes equations with the procedure of extracting the acoustic components of flow parameters, in particular, for unsteady mean flows.

A theoretical model of sound generation and propagation in unsteady subsonic flows with nonuniform entropy field is proposed [4, 5]. This model substantially differs from the known approaches of Lighthill, Powell and Howe.

The equations describing unsteady subsonic ideal-gas flows in some spatial domain $G$ are

$$F(p, q, s) = 0,$$

$$\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{u} + \nabla p = 0,$$

$$\frac{\partial q}{\partial t} + \nabla (q \mathbf{u}) = 0,$$

$$\frac{\partial s}{\partial t} + \nabla (q s \mathbf{u}) = 0,$$

where $p$ is the pressure, $q$ is the density, $s$ is the entropy, $\mathbf{u}$ is the velocity and the tensor $\mathbf{M}_i = (q \mathbf{u}_i \mathbf{u}_j)$ (or $\mathbf{M}_{ij} = q u_i u_j$ if $i, j = 1, 2, 3$).

We assume that

$$\mathbf{u} = \mathbf{u}_v + \mathbf{u}_\alpha, \quad p = p_v + p_\alpha, \quad q = q_v + q_\alpha, \quad s = s_v + s_\alpha,$$

where index $v$ denotes unsteady-mean-flow parameters and $\alpha$ denotes acoustic disturbances. Then we demand for $Z_v(f, t) = \{\mathbf{u}_v, p_v, q_v, s_v\}$ to satisfy the following equations

$$F(p_v, q_v, s_v) = 0,$$

$$\frac{\partial q_v}{\partial t} + \nabla (q_v \mathbf{u}_v) + \nabla p_v = 0,$$

$$\left( \frac{\partial q_v}{\partial s_v} \right)_v \frac{\partial s_v}{\partial t} + \nabla (q_v s_v \mathbf{u}_v) = 0,$$

$$\left( \frac{\partial q_v}{\partial s_v} \right)_v \frac{\partial s_v}{\partial t} + \nabla (q_v s_v \mathbf{u}_v) = 0.$$

We also assume that the mean-flow solution of (6)-(9) is known and the acoustic field $Z_\alpha(f, t) = (\mathbf{u}_\alpha, p_\alpha, q_\alpha, s_\alpha)$ does not affect this mean flow in $G$, this condition not prohibiting sound-vortex interactions outside $G$, for example near the walls - especially near the sharp edges where the viscosity effects cannot be negligible.

It is easy to show that in a medium described by (6)-(9) sound-wave propagation is impossible despite $a_\alpha = (3p_\alpha/3s_\alpha)^{1/2}$ is finite.

Substituting (5)-(9) into (1)-(4) we obtain the following general nonlinear acoustic equations with sound sources $\mathbf{m}$ and $q$:
\[ F(p_v + p_\alpha, q_v + g_\alpha, s_v + s_\alpha) = 0, \]
\[ \frac{\partial \tilde{w}_\alpha^*}{\partial t} + \nabla \tilde{m}_{\alpha}^* + \nabla p_\alpha = 0, \]
\[ \frac{\partial q_\alpha}{\partial t} + \nabla \tilde{w}_\alpha^* = m, \]
\[ \frac{\partial \psi_\alpha^*}{\partial t} + \nabla \tilde{v}_\alpha^* = q, \]

where
\[ \tilde{m}_{\alpha}^* = \tilde{w}_\nu \dot{u}_\alpha + (\tilde{w}_\alpha^* \dot{u}_\nu) + (\tilde{w}_\alpha^* \dot{u}_\alpha), \quad \tilde{w}_\nu = g_v \dot{u}_\nu, \]
\[ \tilde{w}_\alpha^* = q_v \dot{u}_\alpha + g_\alpha \dot{u}_\nu + g_\nu \dot{u}_\alpha, \quad \psi_\alpha^* = g_v s_\alpha + g_\alpha s_v + g_\nu s_\nu, \]
\[ \tilde{v}_\alpha^* = s_v \tilde{w}_\alpha^* + s_\alpha \tilde{w}_\nu + s_\alpha \tilde{w}_\alpha^*, \]
\[ m = -\frac{\partial q_v}{\partial t} + \frac{\partial q_v}{\partial t} \frac{\partial s_v}{\partial t} = -\frac{1}{\alpha_v^2} \frac{\partial p_v}{\partial t}, \]
\[ q = -\frac{\partial q_s}{\partial t} + \frac{\partial q_s}{\partial t} \frac{\partial s_v}{\partial t} = -s_v \left( \frac{\partial q_v}{\partial t} \frac{\partial p_v}{\partial t} \right) = m s_v. \]

We can derive from (10)-(13) the second-order equations with sources \( Q \) and \( \bar{R} \).

\[ \frac{\partial^2 q_\alpha}{\partial t^2} - \Delta p_\alpha - \nabla_1^2 \tilde{m}_{\alpha}^* = Q, \]
\[ \frac{\partial}{\partial t} \left( \frac{1}{\alpha_v^2} \frac{\partial \tilde{w}_\alpha^*}{\partial t} \right) - \nabla_2^2 \tilde{w}_\alpha^* + \frac{\partial}{\partial t} \left( \frac{1}{\alpha_v^2} \nabla p_\alpha - \nabla q_\alpha + \frac{1}{\alpha_v^2} \nabla \tilde{m}_{\alpha}^* \right) = \bar{R}, \]
\[ Q = \frac{\partial m}{\partial t} = -\frac{\partial}{\partial t} \left( \frac{1}{\alpha_v^2} \frac{\partial p_v}{\partial t} \right), \quad \bar{R} = -\nabla m = \nabla \left( \frac{1}{\alpha_v^2} \frac{\partial p_v}{\partial t} \right), \]
\[ \nabla_1^2 = \text{div \ div}, \quad \nabla_2^2 = \text{grad \ div}, \quad \Delta = \text{div \ grad} \]

69
Using exact nonlinear equations (10)-(13), it is not difficult to deduce the nonlinear equations of acoustic-energy balance with a source proportional to \( m \).

Linearizing (10)-(13), we obtain the same system of equations but with \( \mathcal{M}_{\alpha\alpha} = \langle \mathcal{W}_{\alpha} ; \mathcal{U}_{\alpha} \rangle + \langle \mathcal{W}_{\alpha} ; \mathcal{U}_{\alpha} \rangle \), \( \mathcal{W}_{\alpha} = q_{\alpha} \mathcal{U}_{\alpha} + q_{\alpha} \mathcal{U}_{\alpha} \), \( \mathcal{W}_{\alpha} = q_{\alpha} \mathcal{W}_{\alpha} + q_{\alpha} \mathcal{W}_{\alpha} \), \( \xi_{\alpha} = 3 \mathcal{W}_{\alpha} + 3 \mathcal{W}_{\alpha} \mathcal{W}_{\alpha} \) standing for \( \mathcal{M}_{\alpha\alpha}, \mathcal{W}_{\alpha}, \mathcal{W}_{\alpha} \) and \( \xi_{\alpha} \), respectively. It should be noted that for steady mean flows the sources \( m, q, \mathcal{W} \) and \( Q \) are equal to zero and the linearized equations can be reduced to Blokhintsev's type.

We propose also the procedure for estimating \( Z \) using the known solution obtained within the framework of incompressible-fluid approximation[5].

References

NEW TRENDS ON SINGLE NUMBER INDICES AND SHORT METHODS FOR ACOUSTICAL INSULATION
NEW TRENDS ON SINGLE NUMBER INDICES AND SHORT MEASURING METHODS OF ACOUSTICAL INSULATION.
REFERENCE CURVE METHOD VERSUS WEIGHTING NETWORK METHOD

A. Moreno

INSTITUTO DE ACUSTICA
Serrano, 144 - Madrid 28006. Spain

INTRODUCTION

The main topics to deal with in this structured session are: rating procedures (reference curve and weighting network procedures; extension of the frequency range; objective ratings and people reaction; correlation between laboratory and field results; cumputing and prediction), measuring procedures (short tests; absorption measurement; impact sources), consultant point of view, and influences on international standards.

Even a short introduction to each topic should extend beyond the official four pages. Therefore I'll focus my contribution on basic aspects of the reference curve procedure compared to the weighting network procedure.

REFERENCE CURVE OF WEIGHTING NETWORK PROCEDURE?.

The main idea of the reference curve procedure (RCP) is that partitions with an appropriate rating value reduces any incident noise to a sufficient low level. It is assumed that the limit corresponds to a 1' brick wall, which transmission loss curve (somewhat smoothed and modified) was stated as "reference curve". It was thought to be satisfactory for common noises in buildings at the time.

Though the weighting network procedure was initially proposed as an approach to RCP, some time later merged as an independent and well founded method. The basis was the ability of dB(A) ratings to rate almost all kinds of noise. Loudness is then the main characteristic considered.

Several definitions of "insulation in dB(A)" have been proposed. The most interesting (and different) correspond to the following expressions

(1) \[ R_A = - 10 \log \frac{10^{(A_l+R_l)/10}}{10^{A_l/10}} \] (introduced by Gosele is the most widely used)

(2) \[ R(A) = 10 \log \frac{10^{(A_l+R_l)/10}}{10^{A_l/10}} \] (used by L. Taiba)

(3) \[ R_A, tr = 10 \log \frac{10^{(L_u+R_l)/10}}{10^{L_u/10}} \] (used in Nordtest; Lui is traffic noise spectrum, A weighted, normalized to 0 dB(A))

(M.C. Gomperts proposed on 1974 the first definition of insulation in dB(A) as an independent index, but his procedure was not used).
Many authors have found a good correlation between the above quantities and \( R_w \) (or \( \lambda \) in the old ISO 717) for usual partitions in laboratory and in the field \((1, 2, 3, 4, 5, 6, 7)\). Investigations on the intrinsic features of both ratings by deterministic and Monte Carlo procedures conclude that this correlation is mainly due to the nearness between the A weighting curve and the ISO reference curve \((8)\). This conclusion has been confirmed very recently and an analytical expression connecting both quantities has been derived by Parme (9, and this session).

**Approach by deterministic simulation.**

This approach is based on idealized shapes of transmission loss curves \((\text{tlc})\) derived from laboratory results, applying rules of simplification and smoothing. Six main shapes, formed by one or several segments, are found to be the most typical. They can be represented schematically by \(/\), \(/\), \(/\), \(\cdot\), \(/\), \(/\). Within a group the values of the slopes and the positions of the bending frequencies define a particular curve. Most actual tlc can be included into the three first groups: about 70\%. The following three groups gather about 20-25\%.

Indices \( R_w \), \( R_a \) and the arithmetic mean \( R \) of a tlc have a common property with regard to translations. According to this property for any pair of transmission loss curves \( R_1(f) \) and \( R_2(f) \) related by an arbitrary translation along the level axis \( (R_2(f) = C + R_1(f)) \), where \( C \) is a constant) the following relations hold: \( R_{2-RW} - C \), \( R_{2-RA1} = C \). Only the shape (profile) of a tlc seems important in this correlation, the main parameter being the mean slope. The influence of the shape parameters on \( R_w \) and on \( R_a \) follows "parallel" evolutions, independently of the main shape. Differences \( R_{2-RW} - R_{2-RA1} \) are positive, the maximum corresponds to tlc of shapes \(/\) and \(/\), for high slopes. Figure 1 includes \( /R_a \), \( /R_w \) and \( R_{2-RW} - R_{2-RA1} \) for the idealized shapes \(/\), \(/\) and \(/\) ordered in that way from left to right in the frequency scale, according to the values of the bending frequencies \((f_1 \text{ for shape } /\text{ and } f_2 \text{ for shape } /\)) The joining corresponds to shape \(/\), that can be included in the group \(/\), with \( f_1 = 100 \text{ Hz} \) or in the group \(/\), with \( f_2 = 3150 \text{ Hz} \). Similar results are found with the other mentioned shapes. This behaviour indicates that a "in detail" correlation exists between \( R_w \) and \( R_a \).

The statistics of differences \( R_{2-RW} - R_{2-RA1} \) for 1825 tlc, distributed "uniformly" into the above main shapes, looks like a probability density function of a random variable skewed to the right. This figure being independent of the overall level of \( R_a \) (or \( R_w \)) the following expression was suggested:

\[
R_w = R_a + d + \delta
\]

where \( d \) is a measure of differences \( R_{2-RW} - R_{2-RA1} \), and \( \delta \) can be accounted for by a Weibull function centred around \( \delta \). \( R_w \) and \( R_a \) are then related by a translation in statistical sense.
Results by Monte Carlo simulation.

The above results, particularly the expression \((1)\), have been confirmed by means of Monte Carlo simulation models \((10)\). Two types of models have been developed. The first type corresponds exactly to the simplified shapes of transmission loss curves mentioned previously. The second type of models produce \(\text{tlc}\) by means of four uniform random variables: number of segments, position of the bending frequencies, slopes of each segment and sign of slopes. As in the first type, further unevenness, dips etc can be added on a second step. All models include security conditions to ensure that any \(\text{tlc}\) may likely represent an actual transmission loss curve, then conforming an adequate sample of \(\text{tlc}\), a basic condition of Monte Carlo simulation. Models of first type give for \(d\) values around 0.6 dB, few sensitive to uneveness or dips (within the range considered, up to 9 dB); peakedness and skewness do not vary significantly from one model to another. Models of the second type give values for \(d\) neatly higher, 1 to 1.6 dB, and distribution functions also more sensitive to additional uneveness and dips.

Figure 2 gives RW-Ra functions obtained with models of the first \((A3)\) and second type \((B3)\) that include unevenness and dips.

For practical purposes it is more convenient to work a sample of \(\text{tlc}\) in accordance with actual situations. This is a rather complicated item. Nevertheless a first approach has been intended introducing simple assumptions in models of the first type by means of appropriate weightings of the random variables involved \((\text{model A3W})\). Positive slopes are taken following a normal distribution around 6 dB/octave and a standard deviation of 2.5 dB. For the mean value of the second positive slope 12 dB/oct was taken. Normal distributions have been used for the first and second bending frequencies around 200 Hz and 1600 Hz respectively. Some other minor amendments have also been introduced in uneveness and dips. The resulting \((d + \delta)\) function is also represented in Figure 2, for model A3W. The mean is seen to be 1.2 dB and the standard deviation 0.8 dB. These so high dispersions make unpractical in standards and regulations the coexistence of both ratings.

Can the reference curve and the weighting network procedure give equal ratings?.

It is intuitive that the choice of a reference curve equal to the weighting network may give equal ratings. The above mentioned Monte Carlo models permit the computation of the corresponding distribution functions. A mean difference of about 0.5 dB and a standard deviation of about 0.5 dB are obtained for models of several types \((B)\). (Figure 3 corresponds to model A3W). This is a promising result mainly for control purposes.

INFLUENCE OF THE SPECTRUM OF THE INCIDENT NOISE.

It is generally assumed that RW gives very optimistic values when the incident noise is traffic noise and in consequence it is a poor estimator of the sound protection afforded by façades. Similarly RA is inadequate, mainly because it involves incident pink noise. To solve this lack RA, that consider A-weighter traffic noise as incident noise, has been
proposed and at that time it is being widely accepted (11) and constitutes a common practice in French regulations. The histogram of \(R_{W-RA, tr}\), shown in Figure 4 for model A3W, accounts for the failure of \(R_W\) to this aim.

OTHER WEIGHTINGS.

Some studies have analyzed the ratings resulting when the A-weighting curve is substituted by an appropriate curve in an attempt to introduce noisiness to account for the annoying effects of noise (12), but more research should be done.

Acknowledgements. This work has been supported by the program PR84-0118 of the CAICYT (Spain).

(4) L.N. Talbo et al. 1982. JASA 72,449-459.
(6) A. Moreno et al. JASA 75, 476-478.
(12) S.L. Yaniv and S.L. Flinn 1976. Special Publication 499 of NBS (USA)
PREDICTION OF SINGLE NUMBER RATINGS FOR AIRBORNE AND IMPACT SOUND INSULATION

Eddy Gerretsen

TNO Institute for Applied Physics (TPD)
P.O. Box 155, NL-2600 AD Delft, The Netherlands

INTRODUCTION

Requirements for the airborne and impact sound insulation in buildings are given in single number ratings like $D_{n,t}$ and $I_{n,t}$ [1]. Besides these single number ratings according to ISO, used in several countries, there are other rating systems used in other countries, e.g. [2]. All these single number ratings are, however, based on the same quantities in frequency bands.

In the design stage of buildings it is generally desirable to be able to predict the acoustical performance of the design and to predict the effect of design changes. For this purpose prediction models have been developed which describe the airborne and impact sound transmission between rooms, taking all the relevant aspects into account as far as possible [3]. These models perform calculations in octave bands from which the relevant single number rating can be deduced. Under certain restrictions it is also possible to simplify the calculation scheme to a direct prediction of the single number ratings, either by simple calculations or with the use of nomograms. After a short description of the model for the airborne and impact sound transmission, the possible ways of simplifying will be indicated.

CALCULATION MODEL

The sound transmission between two rooms consists of the transmission through several paths, each characterized by an impact on a construction at the source side and radiation of sound by a construction in the receiving room. The transmission of construction noise (vibrations) through the construction is attenuated, particular at the construction junctions. For airborne noise the transmission via a path from construction $i$ to construction $j$ can be described by the flanking sound reduction index $R_{ij}$ (radiated sound power to impacted sound power on the partition):

$$R_{ij} = W_{Ri} + W_{Rj} + W_{D_{ij}} + W_{D_{ji}} + 10 \log S_o / \sqrt{S_i S_j} \quad (1)$$

The total sound transmission is then given by the level difference $D_{n,t}$, normalized to a reverberation time of 0.5 s:

$$D_{n,t} = -10 \log \{10^{-R_p} + 10^{-R_{ij}}/10\} + 10 \log V/3S_o \quad (2)$$

$R$ is the sound reduction index of the constructions, $D_v$ the vibration level difference over the junction, $S$ the areas of the constructions and $V$ the volume of the receiving room. The index $o$ indicates the partition.

In case of homogeneous walls and floors twelve transmission paths are to be considered besides the direct transmission through the partition [3]. More recently the model has been extended to double wall constructions, both for partitions as for flanking walls, in which case a much larger number of transmission paths can be constructed. These can however be reduced to the same twelve paths by adjusting the vibration level difference $D_v$ at the junction to the larger number of paths.
Figure 1: Transmission paths between two rooms.

For impact sound the transmission is much the same, only the impacting of the floor at the source side is different. This impacting is governed by the admittance of the floor Y, so for the transmission path 1 the resulting impact noise level, referenced to an absorption area of 10 m² is [4]:

\[ L_{n1} = L_r - \alpha V_{1} - \alpha V_{21} + 5 \lg \left( S_1 \times Y_1 \times 2 \pi f / \left( S_1 \times e_1 \right) \right) \ - 37.7 \] (3)

and the total impact noise level \( L_{nt} \), normalized to a reverberation time of 0.5 s follows from:

\[ L_{nt} = 10 \lg \ \Sigma 10^{L_{n1}/10} + 10 \lg V/3S_\alpha \] (4)

\( L_r \) is the known force level of the ISO tapping machine, \( m \) the area mass and \( e \) the total damping of the construction. All other parameters have the same meaning as with airborne noise transmission, and ISO reference levels are used.

In principle, all quantities mentioned are a function of frequency, except the areas. So the calculations should be performed for frequency bands. Due to the inherent inaccuracy of predictions there seems not to be much use in calculating for one-third octave bands; octave bands can be more accurate and give sufficient detail.

**SINGLE NUMBER RATING**

As with measurement results, the single number rating can be deduced from the calculated insulation in octave bands. This, at least, is the case for the Dutch rating system [2], where the single number is determined from results in octave bands, either directly measured as such or deduced from measurements in one-third octaves. The ISO rating system however, prescribes results in one-third octave bands - even for field measurements - in order to deduce a single number. Those single numbers can therefore only be estimated on the basis of octave band predictions.

For airborne sound transmission in essentially homogeneous buildings the vibration level difference is hardly depending on frequency. The only quantity in Formula 1 that is frequency dependent then is the sound reduction index. This opens the possibility to use single number ratings for the sound insulation of the elements and apply Formulas 1 and 2 to them.

Although this is not mathematically exact, the results are quite good in most situations (N.B. The result would be exact if the single number rating is based on weighted level differences instead of the ISO system), in fact, this approach was taken with the first version of the model [3], resulting for 75 situations in a prediction of the single number rating which was, on average, correct with a standard deviation of 1.5 dB.
For impact noise transmission no efforts as yet have been taken to simplify in this way, though in principle the same sort of approach would be possible.

**SIMPLIFIED PREDICTION SCHEMES FOR SINGLE NUMBER RATING**

In order to further simplify the calculations the number of paths which have to be taken into account should be reduced. At each junction between the partition and a flanking wall three transmission paths are to be considered, the description of each mainly varying in the vibration level difference to be applied. Formulas 1 and 2 could be rewritten in such a way that for each flanking construction only one term is to be considered.

\[ R_{r1=1} = R_1 + 10 \log S_0 / \sqrt{S_{1s} S_{1r}} + D_{v*} \]  \hspace{1cm} (5)

\[ D_{nt} = -10 \log \left( 10^{-R_{r1}} + \sum 10^{-R_{r1}/10} \right) + 10 \log V / 3S_0 \]  \hspace{1cm} (6)

\( D_{v*} \) is an effective vibration level difference, which takes into account the combined effect of transmission via three paths and thereby depends also on the type of constructions considered and somewhat on the area ratios involved. For some construction types and areas ratios common in dwellings, values for \( D_{v*} \) have been deduced from the complete model of which figure 2 gives an example.

![Figure 2: Effective vibration level difference for the combined effect of three transmission paths at a junction, as function of construction types and area mass ratio.](image)

Since for homogeneous constructions both the sound reduction index of the constructions as well as the vibration level difference are mainly determined by the area mass of the constructions, a further reduction of the calculations is possible. Figure 3 gives an example of nomograms for the determination of the flanking transmission in single number ratings from area mass and type of constructions only. This figure is based on a large amount of calculations using the complete octave band model. For a flanking wall, the flanking sound reduction index can be deduced directly from the figure in relation to the sound reduction index for direct transmission (left scale).
The right scale gives a term $f_t$ which is to be added for all four flanking constructions, resulting in the apparent weighted sound reduction index by $R_w = R_{ow} - 10 \log(1 + f_t)$. Some adjustments (e.g., curve shifts) should be applied to take area ratios and junction type into account.

**Figure 3:** Nomogram to determine the contribution to the sound transmission of a flanking wall.

**CONCLUSIONS**

For airborne sound transmission in traditional building constructions it is possible to predict the single number index in situ in a rather simple way from knowledge of the construction types and masses or from the single number ratings for the partition and flanking structures. Though also possible for impact noise no attempt has been made to deduce such a scheme. Predictions in octave bands, however, have a larger flexibility and allow the prediction of the several single number ratings that (still) exist. Nowadays, use of a PC makes those calculations quite easy to perform, decreasing the need for simplified schemes.

**REFERENCES**

THE DEVELOPMENT OF AN I.S.O. SHORT TEST METHOD FOR THE MEASUREMENT OF AIRBORNE SOUND INSULATION IN BUILDINGS

Robin K. Mackenzie

Department of Building, Heriot-Watt University, Riccarton
Edinburgh EH14 4AS

INTRODUCTION

The development of the short test method for the measurement of airborne sound insulation in buildings has now progressed to the stage of a formal I.S.O. Draft Proposal.

The working group, established in 1982 by I.S.O. Technical Committee 43/SC.2, have considered a wide range of submissions and a Draft Proposal was finally approved in January 1988.

ORIGINAL CONCEPTS

The original concepts established at the outset were as follows:-

SCREENING TEST The short test method should give a result which relates to I.S.O. 717 values so as to indicate whether or not a full I.S.O. 140 measurement is required.

INSTANT RESULT The results should be available in the field.

SIMPLE The method should be suitable for use by acoustically unsophisticated persons and the equipment should be simple and inexpensive.

REPEATABILITY The test should have a repeatability (r) and reproducibility (R) within ± 2 dB.

SOURCE Pink noise within 90-3500 Hz should be used. (However, the advantages of using a shaped random noise were acknowledged).

LOUDSPEAKERS Only one loudspeaker placed in the corner of the room, facing outward, should be used.

MICROPHONE Measurements at six or seven positions arithmetically averaged were considered. Using an L_e meter in a controlled 'walkabout' was an alternative consideration.

BACKGROUND NOISE Should be greater than 5 dB below the receiving room level for a viable result.

MEASUREMENTS 'A' weighted measurements should be made in each room using a non-integrating or integrating sound level meter.

NORMALISATION By measurement of RT using T/0.5.

Many of the original concepts were influenced by the provisions of A.S.T.M. E 597-77T "Tentative Recommended Practice for Determining a Single-Number Rating of Airborne Sound Isolation in Multi-unit Building Specifications". The procedure set down was perhaps too sophisticated for a simple test or perhaps suffered from the variability in building control inspection in the U.S.A. from State to State. Whatever the reason, this method failed to be adopted as a nationally accepted procedure in America and the subsequent
decisions of the I.S.O. Working Group were heavily influenced by the need to maintain simplicity in the measurement.

In parallel with the meetings of the working group, a major research programme was undertaken between 1982 and 1985 (Refs. 1-3). The principal objective of this research was to study the influence which the shape of the noise source had upon the correlation between the $D_{1W}$ and the 'A' weighted level difference. In addition to establishing an optimum source shape, the effect of normalisation upon the accuracy was investigated and ultimately a study of measurement procedure was carried out.

DRAFT PROPOSAL

INTRODUCTION
The International Standard describes a short test method which can provide an estimate of the single number rating which would be obtained by using the more complicated procedures specified in I.S.O. 140-4 and I.S.O. 717-1.

PRINCIPLE
The method is based upon the determination of the difference of the 'A' weighted sound pressure levels between adjacent rooms without regard to the paths of transmission.

EQUIPMENT
Sound Level Meter - Either a sound level meter with 'A' weighting network complying with I.E.C. 651 Type 1 (Precision) or an integrating type complying with I.E.C. 804 Type 1 (Precision) should be used.

Sound Source - This shall comprise of an integral loudspeaker, amplifier and random noise generator. The spectrum for the third octave band sound power level of the radiated random noise shall fall within the shaded limits of Figure 1. In order to overcome the problems of background noise on building sites a broadband sound power output of at least 115 dB re 10-12 is required. It is desirable that the sound source is capable of being switched on and off remotely by a transmitter. Other recommended physical dimensions are included in the Annexes.

Absorbent Pack - Earlier requirements on the use of added absorption for the purposes of sound level spectral correction have now been discarded.

TEST ARRANGEMENT
The building shall be accepted as designed with no special preparation permitted except to allow for the curing or drying of specific materials (see Table 1).

A note shall be made of whether or not the rooms are furnished.

TEST PROCEDURE
Sound Source Position - The loudspeaker shall be placed on the floor on the side of the room opposite the party wall, 1.0m from one of the corners. When testing floors the lower room shall be chosen as the source room.

Sound Level Measurement - The sound level in the receiving room shall be adjusted so as to produce an A-weighted sound level in the receiving room of at least 10 dB above background level. Should this not be possible then a correction shall be applied in accordance with Table 2.

The 'A' weighted equivalent sound pressure level $L_{Aeq}$ as defined in I.E.C. 804 shall be measured directly using an integrating sound level meter or estimated using a sound level meter and
procedures as defined in I.S.O. 1960-1. The measurement time interval as defined in I.S.O. 1960-1 shall be approximately 30 s. The sound level meter shall be held out at arms length with the operator standing as near to the centre of the room as possible. The microphone shall be rotated four times through 180° by moving the arm up and down in a gentle movement during the traverse. The four rotations shall be completed in the measurement time interval.

CALCULATION OF RESULTS
Having calculated the sound level difference $D_A$ between the source and receiving rooms the appropriate value for the estimated $D_{nT,W}$ shall be read directly from a Conversion Table.

TEST REPORT
This shall include:
(a) Reference to the Draft International Standard
(b) The result, e.g. $D_{nT,W}$ or $R'_W$
(c) Unambiguous description of location of source and receiving room
(d) Date of Test
(e) Name of Person carrying out test
(f) Type of Sound Level Meter used
(g) Description of room furnishings, e.g. fully furnished, carpet only or unfurnished
(h) Approximate room dimensions
(i) Background noise level in receiving room
(j) The source and receiving room values and $D_A$

FURTHER RESEARCH
The omission of normalisation undeniably affects the accuracy of the estimated result, which is now considered to be in the order of $\pm 3\, \text{dB}$. Evaluation of the new Draft Proposal is currently being undertaken in a programme of field trials due to be completed by December 1989.

REFERENCES
Figure 1
Tolerances for the 1/10th Octave Band Sound Power Levels of the Sound Source

<table>
<thead>
<tr>
<th>Material</th>
<th>Recommended Minimum Aging Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masonry</td>
<td>7 days</td>
</tr>
<tr>
<td>Plaster:</td>
<td></td>
</tr>
<tr>
<td>Thicker than 3 mm</td>
<td>7 days</td>
</tr>
<tr>
<td>Thinner than 3 mm</td>
<td>3 days</td>
</tr>
<tr>
<td>Wall Board Partitions:</td>
<td></td>
</tr>
<tr>
<td>With water-base laminating</td>
<td>7 days</td>
</tr>
<tr>
<td>adhesives</td>
<td></td>
</tr>
<tr>
<td>With non-water-base laminating</td>
<td>3 days</td>
</tr>
<tr>
<td>adhesives</td>
<td></td>
</tr>
<tr>
<td>With typical joint and finishing compounds</td>
<td>12 h</td>
</tr>
<tr>
<td>Other</td>
<td>as appropriate for caulking and adhesive compounds involved</td>
</tr>
</tbody>
</table>

Table 1
Recommended Minimum Aging Period Before Test

Table 2
Difference between sound pressure level measured with sound source operating and background level alone
Correction to be subtracted from sound pressure level measured with sound source operating to obtain sound pressure level due to sound source alone

<table>
<thead>
<tr>
<th>Correction to Sound Pressure Level Readings</th>
<th>dB</th>
<th>dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 - 5</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6 - 9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Measurement invalid and should be repeated
DOIT-ON CHANGER DE COURBE DE REFERENCE POUR EVALUER LA QUALITE DE LA PROTECTION CONTRE LES BRUITS D'IMPACT ?

R. Josse

Centre Scientifique et Technique du Bâtiment
24 rue Joseph Fourier, 38400 Saint-Martin-d'Hères, France

Toutes les enquêtes qui ont porté sur les immeubles collectifs d'habitation ont montré que c'est toujours le voisin du dessus, en moyenne, qui est considéré comme étant la cause du plus de gêne en matière de bruit. Or qu'a donc de particulier ce voisin si on le compare à celui du dessous ? Essentiellement le fait que le bruit des impacts qu'il produit en marchant ou en déplaçant du mobilier est nettement mieux perçu que celui produit par le voisin du dessous. La protection contre les bruits d'impact dans le collectif devrait donc être meilleure qu'elle ne l'est actuellement afin de ne pas sensibiliser les occupants à ce type de perturbation. Cette protection est l'œuvre du constructeur, le bâtiment étant livré avec une certaine insonorisation, et éventuellement celle de l'occupant qui met en place des tapis supplémentaires. Dans la réglementation, seule la part incombant au constructeur peut être prise en compte.

Malgré leur variété apparente, les réglementations ont tous en commun le fait qu'elles jugent de la sonorité d'un plancher en prenant pour source de référence la machine à chocs normalisée telle que définie dans la norme ISO 140. L'idée de contrôler la sonorité d'un plancher en utilisant une machine équipée de marteaux frappant le sol n'est pas récente puisque le premier laboratoire à avoir mis cette idée en pratique est le National Bureau of Standards, Washington, DC en 1928. Toutefois, la machine d'alors n'était pas identique à celle utilisée aujourd'hui et dont la normalisation remonte à 1938 (DIN 4110) pour l'Allemagne et à 1960 pour l'ISO (ISO R140).

De longue date la machine a été critiquée [1, 2] car produisant des impacts bien plus forts que ceux de la vie courante, en particulier que ceux dus à la marche des personnes. Des machines reproduisant plus fidèlement les impacts des talons de chaussures ont été conçues [3, 4] et fonctionnent bien. Malheureusement, elles présentent l'inconvénient majeur de n'engendrer que des bruits de niveau faible donc difficilement mesurables de jour dans des bâtiments usuels. Par contre, elles sont facilement utilisables en laboratoire.

Pour les besoins de test ou de contrôle in situ et corrélativement pour l'expression des règlements, on n'a donc rien trouvé de mieux que la machine normalisée et il est probable que celle-ci sera utilisée encore pendant de nombreuses années.

On sait que, pour la plupart des règlements, la sonorité d'un plancher aux impacts est jugée en comparant le spectre du bruit produit par la machine à un spectre de référence (ISO 717-2) correspondant à un plancher traditionnel réputé donner satisfaction. Si cette manière de faire se justifie par des raisons historiques, elle n'est pas moins peu scientifique et la question se pose de savoir si avec la même machine on peut faire mieux.

Le problème posé est celui de déduire la sonorité d'un plancher aux impacts de la vie courante de la sonorité mesurée à l'aide de la machine normalisée et de la comparer au niveau de bruit limite acceptable. Or on sait qu'en toute rigueur, le problème n'est soluble que si l'ensemble plancher-revêtement de sol se comporte d'une manière linéaire au cours des impacts qu'il subit. Malheureusement, dès que le plancher est couvert d'un revêtement mince, cette condition de linéarité n'est pas respectée. Alors, que faire ?
La seule manière scientifique de résoudre le problème si l'on désire conserver l'usage de la machine normalisée est de limiter cet usage aux planchers non recouverts d'un revêtement mince souple. Cette proposition avait déjà été faite par E. GERRETSEN en 1976 [5].

En complément, les revêtements minces souples, qui sont des produits industriels de caractéristiques bien définies mais de comportement non linéaire, auraient leur efficacité à réduire les bruits d'impact mesurée en laboratoire à l'aide d'une machine simulant les impacts de la marche des personnes. Bien que ces derniers ne soient pas les seuls de la vie courante, il semble qu'il y ait un consensus international pour affirmer qu'ils sont dominants.

Le règlement pourrait donc alors imposer une performance minimale du plancher nu excité à l'aide de la machine normalisée ainsi qu'une performance minimale du revêtement mince associé mesurée en laboratoire sous une excitation par une machine simulant un marcheur.

Les résultats des expériences de FASOLD (1965) [6] et d'autres similaires, permettraient facilement le passage de la connaissance de la sonorité du plancher nu excité par la machine normalisée à la sonorité du même plancher excité par des marcheurs. De là, il serait facile de déduire les caractéristiques minimales d'un revêtement mince mesuré à l'aide d'une machine simulant la marche de personnes pour respecter un objectif de niveau de bruit de marche défini au préalable.

Passer des méthodes actuelles pratiquées sur le plan réglementaire à la méthode plus rigoureuse venant d'être exposée constitue un saut que beaucoup ne voudront pas tenter.

Ceux-ci voudront alors peut-être se laisser convaincre que la courbe de référence actuellement utilisée dans la méthode ISO devrait être remplacée soit par une droite horizontale soit par une droite légèrement ascendante avec la fréquence (3 dB/octave). En effet, toutes les études tant en laboratoire qu'en terrain ont montré que le niveau de bruit des planchers excités par la machine normalisée est plus faible que le niveau de bruit des planchers excités par les marcheurs. Le calcul de l'indice global tenant compte d'une nouvelle courbe de référence, il est possible qu'un moins cela se fera.

* il s'agit des ensembles plancher-revêtement de sol mince ou épais.
LABORATORY SOUND INSULATION INDICES NEEDED TO ACHIEVE THE REQUIREMENTS IN THE FIELD.

P.H. Heringa
Peutz and Associés B.V.
P.O. box 407
6500 AK NIJMEGEN
The Netherlands

As an acoustical consultant we often encounter the problem of the difference between laboratory sound insulation indices and the indices measured with the same partition walls in the field. Under field conditions the sound insulation of a wall is influenced by a large number of differences comparing to the laboratory, such as:
- Flanking transmission
- Edge constraint conditions and resulting loss factor of the wall
- The occurrence of other transmission paths, such as through a suspended ceiling.

A consultant has the task to state the acoustical requirements for the building elements in order to achieve the result required in the field. From a compilation of results of laboratory measurements and "good" measurements of the same separation wall in the field - without sound leaks etc. - we have determined the minimal difference between field requirements and laboratory sound insulation, for several quality ranges and situations.

Table I shows a number of field situations and the sound insulation required.

Table I: Global requirements for the single-number quantities of airborne sound insulation in the field \( R'_w \) in different situations (2)

<table>
<thead>
<tr>
<th>Normal offices</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offices with higher requirements for speech privacy</td>
<td>43</td>
</tr>
<tr>
<td>Offices with high requirements for speech privacy</td>
<td>48</td>
</tr>
<tr>
<td>Between classrooms</td>
<td>45</td>
</tr>
<tr>
<td>Between dwellings</td>
<td>52</td>
</tr>
<tr>
<td>Between music classrooms</td>
<td>54</td>
</tr>
<tr>
<td>Between rooms in dwellings</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospitals: From/to</th>
<th>Bedroom</th>
<th>Consultation</th>
<th>Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedroom</td>
<td>43</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Consultation</td>
<td>48</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>Instruction</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>Sanitary</td>
<td>43</td>
<td>43-48</td>
<td>43</td>
</tr>
<tr>
<td>Corridor</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
</tbody>
</table>

* In the Dutch situation expressed after the Dutch code in \( I_{16} = R'_w - 52 \)
Offices

In order to achieve maximum flexibility in offices it has become common practice to stop partitions under the suspended ceilings. Experience has shown that the sound transmission through a suspended ceiling is more difficult to reduce than through a partition wall. With suspended ceilings a sound insulation index above $R'_w = 42 \text{ dB}$ is difficult to achieve. To achieve such a value the sound transmission through the partition must be negligible to the transmission through the ceiling. A sound insulation for the partition wall of at least 5 dB higher than the ceiling is recommended.

Light weight partition walls

In most situations other than offices, partitions reach from floor to floor. Table II shows the differences between laboratory measurements and field measurements for light weight walls reaching from floor to floor in a concrete building.

Table II: Average sound insulation index in the field ($R'_w$) and in the laboratory ($R_w$) of gypsum board walls on metal studs. (1)

<table>
<thead>
<tr>
<th>$R'_w$ lab (dB)</th>
<th>$R'_w$ field (dB)</th>
<th>field standard deviation (dB)</th>
<th>number of field situations</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>36</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>45</td>
<td>40</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>51</td>
<td>46</td>
<td>2.3</td>
<td>39</td>
</tr>
<tr>
<td>55</td>
<td>47</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>59</td>
<td>56</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>
Based on the results shown in table II we recommend a difference of 3-5 dB between the laboratory sound insulation and the field requirement.

**Brickwork walls**

A similar study has been performed for brickwork partition walls. The results are shown in table III.

**Table III: Average sound insulation index in the field (R'\(_w\)) and in the laboratory (R\(_w\)) of sand-lime walls (plastered).**

<table>
<thead>
<tr>
<th>wall thickness</th>
<th>(R'_w)PRESS (dB)</th>
<th>(R'_w) FIELD (dB)</th>
<th>field standard deviation (dB)</th>
<th>number of field situations</th>
</tr>
</thead>
<tbody>
<tr>
<td>214 mm</td>
<td>55</td>
<td>52</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>230 mm</td>
<td>57</td>
<td>53</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>265 mm</td>
<td>58</td>
<td>54</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>cavity wall*</td>
<td>100-60-100 mm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ground floor</td>
<td>65</td>
<td>57</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>first floor</td>
<td>65</td>
<td>61</td>
<td>3</td>
<td>21</td>
</tr>
</tbody>
</table>

* cavity wall with no ties but combined foundation for both leaves.

The results of table III agree with the conclusion based on table II. Only for the cavity wall without ties the laboratory sound insulation is much higher than in the field (ground floor).
Conclusion

The laboratory sound insulation indices needed in order to achieve the requirements in the field differ for different situations. In general three different categories can be distinguished. These three categories are based on the situation whether or not there is a suspended ceiling, and on the required sound insulation index. The difference between laboratory and field increases with increasing sound insulation index (table IV).

Table IV: Difference between "laboratory" and "field" sound insulation indices in concrete buildings.

Situations with suspended ceilings 5 dB

Situations with floor to floor partition walls:
- Requirements $R'_{ww} \leq 40$ dB 3 dB
- Requirements $40 < R'_{ww} < 55$ 5 dB

Even when applying this surplus in sound insulation a sufficient reduction of flanking transmission has to be considered depending on the situation and the requirements.

Reference

(1) R. M. Markemeyer, E. Ph. J. de Ruiter
    Sound insulation of gypsum board walls (in Dutch) to be published by the Dutch Building Research Foundation (Stichting Bouwresearch), 1989 Rotterdam.

THE FUTURE OF SINGLE NUMBER INDICES FOR ASSESSING
SOUND INSULATION AS SEEN BY THE CONSULTANT

James F. Yerges

Yerges Acoustics, 5209 Lee Avenue, Downers Grove, Illinois 60515 USA

For the past 18 years, this author has been a professional acoustical consultant practising almost exclusively in the United States. More of this time has been spent dealing with design, specification, and measurement in real buildings than in laboratories. From this perspective, does this consultant see a future for single number airborne and impact sound insulation ratings in the United States? Absolutely. The nature of the design, specification, and regulation process virtually dictates the use of such standards.

The programming documents which become part of the contract between an owner and an architect often stipulate minimum acceptable STC and IIC values for constructions separating various adjacencies. The architect must then select and specify constructions capable of providing the required degree of sound insulation, and specify minimum acceptable performance. This performance requirement, which may be verified by a field test, insures that the installing contractor complies with the design drawings and maintains an adequate standard of workmanship. In the event of subsequent litigation - an increasingly common occurrence in the U.S. - the single number criteria become central to legal arguments. Were the chosen values appropriate for the application? Were the specified constructions capable of achieving the specified performance? Were the materials and workmanship adequate to achieve the potential performance? The acoustical consultant may be pressed into service as an expert witness, to conduct field measurements and to explain the rating systems and measurements to the court.

Sound insulation rating schemes have become a part of the legal fabric in the U.S., and will be with us into the foreseeable future. With this in mind, what are we measuring, and what should we be measuring? In the U.S., laboratory airborne sound insulation measurements are conducted in accordance with ASTM E 90-87 and expressed as STC values, in accordance with ASTM E 413-87. The STC, or Sound Transmission Class, represents a "recipe" in which the sound transmission loss measured in 1/3 third-octave frequency bands is fitted to a contour which crudely approximates the A-scale response of a sound level meter. The A-scale, in turn, crudely approximates the venerable equal loudness contours at moderate sound pressure levels. Significantly, the lowest third-octave frequency band used in the computation is centered at 125 Hz.

"Real-world" airborne sound insulation measurements are conducted in accordance with a variety of methods which derive from the STC. The FSTC (Field Sound Transmission Class), NIC (Noise Isolation Class), and NNIC (Normalized Noise Isolation Class) are all methodologies based on third-octave sound transmission loss or noise reduction data fitted to the STC contour, per ASTM E 336-84. Again, the lowest third-octave frequency band used in these computations is centered at 125 Hz.

Until very recent improvements in instrumentation, gathering the data necessary for computation of these descriptors was so tedious and time
*13th INTERNATIONAL CONGRESS ON ACOUSTICS*
*YUGOSLAVIA* *1989*

-consuming that it was simply not feasible for large-scale proof-of-
performance testing. As a result, a method based on the A-weighted
difference of band-limited pink noise was developed and promulgated as
ASTM E 597-81. Once again, the source spectrum used in this procedure is
severely attenuated below the 125 Hz octave frequency band center.

It will probably come as a surprise to no one that the SLC and its
derivative ratings have not proven to be universally successful as
predictors of subjective judgements of the adequacy of sound insulation.
Philosophically, the STC contour anticipates a speech spectrum with
limited low-frequency energy. As a result, it is a fairly good descriptor
of the sound insulation required to separate adjacent offices. However,
it woefully underestimates the subjective response to low-frequency sound
energy from sources such as mechanical equipment. More importantly, the
STC ratio utterly fails to predict the strong subjective aversion
to the rhythmic low-frequency component of reinforced music - an
inescapable element in multifamily dwellings in the United States.

When H. Stanley Roller was affiliated with the United States Gypsum
Company, he proposed an alternative method for curve fitting available
third-octave transmission loss data. The STC formula permits a total of
32 deficiencies, with a maximum of 8 deficiencies at any one frequency.
Roller's MTC does not permit any deficiencies at 125 or 160 Hz. This
author is not aware of any large-scale controlled study to validate the
MTC as a sound insulation criterion, but has found it to be very useful
both as a predictor of subjective response in multifamily dwellings, and
as a vehicle for explaining the frequency dependence of sound transmission
loss to owners and architects.

Whether the STC, the MTC, or some other curve-fitting paradigm is
deemed most appropriate for a given situation, a source-independent
measurement methodology should continue to be used in the laboratory, and
the exclusive use of source-independent measurements should probably
become the norm even in the field. This probably means noise reduction
measurements in third-octaves with corrections for reverberation when
appropriate, rather than A-weighted or other weighted measurements using a
source with a shaped spectrum. It would also be appropriate to measure
and report TL at lower frequencies, even if these data were not ultimately
used in computing the single number criterion. These data collection
practices might ultimately lead to a body of field data large enough that
real-world performance predictions could be made confidently, and the
tenuous correlation between laboratory performance and field performance
could be better understood.

Source-dependent measurements such as the A-weighted difference have
served us well for many years. However, they do not yield diagnostic
information. More importantly, they were really only stop-gap measures
while we waited for computer controlled real time analysis equipment to
become portable and affordable. That time has very nearly arrived.

If a practising acoustical consultant composed a "wish list" of
features to be incorporated into a test kit, the first criterion would be
portability. The analyzer should be no larger than a standard brief case
(approximately 45 cm x 35 cm x 15 cm). It is feasible to arrange to have
massive loudspeakers and an amplifier delivered to a job site in a distant
city by a local music rental supplier, but the analyzer must be placed
under the seat of an airplane - even the window seat on a DC9. The
analyzer should weigh no more than 5 kg, because it would be carried
through airports and all over job sites all day long, by engineers who are no longer necessarily young and vigorous. The unit should be completely self-contained, with adequate battery life to provide true freedom from an AC power source. The unit should have adequate memory, in a combination of non-volatile RAM and tape or disk, to provide essentially unlimited data storage capability. The unit should provide hard copy of critically important data in the field.

An analyzer used for TL measurements should be capable of recording a large number of source and receive room spectra, and reporting the average values of these spectra. This would permit both temporal and spacial averaging. Note that a typical pink noise source produces a third-octave spectrum which is flat only in the statistical sense. The portable analyzer presently used by this author reports spectra to a precision of 0.1 dB. It is necessary to average 128 spectra over a time span of approximately 70 seconds to achieve this degree of repeatability from the companion pink noise generator.

Spacial averaging is the only practical way to consistently obtain a sound pressure level spectrum which represents the mean energy density in a real room in a field test. As a practical matter, a continuous or moving average can be obtained much more readily than the average of spectra taken at discrete locations. This opinion is not offered as a refutation of Waterhouse and Lubman (JASA, Vol. 48, No. 1, Part 1, 1970), which has become regarded as something of a sacred text in the U.S. Rather, it is intended to point out that an estimate of the mean can be obtained to any required degree of accuracy, despite the presence of correlation between samples, by enlarging the ensemble adequately. Modern instrumentation permits ensemble sizes scarcely imagined in 1970. Like weighted measurements and shaped spectra, it is time to leave discrete space averaging behind.

The acoustical consultant's ideal analyzer would permit a large number of reverberant decays to be recorded in the receive room and displayed as an average decay for each test frequency band. An ensemble of 16 decays might be routinely gathered, with difficult rooms requiring as many as 32 decays. During analysis, this data could be displayed graphically for inspection, and an interactive program would permit the engineer to select the portion of each average decay curve which best represents the presumed model exponential decay. And this leads the author to another digression.

Except in large, diffuse, reverberant rooms, the notion of a statistically uniform exponential decay of sound energy has proven to be a very crude model. What does the decay of sound energy at a point in space really "look" like? Like the ETC (Energy Time Curve). As its name suggests, the ETC depicts the total energy (potential and kinetic, or real and quadrature) at a point in space as a function of time over the frequency range of interest. It is not simply the rms sound pressure, which represents only potential energy, and it is not "smeared" by the finite slewling rate of a mechanical recorder. The proprietary test apparatus used to measure and display the ETC provides an excellent signal-to-noise ratio without balloons, firearms, or other nonsense. While the existing so-called TEF analyzers have not yet been adapted to meet the needs of engineers conducting sound transmission loss measurements, they employ powerful techniques for analyzing sound fields in rooms. Acousticians should become familiar with these techniques and monitor the development of these devices.
Sound transmission measurements utilizing dual-channel analyzers and sound intensity methods have been much discussed in recent years. These methods might render the whole matter of reverberation corrections moot. This author eagerly awaits the promulgation of a standard for measuring sound transmission loss using intensity methods.

Returning now to the discussion of the acoustical consultant's ideal portable analyzer, it would be helpful if the device could compute the FSTC from test results in the field. This would be useful to placate a nervous owner or contractor. However, it is more important to be able to upload the data into a desktop computer to manipulate and graph it.

The analyzer hardware described herein already exists, from several different manufacturers. However, this author has not yet seen a hardware-plus-software package worthy of being deemed "complete." The only two fully-implemented systems this author has seen utilize hardware which is so bulky that it can scarcely be called portable.

What about impact noise insulation? The IIC (Impact Insulation Class, per ASTM E 989-84 and E 1007-84) is the only standard widely used in the United States. Impact noise generated by a standard tapping machine is measured in 16 third-octaves and plotted against a contour which is essentially the obverse of the STC contour. The lowest frequency band used in the computation is centered at 100 Hz. It will come as a surprise to no one that the little tapping machine does not excite gross, low-frequency deflection of the timber frame construction commonly employed in the U.S. to the same degree that footfalls do. This matter is presently under study by an ASTM subcommittee.

And now a cautionary word is in order about the correlation between published laboratory data and the actual performance which may be anticipated in the field. For decades, there were only a few accredited acoustical testing laboratories in the United States. These labs tended to operate with autonomy and independence. Published laboratory tests tended to correlate reasonably well with field test data. A partition capable of an STC 50 in the lab was thought to be consistently capable of FSTC 45 in a good field installation.

In recent years, an egalitarian accreditation procedure has made it possible for individual manufacturing corporations and entrepreneurs to own and operate test labs. This seems to have spurred a "ratings inflation." In certain cases, published test results may represent such low-probability events that they may be thought of as statistical accidents. This author has a recent test result from the accredited laboratory of a major U.S. building materials supplier depicting a time-honored drywall-on-wood-stud construction. For decades this construction has been producing approximately FSTC 46 in the field. The new laboratory test claims STC 58. Such a test would be potentially very dangerous in the hands of an unsophisticated architect or developer.

In summary, single-number sound insulation ratings based on third-octave measurement data will continue to be used in the design and specification process in the United States for the foreseeable future. It is entirely possible that multiple ratings will come into use, reflecting differing sound insulation requirements in different applications. The difference between laboratory measures and performance and field test measures and performance should be reflected in more sophisticated specifications and regulatory documents. New measurement techniques can be expected to impact these ratings.
COMPARISON BETWEEN EUROPEAN SOUND INSULATION RATING METHODS FOR DWELLINGS

J-P. Vian

CSTB, 24, rue Joseph Fourier, 38400 SAINT MARTIN D'HERES/FRANCE

1. INTRODUCTION

The sound insulation between dwellings varies with frequency. To cover the range thought to be of most subjective importance for separating partitions measurements are generally made in between 5 and 16 contiguous frequency bands within the range 100 Hz to 4000 Hz. The measurement therefore produces up to 16 values to describe the insulation between each pair of rooms. This is too complicated for most purposes, such as regulations, and a rating method which provides a single index of quality based on these measurements is required. Most developed countries have such an index, but these indices differ and it has seemed very useful to compare those used by member states of the European Community.

The main criterion for a good index is how well it relates physical measurements to subjective responses. Therefore to compare indices subjective assessments of the sound insulation between a large number of dwellings where the insulation has been measured are required. BRE in the UK and CSTB in France have conducted the necessary physical and social surveys and used their data to assess and compare various rating systems.

2. SOUND INSULATION RATING METHODS

2.1. Origins of sound insulation requirements

The recognition of the problem of sound insulation between dwellings began after the work of Sabine at the turn of the current century. It was not until 1930 that European construction regulations began to include sound insulation requirements. The early regulations generally specified minimum acoustical performance for walls and floors relied on the reputation of sufficient quality. In order to obtain a single value for rating the insulation of partitions, the regulatory requirements generally employed arithmetic means of the third octave band transmission loss.

Since this method proved to be less than satisfactory another approach using standard reference curves was introduced in the 1950s.

The first such method appeared in Germany in 1953. The reference curve used in the German standard DIN 4109 was derived from the transmission loss curve of a 250 mm solid brick wall, plastered on each side, which was considered subjectively satisfactory.

Various studies led to the adoption of the DIN curve by ISO (ISO R717-1968).

Most of the European countries adopted methods very similar to the ISO method, with the exception of France.
The French regulation of 1969 adopted a quite different approach, which states that the A weighted sound level in the receiving room must not exceed a given value when the source room is excited by a band limited pink noise.

2.2. Physical rating methods

Rating procedures have three elements:

a) The form of correction applied to measurement in the receiving room.
   These are two basic types of correction. Either a fixed reverberation time of 0.5 s, or a standard area of absorption - usually taken as 10 m².

b) Reference curve of insulation against frequency
   Measured results are compared with the reference curve, the shape of which reflects the relative subjective importance of different frequencies.

c) The procedure to reduce the measured data to a single index
   There are several procedural differences between rating methods. Some give the index only as a broad class. Some show the difference between the measured result and the reference curve.

2.3. European rating methods

Several rating methods are used by member states and they may be broadly classified as follows:

- Denmark, Germany (FR), Greece, Italy: ISO 717 based methods
- Belgium: National method NBN SO1-400 1977
- Republic of Ireland, United Kingdom: AAD method dB(A) method
- France: Draft national method ELOT 493 and ISO 717 method
- Greece: No national method
- Luxembourg: National method NEN 1070 1976

3. DESCRIPTION OF SURVEYS AND ANALYSES

3.1. French Survey

The French survey was undertaken among the occupants of multiple storey buildings to compare the subjective assessment of their party wall insulation with various rating systems.

To test a reasonably wide set of insulation types, we chose a sample of buildings of low, medium and high acoustical quality, including different types of construction in order to take into account varied slopes of insulation vs. frequency. The range of average insulation values was chosen between 44 and 57 dB(A) while the nominal slopes ranged between 4 and 12 dB octave.

A questionnaire was developed to be fully compatible with that used in the UK survey. More details about the survey can be found in the report [1].
3.2. UK Survey

The UK Survey was undertaken among the occupants of new homes to establish their opinions on the adequacy of sound insulation and other environmental factors.

Measurements were made on over 1200 party walls in houses or bungalows, generally in groups of four per site.

The subjective data was obtained from a social survey which has been described by Langdon, Buller and Scholes [5].

3.3. Common methodology

The product moment regression analysis in both studies is based on answers to the question:
"How would you rate the sound insulation of your house from the one(s) next door?" The possible answers were: 1 very good, 2 good, 3 fair, 4 poor or 5 very poor. The results take the form of regression coefficients between the subjective ratings and the corresponding physical ratings obtained by each rating method.

3.4. French analysis and results

The results obtained for all the buildings without the biased sites are shown in Table 1.

The correlation coefficients range from 0.7 to 0.85 and are all significant, but not statistically different from each other. Nevertheless, the French dB(A) seems to give the best result and the Belgian method the worst, because of the large steps between grading curves.

Various modifications have been applied to the ISO, UK and French methods, combining them with different curves. The ISO method was used with the Party Wall Grade (PWG - UK), with the A weighting curve considered as a reference curve, and with the Dutch curve. Because of the need of a weighting curve for the French method, it was only associated with the C and D curves.

Very few changes appear in the correlation coefficients for these combined methods (See report [6] for more details). All these results are not statistically different from each other.

3.5. UK analysis and results

The results are shown in Table 2.

The correlation coefficients range from 0.66 to 0.71 but are not statistically significantly the best result came from the dB(A) method as used in France. The Belgian method using only broad categories of physical insulation seems to under-estimate subjective discrimination. There is little to choose between the results given by the other methods.

Various modifications have been applied to the Belgian ISO, AAD and French methods and none of them gave better results. See report [6] for more details.
CONCLUSION
The main conclusions of this study are:
(1) Of the rating methods currently used through European countries no method was found to be better than the others in a statistically significant way;
(2) no variations on the basic methods tested gave significantly better results than the basic methods;
(3) subjective satisfaction with sound insulation cannot be predicted completely by physical measures of sound insulation alone;
(4) the slope of the insulation vs. frequency curve was found in this work and a laboratory experiment to affect clarity of speech and music more than it affected overall satisfaction with sound insulation.

REFERENCES
[1] AUDREE (A.), VIAN (J-P.)
"Le choix d'un indice d'évaluation de la qualité des isolalements acoustiques des logements entre eux", Rapport CSTB 2.83.006 (Novembre 1983)
[2] VIAN (J-P.), DANNER (W.)
"Assessment of significant acoustical parameters for rating sound insulation of party walls". JASA 73 (4). (April 1983)
[3] VIAN (J-P.)
"Isolement de reconnaissance des sons" Rapport CSTB 2.81.001 (Janvier 1981)
"Sound insulation performance between dwellings built in the early 1970s", BRE Current Paper CP 20/78
[6] POTHERGILL (L.C.), VIAN (J-P.)
"Validation of an Index for the quality of airborne sound insulation between dwellings", BRE CSTB Report (May 1984)

<table>
<thead>
<tr>
<th>METHOD</th>
<th>REGRESSION EQUATION</th>
<th>CORRELATION COEFFICIENT</th>
<th>SIGNIFICANCE LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO</td>
<td>-0.25</td>
<td>0.84</td>
<td>0.001</td>
</tr>
<tr>
<td>BELGIAN</td>
<td>0.73</td>
<td>0.71</td>
<td>0.1</td>
</tr>
<tr>
<td>UK AAD</td>
<td>0.001</td>
<td>0.57</td>
<td>0.001</td>
</tr>
<tr>
<td>FRANCE (dB(A))</td>
<td>0.73</td>
<td>0.84</td>
<td>0.001</td>
</tr>
<tr>
<td>NETHERLANDS</td>
<td>-0.19</td>
<td>2.82</td>
<td>0.001</td>
</tr>
</tbody>
</table>

TABLE 1. RESULTS OF REGRESSION ANALYSIS WITHOUT THE BIASED SITES FOR THE FRENCH SURVEY

<table>
<thead>
<tr>
<th>REF</th>
<th>CURVE</th>
<th>RATING METHOD</th>
<th>REGRESSION EQUATION</th>
<th>CORRELATION COEFFICIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ISO</td>
<td>ISO</td>
<td>-0.10</td>
<td>8.54</td>
</tr>
<tr>
<td>2</td>
<td>BELGIAN</td>
<td>BELGIAN</td>
<td>0.56</td>
<td>1.91</td>
</tr>
<tr>
<td>3</td>
<td>AAD</td>
<td>AAD</td>
<td>0.02</td>
<td>2.72</td>
</tr>
<tr>
<td>4</td>
<td>dB(A)</td>
<td>dB(A)</td>
<td>-0.11</td>
<td>8.54</td>
</tr>
<tr>
<td>5</td>
<td>NETHERLANDS</td>
<td>NETHERLANDS</td>
<td>-0.10</td>
<td>3.58</td>
</tr>
</tbody>
</table>

TABLE 2. RESULTS OF REGRESSION ANALYSIS FOR THE UK SURVEY.
FIELD EXPERIENCES AND RELIABILITY OF SHORT TEST PROPOSALS FOR AIRBORNE SOUND INSULATION MEASUREMENTS

Lucía N. Taibo

Instituto Nacional de Tecnología Industrial, INTI, C. C. No. 157, 1850 Prov. de Buenos Aires, Argentina

Introduction

Laboratory results of airborne sound insulation of building elements by standardized procedures serve as a data base for the architectural design. However, experience shows that the acoustical behavior of final constructions could greatly differ from predictions based on laboratory tests. Even minor mistakes during construction can spoil the best acoustical design, making isolation tests of newly built or in use constructions an essential means to check the observance of building regulations in the field. There is a general consensus that ISO 140 procedure is not adequate for field measurements on a broad scale, and simple tests that can be used by even non-specialized persons are essential as a tool for an effective legal control with the resulting improvement on the quality of buildings.

Short test procedures

There has been a growing activity in the study of simplified proposals for sound insulation measurements during last years not only as an interesting research topic but as a background for standardized proposals. In a broad sense short methods are focussed to two different approaches: 1) those that having at least the same precision as ISO procedure make use of sophisticated electronic equipment in order to simplify the measuring task, thus reducing the involved time and the probability of human mistakes, and 2) those not conceived as a thorough alternative to ISO method, but that at the expense of a restricted precision actually simplify the measurement procedure itself. Depending on the attempted goal, simplified ratings were correlated to subjective reactions [1], to standard ratings [2], [3] intended for legal control, or screening for quality control indicating when fuller measurements are needed, or simply to characterize an actual situation without any further attempt for standardization [4].

Among methods included in the second category those based on global "A" level differences, DA, between rooms when emitter is excited by a pink noise spectrum are most popular, due partly to their simplicity. It has been proved that they are sensible to subjective judgments of acoustical comfort and some building codes establish
requirements on RA indexes.

When a correspondence between DA (or RA) and standard ratings, DnT,w (Rw or STC) is pursued, three main factors have a direct influence on the attainable precision in real situations: 1) Incident noise spectrum shape: no more ideally pink, 2) Receiver room absorption: different from standard T60 = 0.5 s defined for DnT, and 3) Sound insulation curves shape: are encountered in a wide variety some of them greatly differing from "A" weighting or Rw (IA or STC contours). In practice, the extent to which the first two factors are able to be controlled, altogether with the followed experimental design (microphone and loudspeaker positions, type of sampling, etc.), will determine the overall precision of the method.

Field survey of typical rooms

As a first attempt to establish the validity of normalizing the absorptive characteristics of typical rooms, a reverberation time survey has been conducted [5]. For this purpose rooms were classified into two main groups: 1) In use buildings with normal furnishing and 2) Newly built ones. In the present work we shall refer to the first group.

In use buildings: Reverberation times of 180 rooms with normal state of furnishing mainly dwelling-units (except kitchens and baths), and small offices, with volumes ranging from 25 m³ to 100 m³, a minimum floor area of 8 m² and a maximum of 35 m² were determined. Measurements were conducted in 1/3 octaves with a fully automatic processor. Results are shown in Figure 1.

![Figure 1: Average reverberation times of 180 rooms furnished rooms.](image)

No additional absorption was planned for the measurements. As can be observed RTs are almost independant of frequency within 100 Hz to 3150 Hz, with mean values around 0.6 s and s.d. ranging from 0.15 s to 0.1 s above 800 Hz.

Emitter spectrum shapes

Two different factors that may influence field spectra were considered. First, in order to determine the extent of
deviations from a pink noise source with a flat acoustical output (previously equalized in a typically furnished room) due to the unequal absorptive properties of rooms. 1/3 octave average SPL were measured in 70 furnished rooms. The sample had similar characteristics as those described in the previous section, but was taken in addition to get an independent set of data. Besides, the influence of the acoustical response of different electroacoustical systems (measured in the previous typical room without equalization) was observed. Six examples are shown in Figure 2, where each spectrum shape given in 1/1 octaves corresponds to different emitting systems and the shaded area to the range of absorption.

![Figure 2](image)

**Computer simulation**

Collected results of 70 incident spectra, 180 RT's and 100 airborne insulation curves of varied shapes formed the database for a computer program that simulated DAs for all possible combinations of data. DA were subsequently correlated to Dnt,w. Prediction intervals at the 95% confidence level are summarized on Table 1.

**Table 1**

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>( \Delta_{DNT,W} )</th>
<th>( \Delta_{A, 43} )</th>
<th>( \Delta_{DNT,W} )</th>
<th>( \Delta_{A, 43} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.8</td>
<td>1.9</td>
<td>2.3</td>
<td>2.5</td>
</tr>
<tr>
<td>B</td>
<td>1.3</td>
<td>1.7</td>
<td>2.3</td>
<td>1.8</td>
</tr>
<tr>
<td>C</td>
<td>1.9</td>
<td>1.6</td>
<td>2.1</td>
<td>1.6</td>
</tr>
<tr>
<td>D</td>
<td>1.8</td>
<td>1.5</td>
<td>2.1</td>
<td>1.7</td>
</tr>
<tr>
<td>E</td>
<td>1.8</td>
<td>1.5</td>
<td>2.2</td>
<td>1.7</td>
</tr>
<tr>
<td>F</td>
<td>1.9</td>
<td>1.7</td>
<td>2.2</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Experimental results

ISO 140 technique and "A" level differences method have been applied to measure the airborne isolation between 62 pairs of furnished rooms in dwellings and small offices.

For the simplified test, LA in both source and receiving rooms respectively were averaged from 5 random positions, at least 1 m away from the walls and the loudspeaker, with a Type 1 SELM. A baffle emitting pink noise was placed in a corner opposite the dividing wall facing an horizontal diagonal into the room, with the loudspeaker cone about 1 m away from the adjacent wall. Repeatability was determined within ± 2.5 dB.

Final remarks

As might be expected DA ratings can be related to ISO ones only if experimental conditions are carefully controlled, i.e. the emitter spectrum shape and the absorptive characteristics of the receiver must be restricted within prefixed boundaries. In the present work the measured sample comprised typical furnished rooms of local dwellings and small offices without any added absorption. Not intending to get a closely relationship with ISO ratings, the short test may work as a reliable screening procedure, since DnT,w values are systematically higher than DA. Current research is pursued on the normalization of bare rooms by means of additional absorption. This looks as a promising alternative in order to extend the validity of the short method.

References

ANALYTICAL CONNECTIONS BETWEEN THE REFERENCE CURVE 
AND THE WEIGHTING METHODS

Juhani Parmasen

Technical Research Centre of Finland
Kemistintie 3 /LV1/TH1, SF-02130 Espoo, Finland

Introduction

The indices of airborne and impact sound insulation are used for expressing the sound insulation with a single number. They are commonly obtained with the so-called reference curve methods such as that in the well-known ISO 717-1...3 standards. The advantages of the reference curve methods are that a) the index is obtained without complicated calculations, and b) the graphical presentation reveals directly the frequencies to which the possible improvements should be concentrated, in order to get the wanted insulation (index).

A disadvantage of the reference curve methods is that their acoustical content, and therefore especially also the final result (index), are undefined. They are not connected in international standards to any acoustical quantity, such as sound level difference or sound pressure level. On the other hand, it has been shown in many studies, like [1], that there is a fairly good correlation between the indices and the subjective impression of sound insulation. Obviously, this fact is the main reason why the index $R_\text{w}$ has been taken to be a universal and objective sound insulation measure, suitable for all purposes like expressing the sound insulation of outer walls. So, in some studies $R_\text{w}$ has been given with an accuracy of 0.1 dB instead of the usual whole number presentation [2]. In short test methods of sound insulation $R_\text{w}$ has been, of course, the quantity aimed at, and it has been assumed that it will be obtained by a set of sufficiently well-made acoustical measurements.

What is a reference curve procedure?

It has been shown [3] that in the case of airborne sound insulation, one can write the result of the reference curve procedure, for instance for $R_\text{w}$, as

$$R_\text{w} = R_B + D + d$$

where $D$ is a constant difference between $R_\text{w}$ and $R_B$. The values of the term $d$ vary from case to case, but they have a theoretically limited distribution range, in practice a few dB around 0. $R_B$ is a weighted sound reduction, which is written as

$$R_B = 10 \lg \left( \sum_{i=1}^{16} 10^{B_i/10} \right) / \left( \sum_{i=1}^{16} 10^{(B_i - R_i)/10} \right)$$

where $\{B_i\}$ is a weighting consistent with the reference curve (ISO 717-1 or -3) and $\{R_i\}$ is the sound reduction index (ISO 140). $\{B_i\}$ is determined, according to [3], by the incident spectrum, the absorption in the source and receiving rooms, and the annoyance filter in use, in other words,
those factors will determine the shape of the reference curve. According to Eq. (1), the reference curve procedure is to be understood as an algorithm, yielding an approximation of $R_B$. The content of $R_B$ is defined in Eq. (2). As to the parameters in Eq. (1), the value of $D$ is chosen according to centering of the distribution term $d$. As such, $d$ represents the inaccuracy of the algorithm and can be minimized with a sufficiently good reference curve procedure. Otherwise, it does not have any subjective or objective basis on sound insulation facts. Neither it can be predetermined or taken into account, for example in short test methods.

Moreno has deduced by statistical methods [4] that

$$R_w = R_A + D + d$$ (3)

where $R_A$ is like $R_B$ before, but the weighting $[R_B]$ is replaced by the frequency weighting $A$. The basic statistics of Moreno is shown in Fig. 1 [5]. The validity of Eq. (3) has been shown mathematically [6]. Further, by making suitable assumptions on the shape of most generally existing sound reduction curves, $D$ will get a value of 1.1 dB when within a large material $d = 0$ dB on an average.

For the weighted normalized impact sound pressure level holds [7], similarly to (1)

$$L_{n,w} = L_{n,g} - d$$ (4)

where $L_{n,g}$ is a weighted sound pressure level and $d$ is a distribution term as above. The weighting to be used to get $L_{n,g}$ is given in Table 1. The weighting curve shape is obtained by the negatives of the reference curve in ISO 717-2. Thus also in this case, the reference curve method is an algorithm which gives an approximation for the normalized impact sound pressure level, weighted with the frequency weighting.

Figure 1. A correlation statistics between $R_A$ and $R_w$ according to [5].
Table 1. The weighting terms in $L_{ng}$ (dB) at the frequencies 100-3150 Hz.

<table>
<thead>
<tr>
<th>Frequency, Hz</th>
<th>100</th>
<th>125</th>
<th>160</th>
<th>200</th>
<th>250</th>
<th>315</th>
<th>400</th>
<th>500</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Frequency, Hz</th>
<th>630</th>
<th>800</th>
<th>1000</th>
<th>1250</th>
<th>1600</th>
<th>2000</th>
<th>2500</th>
<th>3150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighting, dB</td>
<td>-12.7</td>
<td>-11.7</td>
<td>-10.7</td>
<td>-7.7</td>
<td>-4.7</td>
<td>-1.7</td>
<td>+1.3</td>
<td>+4.3</td>
</tr>
</tbody>
</table>

**Short test methods**

So far, all the short test methods are generally based on overall level measurements, using suitable frequency filtering for the whole frequency range (e.g. 100-3150 Hz). In particular, the short test methods of airborne sound insulation are based on the common finding that $R_n$ correlates fairly well with the difference of the A-weighted sound levels between adjacent rooms. This correlation is explained qualitatively by Eq. (3). However, $R_n$ cannot be measured directly and one must be content with measuring $R_A$, which actually is the main problem in short test methods. Yet, in many short test studies the main question seems to have been to evaluate the correlation between $R_n$ and $R_A$ [8] and so, in fact, only finding an explanation to the "error" inherent in the reference curve algorithm.

The basic equation of short tests for the airborne sound insulation can be taken from [3] to give the difference of the A-weighted sound levels of adjacent rooms as

$$L_{A1} - L_{A2} = 10 \log \left( \frac{10^{(L_{1}+A_1)/10}}{10^{(L_{1}+A_2-A_1)/10}} \right) + K$$

(5)

where $\{L_1\}$ is the sound level spectrum in the source room, added with a term depending on the receiving room absorption. $K$ is a term which depends only on the absorption in the source and receiving rooms, and on the area of the partition wall. In the case that $\{L_1\}$ happens to be pink noise, the first term on the right side of Eq. (5) will be reduced to the weighted sound reduction $R_A$. If one evaluates the term $K$, for instance by using a reference source, $R_A$ can be solved from Eq. (5) on the basis of the measured sound level difference. The accuracy of short test results is affected most seriously by the quality of the spectrum $\{L_1\}$, which depends also on the sound source characteristics in addition to the factors mentioned before. If $\{L_1\}$ differs from the pink spectrum, the test result generally only approximates $R_A$. The sensitivity of short test methods to the spectrum $\{L_1\}$ can be reduced by using the C weighting instead of A in the source room [9].

For impact sound short tests, a basic equation like (5), can also be written, but using the weighting given in Table 1. Then $L_{ng}$ can be measured with an accuracy which depends on how well one can evaluate the receiving room absorption. The impact sound insulation index $L_{ng}$ cannot be measured directly, either, and the problematics concerning it is identical with the discussion on the algorithm questions. However, the impact sound measurements will give, in principle, $L_{ng}$ more accurately than $R_A$ is given by the short tests for airborne sound insulation. This is due to the fact that in the case of $L_{ng}$ the influence of the reference source characteristics and source room absorption are not involved in the procedure.
REFERENCES


CORRELATION BETWEEN OBJECTIVE RATINGS AND HUMAN REACTION TO IMPACT NOISE

B. McKell

Department of Building, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, Scotland, U.K.

Following on from the development of an ISO short test method for measuring the airborne sound insulation of party walls and floors, work has begun on the development of a simple test for measuring the impact sound insulation of party floors.

The criterion of validity for the airborne short test was based on achieving a good correlation with the results obtained using the full ISO 140 procedure. Such an approach for the impact test is less satisfactory for two reasons. The first concerns the underlying doubts in the correlation of measured results with subjective assessments of impact sound. While the second is due to the spectral variations that are to be found with different levels of insulation [1]. This paper deals with the problems in achieving good correlation and looks at factors which may affect the correlation.

Doubts in correlation

In 1959 Gusele [2] commented that "the effectiveness of a floor against walking noise is not always correctly rated by the impact rating". One consequence of this statement was that several research workers decided to investigate the problem. A recent report by BRE/CSTB [3] reviewed literature on the subject and came to the conclusion that objective ratings are not well correlated with subjective assessments, especially when floors have soft coverings. The variation in conclusions of the above mentioned and other reports will now be examined.

In 1967 Marinier and Hehman [4] reported that floors having the same impact noise rating may differ by almost a factor of 4 in loudness. However, they were unable to describe the significance of this range due to a lack of reliable subjective or objective criteria of acceptability of floors. They commented that this range would span from acceptable floors to totally unacceptable floors.

More recently an extensive survey was carried out by Bodlund [5], and this revealed that a good correlation was achieved between the mean impact sound index as determined by ISO 140 and 717 and the corresponding subjective mean score of each housing area and party construction tested. Comment was made on the fact that the shortcomings of the current method of evaluation seem to arise when test results and subjective rating are compared for lightweight timber joist floors and heavy concrete constructions. These results broadly agree with an earlier study carried out by Olynyk and Northwood [6] who concluded that ratings, (in this case fHA rating), exaggerated the differences between floors in the most unacceptable range but that there was a reasonable degree of correlation among floors tested.
Differences in correlation - floor coverings

A possible explanation for the difference in correlations reported above is the variation in floor types used and the use of floor coverings. In Mariner and Hehman’s study and the report by Olynyk and Northwood various floor types were used, including timber, concrete and composite constructions. The initial part of Bodlund’s survey, which yielded very good correlation, was based on only two party floors, two concrete floors and two timber joist structures, both without additional coverings. In the supplementary part of the survey he looked at a greater variety of floor types including concrete floors with hard coverings. Although the correlation coefficient dropped, it still gave better correlation than has been reported in other studies.

Differences in correlation - subjective parameters

It is possible that experimental procedure employed by the various researchers may contribute to the differences in correlation obtained between objective and subjective results. The poor correlation reported by Mariner and Hehman and the results of Olynyk and Northwood’s study may have been influenced by the assumption that loudness is the significant subjective response to be evaluated. The former evaluated the loudness of real footsteps by using the Stevens method, and the latter employed “paired comparisons” of footstep noise.

In Bodlund’s survey the good correlation obtained may be because during the questioning of survey subjects reference was made only to associating satisfaction with the acoustic environment of the home. No reference was initially made to loudness or detectability of footsteps. Therefore, the subjects may be equating annoyance, as opposed to loudness, with the degree of satisfaction. However, it is possible that the results of this survey were in some way affected by investigator-induced response range effects. This may have occurred due to the type of rating scale used for quantifying the subject’s judgement, i.e. the rating scale had an obvious middle and therefore the subject may in a case of indecision select a response too close to the middle of the range.

Loudness versus annoyance

Loudness is defined as the subjective intensity of sound, independent of any meaning the sound might have. Whereas annoyance commonly signifies one’s reaction to sound based both on its physical nature and its emotional content and novelty (which are excluded from perceived noisiness).

Does the rank order of floors based on a subjective assessment of loudness agree with the rank order of floors based on a subjective assessment of annoyance? The works of Stevens in 1972 would seem to suggest that this will be the case. He stated that “the evidence from experiments in several laboratories suggests that loudness and noisiness may be considered to be essentially synonymous”. He was of the opinion that a fair assessment of the annoyance which a noise has, may be made from the evaluation of its loudness.
Despite this statement there is some evidence to suggest that for certain classes of noise, estimates of the loudness are unsatisfactory predictors of the annoyance caused. The classes of noise referred to here are those in which there is a sizeable proportion of low frequency noise. From experiments carried out to date it would appear that footstep spectra contain a higher proportion of low frequency energy in the receiving room than do tapping spectra. For example, 91% of the total energy produced by a footstep on a timber floor occurs within the 31.5Hz octave band. While the tapping machine on the same floor, only 4% of the total energy occurs in the same octave band. To investigate the possibility of loudness not being directly related to annoyance with footstep signals it was proposed to evaluate the subjective response in terms of:

(a) loudness
(b) annoyance

Improvements in correlation?

The initial task was therefore one of determining whether or not the new reference curves suggested in the study by the Building Research Establishment and C578 (1988) would give a significant improvement in the correlation obtained in terms of annoyance and/or loudness.

This stage of the investigation involved looking at different floor constructions in the laboratory. Objective measurements were made in the 25 - 3150Hz range. This resulted in three single figure ratings:

(1) Rating according to ISO 717/2
(2) Rating according to BRE/C578 and this is using a straight line which increases at 3dB/octave starting at 199Hz and going up to 3150Hz.
(3) Rating according to BRE/C578 only this time the straight line starts at 50Hz and goes up to 3150Hz.

Footsteps

Measurements were then made, in the same frequency range, of real footsteps on each of the floor constructions, and these were recorded on tape or synthesised from the original recording.

Evaluation of subjective response

For the evaluation of loudness subjects are presented with footstep signals through earphones. The signal is presented an octave at a time. The subject will also hear pink noise, the level of which is capable of being increased 3dB at a time. The level of the masking sound is noted when the subject indicates that the footstep signal can no longer be heard. In this way an overall estimation of the footstep signal can be made using the regular methods of loudness estimation.

For the assessment of annoyance subjects are presented with footstep signals through earphones. They are given a sheet of instructions and a response sheet. The responses vary from no annoyance to extreme annoyance. The procedure used here was developed by Berglund, used by Broner and Leventhall, and modified for use in this study.
Inter subject variability in assessment of annoyance

The assessment of annoyance is slightly more complex than an assessment of loudness due to the individual susceptibility to noise. Obviously what is considered annoying to one person may be quite acceptable to another. It is therefore necessary to try and determine the individual's susceptibility to noise before drawing any conclusions regarding the rank order of annoyance. Personality testing was found to be the most effective means of doing this. An in depth literature survey indicated that probably the best way of doing this was to use the shortened form of the noise sensitivity test developed by Bergman and Pearson. Prior to taking part in the study all subjects will be asked to complete the short questionnaire.

Annoying aspects of footstep signal

When looking at the possibility of simplifying the impacting device it is essential, when trying to maintain good correlation with full test method, that particular aspects of the footstep signal which people find to be most annoying be identified. This was done by presenting the subjects with shaped footstep sounds through earphones. The procedure is similar to that outlined earlier for the annoyance evaluation of the full signal. The signals are shaped so that they all have the same 'A' weighted level thus enabling variables such as gradient, peaks at different frequencies and amount of low frequency energy in the signal to be evaluated in terms of annoyance.

Conclusions

The correlation between objective ratings and human reaction to impact noise can now be looked at in terms of both loudness and annoyance. The degree of agreement in correlation obtained in terms of loudness can be compared with the results of other studies.

References


FREQUENCY RANGE CONSIDERATIONS ON SINGLE NUMBER INDICES FOR ASSESSING ACOUSTICAL INSULATION

Keith W. Walker
USG Corporation
P. O. Box 460, Round Lake,
Illinois 60073
USA

Introduction

Single number sound insulation ratings have been used for many years to determine if the acoustical performance of interior walls between dwellings, offices, etc. was adequate to provide speech privacy and control of radio and television sounds. The Sound Transmission Class (STC)\(^1\) and ISO 717 Weighted Sound Reduction Index (R\(_w\))\(^2\) were designed for these purposes, but they were never intended to be used to describe sound insulation performance against outdoor traffic and other sounds with strong low frequency content. Despite these limitations, these rating methods have been used many times to select and compare the performance of exterior walls, windows and doors, with resultant failure to achieve satisfactory results. This paper will demonstrate the limitations of STC and similar ratings when compared with the loudness reduction of a series of lightweight design walls in the range STC 30 - STC 69, which were measured for sound transmission loss (TL) from 50 - 5000 Hz, and will offer an alternative rating method based on A-weighted sound reduction\(^3\).

Statistical Studies

Using linear regression, this project studied the correlation between STC and the loudness reduction (D\(_L\)) calculated using ISO 532 B\(^4\) for a series of 42 gypsum board and steel stud walls subjected to three assumed transportation sound sources\(^5\) and speech, see figure 1. Figure 1 also shows an averaged source which is used later. The slope, intercept, correlation coefficient and standard deviation from the slope were calculated for STC vs. D\(_L\) for each sound source. First, the loudness of each source was calculated in PhonGC (F designates a free field condition). The building interior sound levels were then calculated by subtracting the measured TL from the sound source level for the 50 - 5000 Hz range; no correction was made for room sound absorption. The indoor loudness was calculated in PhonGN (D designates a diffuse field condition) and subtracted from the source PhonGC to obtain the D\(_L\). Figures 2, 3, 4 and 5 plot the STC against D\(_L\) for each sound source and display the statistical data. STC is shown to work well for speech but is seriously deficient as a descriptor when used with the other sources. With a t-intercept of 15.2 and a slope of 1.094 on figure 5, STC 50 overestimates D\(_L\), by 15 points with a standard deviation of 0.1 dB. Similar studies by the author on R\(_w\) and the USA Federal Aviation Administration's Exterior Wall Rating EWR\(^7\) have shown little improvement over STC even though R\(_w\) includes the 100 Hz one-third octave band.
Development of a New Rating Method

Several attempts were made to develop an improved version of the STC method. Changing the STC grading curve shape only or changing the curve fit method to be more controlled by the low frequencies was not adequate since the standard test range does not go below 125 Hz. Some worthwhile improvement was achieved by extending the range down to 50 Hz. Few laboratories have rooms of a size which permits reasonable test accuracy down to 50 Hz. Even if large rooms were available, when the wavelength is longer than the test wall dimensions, the transmission loss is largely controlled by the wall stiffness and is often dependent upon how the wall is tied into the surrounding test frame. The low frequency TL dependence on the mounting method is significant since there is no way...
to ensure that the wall stiffness can be replicated in the field, particularly in non-masonry building structures. It is unreasonable to expect laboratories to provide data to 50 Hz on a routine basis and, even if available, the information has a low credibility.

Finally, it was determined that a calculation of A-weighted sound reduction provided a significantly improved correlation with $D_a$. Since it would not be reasonable to routinely perform calculations for every type of sound source, the three selected outdoor sound sources were equalized for dBA level, averaged for spectrum shape and given the levels shown on figure 1. A-weighted sound levels were calculated in eq. 1 by adding the corrections published in IEC 123** to each third octave band sound level in the frequency range of interest (3 ranges, 50, 60 and 100 Hz to 5000 Hz) and summing the corrected levels:

$$ L = 10 \log \frac{(SPL, + W_r)/10}{10} \text{ dB} \quad (1) $$

$L = \text{A-weighted sound level}$
$f = \text{one-third octave frequency bands in the required range}$
$SPL_r = \text{sound pressure level in each frequency band}$
$W_r = \text{A-weighting correction for each frequency band}$

The A-weighted source sound level $L_s$ was calculated for the averaged spectrum of figure 1. $L_s$ for the receiving side of each wall was obtained by substituting SPL$_r$, for SPL, in equation 1, SPL$_r$, being derived from:

$$ \text{SPL}_r = \text{SPL}_r - TL, \text{ dB} \quad (2) $$

where $SPL_r$ is the sound level on the receiving side, SPL$_r$ is the sound level on the source side and TL is the partition sound transmission loss for each one-third octave band.

The A-weighted sound reduction, designated Outdoor-Indoor Transmission Class (OITC) is then:

$$ \text{OITC} = L_s - L_r \text{ dB} \quad (3) $$
available, it is still a significant improvement over STC or R.<br>

Conclusions<br>

The STC and (by implication) R<sub>n</sub> ratings are not effective for characterizing the effectiveness of walls in providing protection from transportation noise. The use of frequency band limited A-weighted sound reduction based on a fixed spectrum shows promise. The calculation of A-weighted reduction is simple and the rating is relatively simple to explain to the layman. Until transmission loss data in the 00 HZ one-third octave band is available, the method could temporarily use the 100 - 5000 HZ range. Further limitation to 3150 HZ would result in little change in the OITC. Since OITC has not been verified with sounds other than those described in this paper, its use should be limited to transportation noise until further statistical work is performed.

References

5. Railroad Noise, unpublished data, courtesy USG Corporation.
COMPARISON OF SINGLE-NUMBER INDICES FOR ASSESSMENT OF SOUND INSULATION OF BUILDING WALLS BASED ON PSYCHOPHYSICAL EXPERIMENT

Masaru Koyasu* and Hideki Tachibana**

*Acoustic Engineering Laboratory
Shinjuku-ku, Tokyo, 160 Japan
**Institute of Industrial Science, University of Tokyo
Roppongi 7-22-1, Minato-ku, Tokyo, 108 Japan

Introduction
For the assessment of sound insulation efficiencies of building walls, there exist international and national standards. These standards specify different ways of rating the sound insulation efficiencies. In this paper, sounds transmitted through walls with various sound insulation characteristics were electronically synthesized, and loudness hearing tests were performed. Based on the experimental results, the appropriateness of various single-number measures was examined.

Experimental System
The experiments were performed in an anechoic room with a loudspeaker source. As shown in Fig.1, the loudspeaker system consisted of eight flat-type woofers and a small loudspeaker for middle and high frequencies. The hearing position was set at a point 2m from the loudspeaker system.

The test sounds were electronically synthesized by using digital technique. That is, the source signals of 11.9 s duration time were put into a computer, and they were divided into signals in each 1/3 octave band from 50Hz to 5kHz by digital filtering. They were weighted in each 1/3 octave band according to the sound insulation characteristic of a hypothetical wall, and then synthesized as the sound transmitted through the wall. At the same time, the frequency characteristic of the loudspeaker was corrected by inverse-filtering technique. Thus synthesized signals were recorded on a digital-audiotaperecorder (DAT) and reproduced for the hearing tests.

Test Sound
Various sound insulation characteristics of building walls were classified and exemplified, and eleven characteristics shown in Fig.3 were assumed. Among them, "H" is the most basic characteristic of 5dB per octave determined by the mass law. "I" is the characteristic based on the reference curve specified in ISO 717-1 and ASTM E413, and "J1" is the one based on the reference curve specified in JIS A 1419.

As the incident noises to walls, two noises were adopted: one was a model noise (M-noise) which represented general indoor noises and had dominant components in middle frequency bands between 250 Hz and 1 kHz, and the other was a rock music. The spectral characteristics in octave bands of these source sounds are shown in Fig.4.

---

Fig.1 Loudspeaker system
Fig.2 Presentation of test sounds

115
Fig. 3 Modelling of sound insulation characteristics of building walls

Test Procedure and Condition

The method of adjustment by subject was adopted as the test procedure. That is, the standard stimulus (Ss) and the comparison stimulus (Sc) were presented alternately as shown in Fig. 2, and the subjects were instructed to adjust the magnitude of Sc with a remote control attenuator until it was perceived to be equally loud as Ss. Each loudness matching test was performed in both conditions that Sc was made louder and enough softer than Ss, and the point of subjective equality (PSE) for each comparison stimulus judged by each subject was obtained as the mean value of such downward and upward tests.

As the standard stimulus, a noise with frequency characteristic of -5 dB per octave was adopted. The magnitude of the standard stimulus was fixed at relatively low levels of 30dB(A) and 20dB(A) in Leq. Since the experimental results obtained under these conditions were almost the same, the results concerning the latter condition (Ss - 30dB(A) A) are presented below. Fourteen male and female subjects in late teens to forties with normal hearing abilities participated in the hearing tests.

Experimental Results and Discussion

The mean value of PSE judged by all of the subjects was calculated. Four examples of test sounds (hypothetical transmitted sounds) which were judged equally loud are shown in the middle of Fig. 4.

The results of loudness matching tests were arranged by various well-known noise assessment measures. As a result, it was reconfirmed that the loudness judgement and the precise noise measures like Loudness Level proposed by Zwicker are in good accordance. Besides, it was found that the arithmetic mean value of sound pressure levels in octave bands from 63Hz to 125Hz to 4kHz is a considerably good measure for loudness estimation of this kind of sounds. This conforms well to the results of our previous studies. [1, 2, 3]

Next, the sound insulation characteristics of the hypothetical walls judged equally efficient were calculated from the sound pressure levels of the assumed sound sources and the matched hypothetical transmitted sounds. In this case, the magnitude of the sound sources was assumed as 90dB(A). Four examples of sound insulation characteristics which were judged equal in sound insulation efficiency are shown in the bottom of Fig 4.

To evaluate all of the experimental results, the following seven single-number measures were applied. (1) D(ISO) : according to ISO 717/1, (2) STC : according to ASTM E413, (3) D(JIS) : according to JIS A 1419, (4) WD : a similar measure as STC, but the
reference curve specified in JIS A 1419 is applied (see Fig.3). (5) $\bar{\Delta}L(125-4k)$ : the arithmetic mean value of sound insulation efficiencies (in dB) in octave bands from 125 Hz to 4kHz, (6) $\Delta L(63-4k)$, similar value from 63Hz to 4kHz, and (7) $\Delta L_A$ : the level difference in A-weighted sound pressure level. The results arranged by these measures are shown in Fig.5. In each figure, eleven hypothetical walls (on the abscissa) were judged equal in sound insulation efficiency. Therefore, it could be expected that each plot would lie on a horizontal line if the assessment measure were proper. A comparison of the results shown in Fig.5 indicates that the results arranged by $O(ISO)$, STC, WD, and $\Delta L_A$ are almost the same, and the results arranged by $\bar{\Delta}L(125-4k)$ and $\Delta L(63-4k)$ are most converged on the horizontal lines among the seven results in both figures.

Conclusions

Regarding the appropriateness of a single-number measure for assessment of sound insulation efficiency of building walls, we may conclude as follows within the experimental results obtained in this study.
1) Among $O(ISO)$, STC and $\Delta L_A$, apparent superiority or inferiority can not be seen.
2) As compared with them, the correspondence between the loudness judgement and $O(JIS)$

Fig.4 Spectral characteristics of sound sources ($M$-noise and a rock music), hypothetical transmitted sounds judged equally loud, and sound insulation characteristics of walls judged equally efficient.
Fig. 5 Comparisons of sound insulation characteristics judged equally efficient, evaluated by seven kinds of single-number indices

is a bit worse. This might be ascribed to the fact that this measure is based on "the peak fitting method". However, if using the reference curve specified in JIS A 1410 and applying "the curve fitting method" as in STC method, the correspondence with the psychoacoustic response is much improved as is seen in the results arranged by MD. (This suggests that the slight difference in the reference curves is not so serious.)

3) Within the results in this study, the arithmetic mean value of sound insulation indices (sound transmission loss or sound pressure level differences) in octave bands from 63 Hz or 125 Hz to 4 kHz is the best. Although this measure is not being used nowadays, it is necessary to be reconsidered as a convenient single-number index for the assessment of sound insulation efficiency of building walls.

4) In this study, loudness response was mainly investigated. However, annoyance is also important in the sound insulation problem, and it should be examined by other sophisticated psychoacoustic investigations.

References
3. A
ZUR ABSCHÄTZUNG DES FREQUENZGANGES DER REFLEKSIONSFAKTOREN BEI RECHTECKIGEN PLATTEN

L. Cremer

Technische Universität Berlin, Institut für Technische Akustik, Einsteinufer 25, 1000 Berlin 10

Als ich bei der Beratung des Herkulessaales in München /1/ von dem Architekten, Prof. Escher, gefragt wurde, ob der von mir über dem Orchester verlangte Reflektor nicht in mehrere kleine aufgeteilt werden könne, war mir klar, damit eine Beschränkung ihrer Wirkung nach tiefen Frequenzen in Kauf zu nehmen. Bei der Suche nach einer Grenzfrequenz \( f_g \), oberhalb welcher eine einzelne Rechteckplatte den geometrischen Reflexionsfaktor 1 einer unendlich ausgedehnten erreicht, verließ ich mich auf die Methode der Plattenunterteilung in Fresnelschen Zonen, wie sie von Lord Rayleigh /2/ beschrieben wurden.

Er geht bei ihrer Einführung von einem kreisrunden Reflektor (siehe Bild 1) aus, bei dem sich zusätzlich Sendepunkt S und Empfangspunkt E auf dessen Mittelachse befinden. Bei diesem rotationssymmetrischen Feld erscheinen die Flächenelemente als Ringelemente:

\[
ds = 2\pi r \, dr = \pi d(r') .
\]  

(1)

Er nimmt wie Fresnel ferner an, daß Senderabstand \( l_S \) und Empfängerabstand \( l_E \) so groß gegen den Reflektorradius \( a \) sind, daß sowohl der Reflektor vom Sender als homogen bestrahlt angenommen werden kann, wie von dort aus gleichstarke Strahlen den Empfangspunkt treffen. Die Beiträge unterscheiden sich aber durch ihre Phasen, der wesentlichen Ergänzung, die Fresnel dem Huygens'schen Prinzip hinzufügte. Unter den gemachten geometrischen Annahmen legt der über \( ds \) laufende Strahl gegenüber der Verbindung \( SE = l_S + l_E \) einen Zusatzweg von:

\[
\Delta l = (r^2/2) \left( 1/l_S + 1/l_E \right) 
\]

(2a)

zurück, bleibt also in der Phase um

\[
\phi = (\pi r^2/\lambda) \left( 1/l_E + 1/l_S \right) 
\]

(2b)

hinter der des Mittelstrahls zurück. Da dieser Phasenwinkel als einzige Variable (bei Zeigerdarstellung im Exponenten) auftritt, und sie genau wie \( ds \) proportional zu \( r^2 \) ist, erhält man bei der Berechnung des Verhältnisses des Druckes bei E zu dem Druck, der dort bei Reflexion an unendlicher Wand zu erwarten wäre, \( P_{pp} / P_{p} \), das wir auch als wirksamen Reflexionsfaktor in Reflektormitte \( R(0) \) ansehen können, ein explizit lösbares Integral /3/:

\[
R(0) = \int_0^a \exp(-j\phi) \, d\phi = -j(1 - e^{-j\phi_a}) ,
\]

(3a)

mit

\[
\phi_a = \pi a^2 \left( 1/l_S + 1/l_E \right) .
\]

(3b)

als Phasenachseilung des Randes gegenüber der Mittel. Bild 2 zeigt die Zeigerkonstruktion des Ergebnisses, bei
der Randstrahl einen Kreis um den Endpunkt des Mittelstrahls mit wachsenden \( \phi \) durchläuft, der bei Null beginnt. Dies leuchtet, wie auch das in Bild 4 gezeigte sinusförmige Anwachsen des Betrages von \( R \) mit wachsender Fläche, physikalisch ein. Aber gerade dieses Anfangsgebiet verstößt gegen die von Fresnels vom Huygens'schen Prinzip übernommene Annahme, daß die Quellenverteilung im Spalt, bzw. auf einem Reflektor durch die einfallende Welle gegeben ist. Wir haben an Rande Abweichungen zu erwarten, die freilich um so weniger wiegen, je mehr Wellenlängen der Reflektor breit ist. Nach Leisner /4/ genügen bereits 2 Wellenlängen. Aber Fresnel war vielmehr daran interessiert, daß mit wachsendem \( \phi \) immer wieder die Werte 0 und 2 erreicht werden bei

\[
\phi_n = n\pi
\]

(4a)

und zwar das erste bei geraden, das zweite bei ungeraden \( n \). Diesen Werten entsprechen bei gegebenen \( \lambda, l_S \) und \( l_E \) Radien

\[
a_n = \left[ n\lambda/(1/l_S + 1/l_E) \right]^{1/2}
\]

(4b)

die gleiche Teilflächen mit einander entgegengewirkenden Phasen begrenzen, die Rayleigh als Fresnelsche Zonen bezeichnete und die Interferenzerscheinungen leicht zu beschreiben gestatten.

Diese Unterteilung in gegenphasige Gebiete erleichtert die Übersicht auch dann, wenn die zunächst gleichen Flächen durch eine nicht konzentrische Begrenzung, für die die in Bild 4 wiedergegebene unendlich breite Brüstung als einfachstes Beispiel wiedergegeben ist, allmählich abnehmen. Rayleigh weist ausdrücklich darauf hin, daß dabei jede Zone durch die Hälfte ihrer Nachbarn kompensiert wird, und daß so die erste halbe Zone zum Representanten der vom Rande unbeeinflußten Reflexion an unendlicher Platte wird. Dies wird in Bild 4 erstmals erreicht, wobei der Randstrahl dem Mittelstrahl um \( \pi/2 \) nacheilt. Diese Nacheilung hatte ich auch (unter zusätzlicher Berücksichtigung schrägen Einfalls) der Berechnung der Grenz- frequenz der Rechteckplatten im Herkulessaal zugrunde gelegt, die bei dem in Bild 5 gezeigten Quadrat hinauskommt auf:

\[
f_g = c/\lambda g = c/\left[ 2a^2 \left( 1/l_S + 1/l_E \right) \right]
\]

(5)

Rindel /5/ hat sowohl durch die auch schon von Fresnel durcharbeitete integrierende Berechnung /6/, sowie durch Messungen im schalllosen Raum nachgewiesen, Bild 6, daß der Wert \( R = 1 \) bereits eine Oktave tiefer erreicht wird. Man braucht nur Bild 4 und 5 zu vergleichen, um zu erkennen, daß in Falle des Quadrates gar keine gegenphasigen Zonen in Erscheinung treten, die ganze erste Zone, wenn der eingeschriebene Kreis die halbe begrenzt. Das Zonen-Prinzip würde den Näherungswert

\[
2 \left( 4a^2/2\pi a^2 \right) = 1,27
\]

liefern.

Rindel's exakte Integration liefert einen Pegelzuwachs von 4 dB, was einem R von 1,6 entspricht.

Ich habe nun die Zonenschätzung dadurch verbessert, daß ich die Ringelemente beibehalte, aber berücksichtigte, daß sich ihr Bogenvinkel nach Berührung mit irgendeinem vom konzentrischen Kreis abweichenden Rand von \( 2\pi \) um einen vom Radius \( r \), also von \( (\phi)^{1/2} \) abhängigen Faktor verringerte. Die Inte-
grale (3) nehmen dadurch die Form an:

\[
R = \int_0^\pi \gamma(VF) e^{-j\phi} \, d\phi
\]  

(6)

Beim Quadrat führt das auf:

\[
\gamma(VF) = 1 - \frac{4}{\pi} \arccos \sqrt{\frac{r}{\phi}}
\]  

(7)

Da das Integral außerhalb des eingeschriebenen, durch \( \phi \) gekennzeichneten Kreises keine explizite Lösung zulässt, ist man im Bereich \( \phi \), bis \( \phi = 2 \phi \), auf die Summierung von Beiträgen genügend schneller Ringgebiete angewiesen. Ich konnte so nicht nur den Fall \( \phi = \pi/2 \), sondern auch den zur Rindel'schen Grenzfrequenz gehörigen Wert \( R = 1 \) verifizieren.

Die bei den Fresnel'schen Integralen verwendeten kartesischen Koordinaten \( x, y \) mit dem Flächenelement \( dx \) sind nicht nur den rechteckigen Rändern besser angepaßt, sie erlauben sogar, da sich \( r' \) additiv in den Exponenten in \( x^2 + y^2 \) aufspaltet, eine multiplikative Aufspaltung des Gesamtintegrals in ein nur von \( x \) und ein nur von \( y \) abhängiges Integral, deren Grenzen von den Abständen des Reflexionspunktes von den zu \( x \) und \( y \) senkrechten Kanten bestimmt werden. Dadurch war Rindel in der Lage, auch außerhalb der Reflexionspunkte zu untersuchen.

Während die obere Kurve in Bild 7 sich auf 20 lqR in der Mitte bezieht, gehört die untere zu einem Punkt in Randmitte des Quadrats. Beide zeigen Schwankungen zwischen maximalen und minimalen Werten, die bis in den Ultraschallbereich kaum abnehmen. Nun dürfte bei den oberen Kurven das 4 Oktaven über Rindel's Grenzfrequenz erscheinende erste Minimum bereits keine praktische Bedeutung mehr haben. Dagegen machen sich diese Interferenzen am Rande bei tieferen Frequenzen bemerkbar, was Rindel auch an Hand berechneter Querverteilungen zeigte. Auf Bitte des Autors berechnete auch Heckl für engere Schritte solche (s. Bild 8) und ermittelte daraus mittlere R-Werte (Bild 9), die für die Dimensionierung der Reflektoren noch wichtiger sind als die in der Mitte gewonnenen. Sie schwanken oberhalb der Rindel'schen Grenzfrequenz \( f/2 \) so wenig, daß man sagen kann, auch für sie wird ein durch \( R^g = 0,85 \) gegebenes Plateau bei der Brüstung und wegen der multiplikativen Aufspaltung dessen Quotient 0,72 bei dem Quadrat erreicht /7/.

Literatur

/1/ Cremer, L., (1953) Schalltechnik 13, Nr. 5, S. 1
/2/ Lord Rayleigh, (1896) Theory of Sound, Vol. 2, § 283

123
THE INFLUENCE OF DIFFUSE SURFACE REFLECTIONS ON ROOM SOUND FIELDS

Murray Hodgson

Institute for Research in Construction, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6

INTRODUCTION

In recent years considerable interest has been shown in the influence of diffuse surface reflections on room sound fields. With respect to factories, it has been suggested that such diffusion must be taken into account in order accurately to predict the reverberation time and noise levels at large source/receiver distances [1]. With respect to rooms for speech and music, it is considered that diffuse reflections improve the acoustic environment and its uniformity; considerable effort has been made to include diffusely reflecting surfaces in new concert halls.

The objective of the present study was to use ray-tracing techniques to investigate the influence of surface diffusion on various aspects of the sound field in three very differently-shaped rooms.

PREDICTIONS MADE

Predictions were made using a ray-tracing technique developed by the INRS in France to predict factory noise levels [2]. It was modified to incorporate variable diffuse surface reflection and predict the unweighted echogram, \( I(t) \), and the cos-weighted echogram, \( I'(t) \). The cos-weighted echogram is obtained by weighting the energies of rays incident on the receiver position by \( \cos \) (angle between the ray and the horizontal axis perpendicular to the line joining the source and receiver); this is used to determine the lateral energy. From these echograms the sound decay (SD), reverberation time (RT) and the following variables used to describe the acoustic conditions in rooms were then determined (times in ms):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol and definition</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early energy</td>
<td>( E_T = \int_0^T I(t) , dt )</td>
<td>Measure of early energy</td>
</tr>
<tr>
<td>Cos-weighted early energy</td>
<td>( E'_T = \int_0^T I'(t) , dt )</td>
<td>Measure of early lateral energy</td>
</tr>
<tr>
<td>Sound propagation</td>
<td>( SP = 10 \log (E_{in}) )</td>
<td>Steady-state noise level minus the source power</td>
</tr>
<tr>
<td>Loudness</td>
<td>( G = SP + 3l , dB )</td>
<td>Concert hall sound strength</td>
</tr>
<tr>
<td>Early decay time</td>
<td>( EDT = ) time of sound decay from 0 to -10 dB</td>
<td>Subjective reverberance</td>
</tr>
<tr>
<td>Early/late energy</td>
<td>( C_T = 10 \log \left( \frac{E_T}{E_{in} - E_T} \right) )</td>
<td>( C_{50} ) - speech intelligibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( C_{80} ) - concert hall clarity</td>
</tr>
</tbody>
</table>
Spatial impression

\[ SI = 14.5 \left( \frac{E_8}{E_0} - 0.05 \right) \]

Concert hall quality

E and E' predictions were made for T = 80 ms and T = 200 ms. Predictions were made of all of the above variables for three very differently-shaped rooms as follows (see also Fig. 1):

'CUBE'- 27.5 m x 27.5 m x 27.5 m, absorption coefficient = 0.1,
source and receiver at half height and width,
R = 5 m and 20 m (echogrammes), R = 5, 10, 15, 20 m (SP, G);

'FLAT'- (idealized factory), 110 m x 55 m x 5.5 m high, absorption coefficient = 0.1,
source and receiver at half height and width,
R = 5 m, 20 m, 80 m (echogrammes), R = 5, 10, 15, 20, 30, 90 m (SP, G);

'HALL'- (idealized concert hall), 40 m x 24 m x 20 m high, absorption coefficient =
0.07 except sloping floor and rear wall = 1.0, source and four receiver positions as shown in Fig. 1.

The air absorption was assumed to be zero. Predictions were made with all surfaces 100% specularly reflecting, 100% diffusely reflecting and, except 'HALL', 50% specularly and 50% diffusely reflecting.

RESULTS

Results are presented qualitatively since the tendencies are of greatest interest. Regarding the single-valued variables, the arrows in the table indicate how they changed with increasing surface diffusion. Note that the variables related to fixed times (i.e., SI, C_{50}, C_{80}) depend strongly on the room size making general conclusions difficult to draw from specific cases.

Fig. 2 shows the unweighted and cos-weighted echogrammes for D = 0, 0.5 and 1.0 for FLAT, R = 20 m as an example. In all cases considered, increased diffusion caused a smearing and smoothing of the echogrammes after the initial time gap. Increased diffusion tended to decrease the early energies and the spatial impression in CUBE and HALL at all positions and in FLAT at R = 80 m; these variables increased in FLAT at R = 5, 20 m.

Figure 3 shows the variation of SD with diffusion in CUBE and FLAT at R = 20 m. In all cases considered, increasing diffusion increased the rate of sound decay and, especially in FLAT, made the decay more exponential. That diffusion increases the decay rate in a cubic room is somewhat surprising and shows that the sound field in a cubic room with specularly reflecting surfaces is not diffuse. In CUBE and FLAT increasing the diffusion tended to decrease the EDT and RT; however they increased with increasing diffusion in HALL. This behaviour is probably related to the distribution of surface absorption.

Increasing diffusion resulted in lower SP and G levels in all cases. This is again surprising in the case of CUBE and is again related to the non-diffuse field.
associated with specular reflection. Increasing surface diffusion had less consistent effects on $C_{50}$ and $C_{80}$. In CUBE and FLAT at $R = 5$ and 20 m the initial increase increased these quantities; the final increase then decreased them. In FLAT, $R = 80$ m the effect was small. In HALL diffusion decreased $C_{50}$ and $C_{80}$ at all positions.

CONCLUSIONS

The results have the following implications for factory noise prediction and the design of factories, and rooms for speed and music:

FACTORIES - the SP and RT decrease with diffusion in CUBE as in FLAT suggests that surface diffusion is not the (entire?) explanation for the results in [1]:

SPEECH ROOMS - wall diffusion may increase or decrease the speech intelligibility depending on the room, and decreases speech levels:

MUSIC ROOMS - wall diffusion may increase or decrease early energies and clarity, and reduces spatial impression and sound strength.

In conclusion, the predictions made in this study do not demonstrate that uniformly-distributed wall diffusion is a generally useful design feature of rooms.

REFERENCES


**TABLE**

Predicted change of variable with increasing surface diffused. ($\uparrow$ = increase, $\downarrow$ = decrease, $\rightarrow$ = decrease then no change, etc.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$R = 5$ m CUBE</th>
<th>FLAT</th>
<th>$R = 20$ m CUBE</th>
<th>FLAT</th>
<th>$R = 80$ m FLAT</th>
<th>HALL R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{80}$</td>
<td>$\rightarrow$</td>
<td>$\rightarrow$</td>
<td>$\downarrow$</td>
<td>$\rightarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$E_{200}$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$E'_{80}$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$E'_{200}$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>SP/G(dB)</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\rightarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>EDT (s)</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
</tr>
<tr>
<td>RT (s)</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
</tr>
<tr>
<td>$C_{50}$ (dB)</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\rightarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$C_{80}$ (dB)</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\rightarrow$</td>
<td>$\uparrow$</td>
<td>$\uparrow$</td>
<td>$\rightarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$SI$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\uparrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
</tbody>
</table>
Fig. 1 Plans, sections and dimensions of the rooms CUBE, FLAT and HALL.

Fig. 2 Predicted unweighted (top) and cos-weighted (bottom) echograms for FLAT, $R_a=20$ with $D=0$, 0.5, 1.0. Note the change of vertical scale.

Fig. 3 Predicted sound decay in CUBE and FLAT, $R_a=20$ m with $D=0$ (-----), $D=0.5$ (---) and $D=1.0$ (-----).
INTRODUCTION OF IMPULSE RESPONSE IN A SOUND FIELD BY THE FINITE ELEMENT METHOD

Seokjoo CHOI and Hideki TACHIBANA

Institute of Industrial Science, University of Tokyo
Roppongi 7-22-1, Minato-ku, Tokyo, 106 Japan

INTRODUCTION

The impulse responses in a room are very informative for the assessment of room acoustics. To calculate the impulse responses, various theoretical investigations have been made by the computer simulation technique based on the ray theory [1-2], the virtual source method [3], and the boundary integral equation method [4].

The finite element method (FEM) has been widely applied to the numerical analysis of acoustical problems [5-6], but it has not been applied to the estimation of impulse response in a sound field. In this paper, a numerical simulation by the FEM has been applied to the estimation of impulse responses in a two-dimensional sound field. First, the transfer function in frequency domain of steady state sound field was calculated by the FEM, and then, the impulse response was obtained by making the inverse Fourier Transform of the transfer function.

To examine the validity of the calculation, 1/4 scale model experiment was conducted, and these results were compared.

EXAMPLES OF NUMERICAL SIMULATIONS

Case-1: Verification of Radiation Problem

To examine the accuracy of the FEM calculation, a sound field around a vibrating 2-dimensional (cylindrical) sound source with 0.5 cm radius was calculated. As shown in Fig.1, the fan-shaped area of the sound field was divided into 96 bi-quadratic elements and 567 nodes. The vibration velocity of the surface of the hypothetical sound source was assumed to be 1 m/s over all frequencies. The sound pressure level at the point A (1 m from the origin) calculated by the FEM is shown in Fig.2, being compared with the theoretical value. They are in fairly good agreement.

![Fig.1 Sound field around a vibrating surface.](image1)

![Fig.2 Sound pressure level at the point A.](image2)
Fig.3 2-dimensional sound field studied by FEM calculation and 1/4 scale model experiment.

These results indicate that the sound pressure increases in proportion to the square root of frequency. This means that, to obtain a flat response of sound pressure, the frequency characteristic of the vibration velocity of the sound source should be inversely proportional to the square root of frequency. The following calculations were made under this condition.

Case-2: Impulse Responses in a 2-dimensional Sound Field

The 2-dimensional sound field shown in Fig.3 was analyzed. In the calculation, the sound field was divided into 248 elements and 1047 nodes. Each element consisted of 9-node bi-quadratic elements. As the boundary conditions, a sound pressure reflection coefficient of 0.85 was assumed over the frequency range from 81 Hz to 400 Hz (The phase shift at the boundary was ignored). To obtain the impulse response of 1 second duration, the transfer function was calculated at an interval of 1 Hz. From the results, the impulse responses were obtained. In this calculation, the frequency weighting according to the frequency characteristic of the sound source used in the model experiment was made to compare the results with the experimental results (see Fig.4). The temperature when the experiment was performed was considered in the calculation.

For experimental study, a 1/4th scale model of the 2-dimensional sound field was constructed with vinyl chloride plate. A loudspeaker was attached to the outside as shown in Fig.3, and a rectangular impulse of 0.1 ms duration was radiated. The sound pressure responses in the scale model were measured at two points (R-1, R-2) through 1/2 in. condenser microphone, and the frequency components from 81 Hz to 400 Hz were filtered by digital filtering technique.

Fig.5 and Fig.6 show the comparisons of the transfer function and the impulse response obtained by the FEM calculation and the model experiment. In these figures, we can see considerably good agreement between the calculated results and measured results on the whole. However,
the amplitudes of the calculated transfer functions are slightly higher than the measured ones. This discrepancy may be caused by the errors in the assumptions of frequency characteristic of the hypothetical sound source and the boundary conditions.

As another study for a reference, the impulse response at the point R-2 was calculated by the geometrical image source method. Fig. 7 shows the comparison of the impulse responses obtained by the image method and by the FEM. In this case, the boundary condition is relatively simple, and these results are in good agreement.

Conclusions

Concerning the calculation of impulse responses in a sound field by the FEM, a basic investigation was made for 2-dimensional sound field, and the calculated results were compared with the experimental ones. As a result, it has been found that this method is effective, in principle. For practical application, however, there are several problems such as the way of setting the characteristics of the sound source and the boundary condition. In addition, the memory size and the CPU time during computation are also serious problems.
Fig. 5 Transfer function and impulse response at R-1.

Fig. 6 Transfer function and impulse response at R-2.

Fig. 7 Impulse responses by FEM calculation and image method at R-2.

References


MEASUREMENT OF IMPULSE RESPONSES IN REAL AND SCALE MODEL AUDITORIUMS

Yoshito HIDAKA and Hidoki TACHIBANA

Institute of Industrial Science, University of Tokyo
Roppongi 7-22-1, Minato-ku, Tokyo, 106 Japan

Introduction

In auditorium acoustics, various quantities such as reverberation time, sound pressure distribution, echo time pattern and other parameters like Deutrichkeit are measured by different methods. Almost all of these quantities, however, can be obtained from impulse response, in principle. Further, by using the digital convolution technique, the "hall tones" can be synthesized from impulse responses and arbitrary sound source signals such as music, speech and so on. According to this technique, we can directly hear and compare different "hall tones" by using arbitrary sound sources. Therefore, we have been investigating the method of measuring impulse responses with high accuracy in real and scale model auditoriums. Some results are presented in this paper.

Measurement of Impulse Responses in Real Auditoriums

In general room acoustic measurements, the sound source with omnidirectionality should be used, basically. For this purpose, we have contrived a dodecahedral loudspeaker system (TS-12M) shown in Fig.1. In the acoustic surveys of auditoriums that we have made so far, this source system have been used as an omnidirectional sound source. \(^1\)

For the measurement of impulse responses, the most basic and direct method is to use an impulsive signal. In this method, however, it is rather difficult to get sufficient energy for high S/N ratio over a wide frequency range. For this reason, other methods have been contrived so far, such as the time-stretch and compression technique (the sweep pulse method) \(^2,3\) and the cross-correlation technique using pseudo-random noise. \(^4\)

In addition to accurate measurement of impulse response, it is desirable that we can directly hear the transient room response and judge acoustic conditions during the measurement. For this reason, the cross-correlation method using stationary noise is not suitable, though the highest S/N ratio can be realized. In former days, we used the direct method, but recently we are mainly using the sweep pulse method. These two methods were compared in this study as follows.

\[
H(k) = \exp \left[ -\frac{(k-500)^2}{800} \right] \cdot \exp \left[ j\frac{12(k+1)}{10000} \right].
\]

\[
H(k) = 0, \quad k = 2048
\]

\[
H(k) = H'(4096 - k), \quad k = 2048 \pm 4095
\]

Fig.1 Dodecahedral loudspeaker system (TS-12M).
The source signal for the sweep pulse method is expressed by Eq. (1) in frequency domain. The time function was obtained by calculating the inverse Fourier transform of Eq. (1) in a computer. Figure 2 shows the time functions of the stretched pulse (a), its time-compressed signal by deconvolution (b), and their energy spectrum (c). This signal was recorded on a digital audio tape recorder (DAT) repeatedly at an interval of 3 seconds. In this recording, a spectrum equalizer was used to roughly correct the frequency characteristic of the dodecahedral loudspeaker. Figure 3 shows the waveform of the stretched pulse fed to the loudspeaker, its deconvolved waveform, and their energy spectrum. Figure 4 shows the waveform of the sound radiated from the loudspeaker measured at a distance of 8 m in an anechoic room, and its deconvolved signal. In the actual measurements, the trigger pulses for synchronous averaging were supplied to a receiving DAT simultaneously. The digital data recorded on the DAT were transferred to a computer, and then synchronous averaging, deconvolution and other various processing were performed there.

Figure 5 shows an example of the measured result obtained in a concert hall; (a) is the room response to the stretched pulse and (b) is the impulse response derived from it, whereas (c) is the impulse response measured by the direct method using an impulse with 20 μs duration. In these measurements by the two methods, the amplitudes of the source signals fed to the loudspeaker were almost the same, and the averaging was done 32 times in both cases. To compare the two methods from the viewpoint of S/N ratio, the reverberation decay curves were calculated from the impulse responses according to the Schroeder's method. The results are shown in Fig. 6. From the results, it is apparent that relatively high S/N ratio could be obtained by the sweep pulse method.

To correct the frequency characteristic of the sound source system more precisely, we are now investigating the application of the inverse filtering technique.
For subjective tests, the impulse responses were measured through a dummy head system. By convolving the binaural impulse responses with dry music, high quality signals have been synthesized.

**Measurement of Impulse Responses in 1/10 scale model Auditoriums**

The sweep pulse method is effective for the measurement of impulse response when a loudspeaker is available. In the case of scale model studies, however, it is rather difficult to get small and omnidirectional loudspeakers. Therefore, in our present study of 1/10 scale model experiment on auditorium acoustics, the direct method is being applied by using a spark discharge impulse source with high repeatability. Figure 7 shows the waveforms of sound pressure superposed 64 times measured at a distance of 50 cm from the electrodes. As is clearly seen, a fairly good repeatability has been achieved. By averaging these waveforms and cutting off the reflections, the direct sound shown in Fig.8(a) was derived. Although its energy spectrum (b) is not flat, no dips are observed over a wide frequency range, and so the frequency characteristic can be corrected by the inverse filtering technique. Figure 9 shows the impulse response and the frequency characteristic of the inverse filter to realize a flat frequency characteristic from 1.3 kHz to 80 kHz.

We are now making several 1/10 scale model experiments for acoustical design of concert halls and opera theaters in Japan. In these experiments, impulse responses in the model halls were measured by using the spark discharge and a binaural receiving system using a model dummy head. Each pair of impulse responses were convolved with various dry music signals for subjective hearing test. The signal processing of synchronous averaging, inverse filtering and convolution were performed on a personal computer and a host computer. As a result of the preliminary hearing tests, it was found that the sound quality of the synthesized signals are passably satisfactory, and such subjective differences as reverberancy and loudness can be clearly judged and spatial impression can be sensed to some extent. (H.Els et.al. made the first trial of this kind of modeling by using an electrostatic loudspeaker.)
Conclusions

The measuring methods of impulse responses in real and scale model auditoriums were investigated. From the results of experimental studies, the following has been concluded.

The sweep pulse method is useful for the measurement of impulse response with high S/N ratio in real halls. In addition, room response can be subjectively judged at the measurement.

For 1/10 scale model experiment on room acoustics, the way of measuring impulse responses by using a spark discharge impulse source and a model dummy head has been realized. According to this measuring method and digital convolution technique, we can hear the model "hall tones". To improve this simulation technique, we are making further study by using a dome type loudspeaker and the cross-correlation technique.

References

MEASUREMENT OF VERY SHORT REVERBERATION TIMES

Jens Holger Rindel\textsuperscript{1} & Birgit Rasmussen\textsuperscript{2}

\textsuperscript{1) The Acoustics Laboratory, Technical University of Denmark, DK-2800 Lyngby, Denmark
\textsuperscript{2) Bruel & Kjær, 18 Nærum Hovedgade, DK-2850 Nærum, Denmark

Introduction

In speech studios with reverberation times around 0.1 – 0.2 s it is not possible to use traditional measuring technique because of the ringing of 1/3 octave filters. A new measuring technique, which is based on a time reversed analysis, has been used in speech studios in the Danish Broadcasting House and the results have been compared with those from traditional methods. The interrupted noise method and the integrated impulse response method have been used, and the relation between limits for reliable results and the choice of measurement setup parameters have been closely regarded.

Measurement method

A suitable and convenient measurement method is not found in any standard. The international standard ISO 3382 (1975) \cite{isos} describes field measurements of reverberation time, but guidelines for the measurement of short reverberation times are not included. Further, the standard is completely out-of-date, since the main contents is more than 15 years old, and the measurement technique has developed considerably during this period. Some useful hints are found in ISO 354 (1985) \cite{isos2}, but measurement of short reverberation times in studios is outside the field of application and not possible using the method described. Consequently, the measurement procedure must be defined utilizing experience from other practical measurements and the new possibilities offered by modern instrumentation.

In the traditional method a broadband noise source is used and after the source has been switched off the decay curve can be recorded. However, this curve can have strong fluctuations due to the stochastic character of the excitation noise, and several decay curves should be evaluated in each position. Using the measuring technique of today, the decay curves from repeated excitations can be averaged into an ensemble averaged decay curve with much reduced fluctuations. Extending the ensemble averaging to include spatial averaging gives rather smooth decay curves, and it seems reasonable to assume that these curves are a good basis for the evaluation of reverberation time.

Another traditional method is the use of a pistol shot as an excitation signal. By this method the impulse response of the room is measured. As shown by Schroeder \cite{schroeder}, the decay curve can be calculated from a backward integration of the squared impulse response. One major quality by this method is, that there will be no stochastic fluctuations, so one excitation in each position will be sufficient to get a decay curve equivalent to the ensemble averaged decay from an infinite number of excitations using interrupted noise. Spatial averaging can be made by ensemble averaging over all source and receiver positions as described above.

Evaluation of decay curves

Within the last years there has been a tendency to prefer an evaluation range of 20 dB. The last version of ISO 354 \cite{isos2} has followed that tendency. However, the first part of the decay curve (5 dB or so) should not be used for the evaluation.

Limitations caused by bandpass filters

Reverberation time measurements are usually analyzed in 1/3 or 1/1 octave bands. However, such bandpass filters can influence the measurement due to the filter ringing, which can give rise to a characteristic wavering of the decay. According to Ref. \cite{ref4} reliable decay curves are obtained only if

\[
R \cdot T_{60} > 16
\]
where $B$ is the bandwidth of the filter and $T_{60}$ is the reverberation time to be measured. If requirement (1) is not met the evaluated reverberation time can be too short or too long, i.e. the sign and size of the error is not predictable.

However, in Ref. [5] it has been demonstrated that reversing the time signal to the filter leads to much less distortion of the decay curve. It has been found that if the upper 5 dB are excluded from the evaluation the requirement (1) can be replaced by

$$B \cdot T_{60} > 4$$  \hspace{1cm} (2)

It should be noticed that there is no distinct limit between acceptable and unacceptable decay curves.

For measurements in $\frac{1}{3}$ octave bands the limit for reliable results at 100 Hz is changed from 0.7 s to 0.17 s when time reversed analysis is used instead of forward analysis.

**Limitations caused by detector**

When measuring short reverberation times it is important to choose the averaging time of the detector short enough to avoid any influence on the slope of the decay curve. Using a device with exponential averaging (time constant $\tau$) it has been shown in Ref. [4] that the averaging time should obey the requirement

$$T_{av} = 2\tau < T_{60}/14$$  \hspace{1cm} (3)

Here again $T_{60}$ is the reverberation time to be measured. If the requirement (3) is not met the evaluated reverberation time will be too long.

However, since response of the detector is much faster when the signal increases instead of decreasing, it will be of great advantage to use time reversed analysis. According to Ref. [5] requirement (3) can then be replaced by

$$T_{av} = 2\tau < 8 \cdot T_{60}/14$$  \hspace{1cm} (4)

Only the upper part of the decay curve is influenced when time reversed analysis and long averaging times are used. Further investigations reported in Ref. [6] have indicated, that the reverberation times deviate less than 1% from the correct values, if conditions (1) - (4) are fulfilled, and the evaluation range does not include the upper 5 dB of the decay.

The averaging time is connected to the frequency bands so that $B \cdot T_{av} \geq 1$ to obtain reliable results. This is a well known condition for the analysis of stationary signals, but at the first glance it could seem to be difficult to meet in decay measurements. However, ensemble averaging offers a solution to the problem: using this technique a requirement for the number of repeated and ensemble averaged excitations $N_{av}$ should be

$$N_{av} \geq \frac{1}{B \cdot T_{av}}$$  \hspace{1cm} (5)

To minimize the number of excitations, the averaging time should be as large as possible but still fulfill (3) or (4). Looking at these requirements, it is seen that a much larger averaging time is allowed for time reversed analysis than for forward analysis. Hence, fewer excitations are necessary, which is an additional feature of the time reversed analysis. For the actual measurements $T_{av} = 1/128$ sec. and $N_{av} = 12$ were used for forward analysis, while $T_{av} = 1/16$ sec. and $N_{av} = 2$ were used for time reversed analysis.

When the method of integrated impulse response is regarded, this is equivalent to the use of a very long averaging time. So, in that case it will not be relevant to consider (5), and one excitation in each position will be sufficient.

Examples of measured decay curves in the actual studio at the 100 Hz $\frac{1}{3}$ octave are shown in Fig. 1 for noise excitation and in Fig. 2 for impulse excitation.
Fig. 1. Interrupted noise excitation in a 68 m$^3$ studio. 100 Hz 1/3 octave band

A. Ensemble averaged decay curve using six source-microphone positions and forward analysis. 72 excitations in total. The requirement (1) is not met

B. Ensemble averaging as in A, but using time reversed analysis. 12 excitations in total. All requirements are met

Fig. 2. Impulse excitation in a 68 m$^3$ studio. 100 Hz 1/3 octave band. Same positions as in Fig. 1. In both examples the decay curve (solid curve) has been calculated by backward integration of the impulse response (bar graph)

A. Ensemble averaged impulse response and the corresponding decay curve using six source-microphone positions and forward analysis. The requirement (2) is not met

B. Ensemble averaging as in A, but using time reversed analysis. All requirements are met

Description of the studio

A speech studio in the Danish Broadcasting House were selected as a measuring object because the reverberation time was expected to be very low. The volume of the studio was 68 m$^3$ and the studio was furnished with a table and 6 chairs. Two source positions were used, one of them close to a corner with the centre of the loudspeaker 0.4 m from each of the nearest surfaces. Three microphone positions were distributed in each room, 1.0 m, 1.2 m and 1.4 m above the floor. Each source position was used in combination with each microphone position, making a total of six combinations, which were used for spatial ensemble averaging.

Instrumentation and measurement results

The measurements were carried out with the Bruel & Kjær Real Time Analyzer Type 2133. A more detailed description of the use of this instrument for measurement of short reverberation times is given in Ref. [6]. A 6 mm pistol was used for impulse excitation in combination with a -6 dB/octave filter. The latter did improve the signal-to-noise level at low frequencies, and so the pistol was usable in a wide frequency range.

The measured reverberation times in the studio are shown in Fig. 3. Compared with the limits discussed above it is seen, that only the values based on time reversed analysis are reliable at the lower frequencies (below about 500 Hz). There is a clear tendency that when the requirements for the forward analysis are not met, the results are longer than those from the time reversed analysis.
Fig. 3. Measured reverberation times in a 68 m³ studio using noise and impulse excitation. The results are compared with the lower limits for reliable results, (1) and (3) for forward analysis, (2) and (4) for time reversed analysis.

--- Using interrupted noise
--- Using integrated impulse response
--- Lower limit for reliable results

A. Forward analysis
B. Time reversed analysis

Conclusion

The measurement results from a studio have demonstrated that the ringing of 1/3 octave filters can introduce errors in the reverberation times evaluated from the decay curves. These difficulties can be overcome when a time reversed analysis is used. In addition this method allows the number of excitations at each position to be reduced considerably if the interrupted noise method is used.

It has been demonstrated that the interrupted noise method and the integrated impulse response method do agree very well and that both methods can benefit from the time reversed analysis when very short reverberation times have to be measured.

References

LATE DECAY TIME IN CONCERT HALL ACOUSTICS

Blago Guzina
Radio Beograd, Hilandarska 2, 11000 Beograd, Yugoslavia

1. Introduction

An investigation in order to obtain an insight into the relations existing between the general subjective judgments of the acoustic quality of three major concert halls in Beograd and the objective measurements of the so-called Late Decay Time was carried out.

The objective investigation included the measurements of Reverberation Time for a sound decay of first 10 dB (Early Decay Time), first 15 dB (Initial Reverberation Time), second 10 dB (named Late Decay Time) and first 30 dB. The corresponding denotations were $T_{10}$, $T_{15}$, $T_{II-10}$ and $T_{30}$ respectively.

All the measurements were performed in the same manner, using the same instrument, Microprocessor Audio Analyzer, IVIE Electronics, IE-30A/17A.

The measuring procedure was repeated ten times at each of six microphone positions in the octave frequency bands from 63 to 8000 Hz.

Concert halls were empty during the measurements. That was also the case with the Studio 6 of Radio Beograd, in which the same measurements were performed, in order to be compared with those of concert halls. Studio 6 is the greatest music recording premise in Radio Beograd.
2. Subjective Assessment of Acoustical Quality of Concert Halls

In order to subjectively judge the acoustical quality of concert halls in Beograd written questionnaires were distributed to 365 musically competent listeners. Among them 80 were professors and assistants of the Faculty of Musical Arts in Beograd, 20 students of the same faculty, 175 members of the Radio Television Beograd Chorus and Symphony Orchestra, Jazz Orchestra, recording engineers, producers and assistants of the Musical Production Department of Radio Beograd, 90 music critics, journalists and editors of Radio Beograd.

The participants of investigation were asked to judge the general musical acoustical quality of concert halls, based on their average total aural impressions obtained in a considerable long period. Rating numbers 1 - 5 were deliberately chosen concerning that the same way of rating is used in schools in Yugoslavia and people are familiar to it.

Musicians were asked to judge the concert halls strictly upon the impressions they had as the listeners in the auditorium, not as the performers on the stage.

The result of subjective acoustical assessment is shown in Table 1.

3. Relations Between the Results of Objective Measurements and Subjective Ratings

Measured values of the so-called Late Decay Time, $T_{10}$, were smaller than the values of Reverberation Time, evaluated for the first 30 dB of a sound decay, $T_{30}$, in the Kolarčev narodni univerzitet hall and Dom sindikata Jugoslavije. These two halls were subjectively judged better (mean values 4.7 and 2.4, respectively) than the Sava Centar hall. As shown in Table 2, the similar tendency $T_{10} < T_{30}$ was noticed in Studio 6, according to many opinions, the best in Radio Beograd.

In the Sava Centar hall, subjectively rated only 2.3, the Late Decay Time was 11% greater than Reverberation Time,
Also, by comparing the measured results of $T_{II-10}$ and $T_{10}$, it was noticed that only in Sava Centar hall, having the smallest rating, the Late Decay Time was greater than Early Decay Time, $T_{II-10} > T_{10}$, at almost all frequencies.

In other two halls and Studio 6, $T_{II-10}$ was smaller or equal to $T_{10}$.

The possible theoretical explanation of this is the fact that both in the Studio 6 and Kolarčev narodni univerzitet hall, whose geometric dimensions are manifold smaller than those of Sava centar hall, there were enough strong early reflections within the initial part of the reverberation process. The influence of the early reflections is obvious and the shape of reverberation curves is concave at the beginning of a slope.

In the halls with the subjectively good ratings, the values of $T_{10}$ were greater than $T_{30}$ in average 13% (Kolarčev narodni univerzitet hall) and 20.5% (Studio 6).

The values of $T_{15}$ were greater 1% in average than $T_{30}$, in Kolarčev narodni univerzitet hall, and 14% in Studio 6.

It was particularly noticed that the values of $T_{II-10}$ compared to those of $T_{10}$, $T_{15}$ and $T_{30}$ were in inverse relationship when measured in the halls subjectively judged good, compared to those judged bad. This very result was sufficient to allow the conclusion that the idea of measuring the Late Decay Time was justified.

4. Conclusion

In the course of objective investigation of acoustical properties of concert halls, besides the $T_{30}$ and $T_{10}$, it is worthwhile to measure the values of $T_{II-10}$. The results of $T_{II-10}$ should be compared to those of $T_{10}$.

In the halls, acoustically judged good, the values of the Late Decay Time are smaller than Early Decay Time.
# Table 1. Results of subjective investigation

<table>
<thead>
<tr>
<th>Hall / Studio</th>
<th>Number of participants</th>
<th>Average rating</th>
<th>Percentage of best marks (5)</th>
<th>Volume (m²)</th>
<th>T₃₀ (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNU</td>
<td>149</td>
<td>4.7</td>
<td>71</td>
<td>3900</td>
<td>1.6</td>
</tr>
<tr>
<td>DSJ</td>
<td>146</td>
<td>2.4</td>
<td>0</td>
<td>12000</td>
<td>1.3</td>
</tr>
<tr>
<td>SC</td>
<td>146</td>
<td>2.3</td>
<td>1</td>
<td>25000</td>
<td>1.5</td>
</tr>
<tr>
<td>SG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1550</td>
<td>0.5</td>
</tr>
</tbody>
</table>

KNU - Kolarčev narodni univerzitet Hall  
DSJ - Dom sindikata Jugosloviye Hall  
SC - Sava Centar Hall  
SG - Studio 6, Radio Beograd

\[
\frac{\Delta T_{11-10}}{T} = \frac{T_{11-10} - T_{30}}{T_{30}} \times 100\% 
\]

<table>
<thead>
<tr>
<th>f(Hz)</th>
<th>63</th>
<th>125</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
<th>8000</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNU</td>
<td>-41</td>
<td>-5</td>
<td>-20</td>
<td>0</td>
<td>-11</td>
<td>-4.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DSJ</td>
<td>-10.5</td>
<td>-7</td>
<td>-3</td>
<td>0</td>
<td>-8</td>
<td>-9</td>
<td>-9</td>
<td>-22</td>
</tr>
<tr>
<td>SC</td>
<td>-5</td>
<td>13</td>
<td>33</td>
<td>0</td>
<td>7</td>
<td>13</td>
<td>11</td>
<td>12.5</td>
</tr>
<tr>
<td>SG</td>
<td>-25</td>
<td>-14</td>
<td>-20</td>
<td>-20</td>
<td>0</td>
<td>-17</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Values of relative percentage difference between T₁₁₋₁₀ and T₃₀

KNU - Kolarčev narodni univerzitet Hall  
DSJ - Dom sindikata Jugosloviye Hall  
SC - Sava Centar Hall  
SG - Studio 6, Radio Beograd
ACOUSTIC ENERGY PROPAGATION IN ROOMS OF DIFFERENT FORM
Veska Topalova
Research Institute of Labor Safety and Ergonomics
Sofia, Bulgaria

The influence of the form of the room and the number of sources on the acoustic energy propagation and the reverberation time have been studied in production premises and in their models, scaled down at Μ = 1:15 for constant volume. No additional sound absorption has been introduced in the rooms/premises and in their models. Their form was: flat, cubic and elongated in accordance with the following requirements:

\[
\begin{align*}
L & > 5; & B & > 4 \text{ for flat rooms} \\
L & > 5; & B & < 4 \text{ for elongated rooms} \\
L & = 1; & B & \approx 1 \text{ for cubic rooms.}
\end{align*}
\]

Point sound sources with a circular frequency response were used.

I. In order to determine the influence of the form of the room on the acoustic energy distribution there, two types of investigations were carried out: first - with the source in the center of the room and second - with the source in one of the trilateral corners (Fig. 1 and Fig. 2).

The results obtained showed that:
- when the source is in the center of the room, there is uniform and substantial attenuation (4L=10 dB) along the three directions (the longitudinal and the vertical axes and the diagonals) only in the case of a flat room;
- when the source is in the corner, the shape of the room has no influence on the distribution of the acoustic energy;
- in elongated rooms the energy reflected from the longer walls and from the ceiling is added on to the incident acoustic energy with the diffusivity of the acoustic field (A 1<3dB) remaining the same for distances between the sound source and the reflecting surface (the wall) equal to or smaller than 5Lmax (Lmax is the maximum dimension of the source).

II. The influence of the number of sources on the acoustic pressure level at a given point has been studied in a flat room. The sound sources are of the same type, they emit the same level of acoustic power and are point sources. The increase in the acoustic pressure level with the increase of the number of sound sources is given on Fig. 3. The acoustic pressure level increases with 9 dB when the number of sound sources increases to 10. The acoustic pressure remains constant when the number of sources increases over 10. The point of investigation is in the center of the room and the sources are distributed uniformly around it and are at distances up to 10 m.

III. In statistical theory formulae, the room constant, assumed to be dependent on the room reverberation time and its volume influences the acoustic pressure level. The reverberation time has been studied for rooms of different form (Fig. 4).

According to Straszewicz (Poland) the reverberation time for the mean geometric frequency 1000 Hz depends on the room height, the sound absorption factor of the ceiling of the room and the factor Kt - depending on the height of the hardware.

\[
T = 0.15H - 1.8K_{ce} + 4.8K_t
\]

Since this author's investigations were carried out in empty rooms, without the introduction of additional sound absorption (Kt<<1, Kt=0, T=H)
and the height of the flat room is equal to the height of the elongated room, the form of the room, therefore influences substantially the reverberation time, which necessitates a correction in the statistical theory formular.
3. B
RELATION BETWEEN ABSORPTION COEFFICIENT $\alpha_{\text{stat}}$ DERIVED FROM $\alpha(\theta)$ AND $\alpha_n$ MEASURED IN THE REVERBERATION ROOM.

J. Kaku and J. Yoshimura

Kobayashi Institute of Physical Research, 3-20-41 Higashimotomachi, Kokubunji, Tokyo 185, Japan

1. Introduction
The oblique-incidence sound absorption coefficients $\alpha(\theta)$ of four sorts of porous materials were measured. They were compared with those calculated from the respective specific normal acoustic impedances by use of the Paris’ formula. On the other hand, the reverberant absorption coefficients $\alpha_n$ of the same materials were also measured in a reverberation room, the excess absorption due to the edge-effect being excluded.

The $\alpha_n$ was compared with the random incidence sound absorption coefficient $\alpha_{\text{stat}}$ obtained from the measured and calculated $\alpha(\theta)$ being averaged over the incident angle $\theta$ to $\pi/2$ respectively.

2. Materials
Acoustic characteristics of the test specimens are listed in Table 1. Polyurethane foam was selected as a material of isotropy and glassfiber as a material of anisotropy.

3. $\alpha(\theta)$

3.1 $\alpha(\theta)_M$ measured
$\alpha(\theta)_M$ was measured by the cancellation method1). A sheet of the material under test (size:4x3 m$^2$) was arranged on a vertical wall with solid backing in a dead room (volume:1600 m$^3$). The sound source was set at a distance of 2.5 meters from the reflection point on the test material so that the errors due to deviations from the plane wave incidence can be neglected in the frequency range above 500 Hz. The accurate measurement of $\alpha(\theta)_M$ at an incident angle more than 75° is difficult owing to the errors by diffraction at the edge of the test specimen.

3.2 $\alpha(\theta)_C$ calculated
The specific normal acoustic impedance $z$ of every sort of materials was measured in an impedance tube by a single microphone method with a signal of periodic pseudorandom sequence2). When the behavior of the material is assumed as extended reaction, $z$ was calculated from the

<table>
<thead>
<tr>
<th>Table 1. Polyurethane foam and glassfiber sheet for test.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index of Thickness</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>specimen</td>
</tr>
<tr>
<td>A$^{25}$</td>
</tr>
<tr>
<td>A$^{50}$</td>
</tr>
<tr>
<td>B$^{25}$</td>
</tr>
<tr>
<td>C$^{50}$</td>
</tr>
</tbody>
</table>

149
characteristic acoustic impedance \( W \) and the propagation constant \( \gamma \) which were determined by the two thickness method. \( \alpha(\theta) \) was calculated from the following Paris' formula,

\[
\alpha(\theta) = \frac{\cos \theta - 1}{\cos \theta + 1} + \frac{4 \text{Re}(c) \cos \theta}{|c|^2 \cos^2 \theta + 2 \text{Re}(c) \cos \theta + 1}.
\]

Figure 1 shows \( \alpha(\theta) \) together with \( \alpha(\theta) \).

4. \( \alpha_{\text{stat}} \)

Two kinds of \( \alpha_{\text{stat}} \) are given by averaging \( \alpha(\theta) \), which are obtained in 3.1 or 3.2, over the incident angle \( 0 \) to \( \pi/2 \) as follows,

\[
\alpha_{\text{stat}} = 2 \int_0^{\pi/2} \alpha(\theta) \sin \theta \cos \theta \, d\theta.
\]

4.1 \( \alpha_{\text{stat}} \) from \( \alpha(\theta) \)

\( \alpha_{\text{stat}} \) which was calculated from \( \alpha(\theta) \) is denoted by \( \alpha_{\text{stat}} \).

4.2 \( \alpha_{\text{stat}} \) from \( \alpha(\theta) \)

\( \alpha_{\text{stat}} \) is differentiated into \( \alpha_{\text{stat}} \) and \( \alpha_{E} \) corresponding to the cases when the material is assumed to be locally reaction or extended reaction.

Fig.1. Comparison between measured \( \alpha(\theta) \) (\( \bigcirc \), \( \bigtriangleup \), \( \square \), \( \bigstar \)) and calculated \( \alpha(\theta) \) (—— locally, —— extended).
5. \( \alpha_\infty \)

A reverberation room of Kobayasi Institute of Physical Research (volume: 513 m\(^3\), surface area: 382 m\(^2\), with suspended diffuse) was used for measurements of sound absorption coefficient \( \alpha_0 \). The decay curves were observed by using the reverberation-meter (YAMAHA AS-1)* in which Schroeder's impulse integrating method was applied.\(^2\)

Figure 2 shows the relation between \( \alpha_0 \) and the specific length \( E (= L/S) \), where \( L \) is the perimeter of the test specimen and \( S \) is the total area. \( \alpha_0^\infty \) for an infinitely large sample corresponds to the value of \( \alpha_0 \) at \( E=0 \) by extrapolation of the regression line. The symbols with asterisk at \( E=0 \) in Fig.2(a) mean the values which were measured by covering the whole floor of the reverberation room with test material.

Figure 3 shows \( \alpha_0 \) and \( \alpha_{\text{stat}} \) of four specimens.

**Fig.2.** The reverberant absorption coefficient \( \alpha_0^\infty \) without edge-effect.

6. Discussion

The \( \alpha_0^\infty \) of four specimens may be regarded as reliable from the linearity of decay curves. \( \alpha_{\text{stat}} \) obtained from the measured \( \alpha(0)_m \) agrees well with \( \alpha_0^\infty \), accordingly the \( \alpha(\theta)_m \) measured at an incident angle of 45° to 60° may also be regarded as reliable one.

On the other hand, as is shown in Fig.1, the disagreement between \( \alpha(\theta)_m \) and calculated \( \alpha(\theta)_c \) is remarkable with increasing incident angle for every specimen.

This discrepancy may be ascribed to the invalidity of the Paris' formula, in particular at large incident angle.

* The authors are deeply indebted to Mr. Kawakami for the use of his instrument.
Fig. 3. Comparison between $\alpha_{a}^{\omega}(\bullet), \alpha_{stat}^{M}(\circ), \alpha_{stat}^{L}(\triangle)$ and $\alpha_{stat}^{E}(\square)$.

Acknowledgements:

The authors wish to express their gratitude towards Dr. J. Igarashi and Dr. M. Yamashita for their support to this study.

References


EDGE EFFECT INFLUENCE ON THE STATISTICAL ABSORPTION COEFFICIENT

A. Cops, W. Lauriks

Laboratorium voor Akoestiek en Warmtegeleiding, Departement Natuurkunde, Katholieke Universiteit Leuven, Celestijnenlaan 200 D, 3030 Leuven, Belgium

1. Introduction

From impedance measurements with two-microphone techniques and sound pressure measurements above a clamped foam material, in free field conditions, the statistical absorption coefficients have been calculated and compared with experimental data, calculated with Eyring's formula, for different sizes of foam samples, by taking the diffraction of sound waves at the edges into account. This numerical method may be of use to compare the results, obtained in different reverberation rooms with one another, in order to explain the differences obtained.

2. Theory

2.1. An infinite sample

The statistical absorption coefficient for an infinite sample is given by the well known Paris formula:

\[ \alpha_\infty = \frac{4}{\pi} \int_0^{\pi} \frac{Re(Z_0) \sin(\theta) d\theta d\phi}{\int_0^{\pi} |Z_0 \cos(\theta)|^2} \]  

(1)

where Z and Re(Z) are the normalized impedance respectively the real part of the normalized impedance of the material. θ and ϕ are given in figure 1.

2.2. Finite samples

We used a variational method described in (1) and developed by Thomasson (9) to calculate the pressure distribution on a rectangular patch with dimensions a and b. Let

\[ p_1(r) = A e^{ikr - i\omega t} \]  

(2)

be the incident sound pressure. In this case the pressure just above the surface is given by:

\[ p(x,y,0) = 2p_1(x,y,0) \pm \frac{ik}{Z} \int_{-a}^{b} \int_{-b}^{b} G(x_0, y_0, 0) dx_0 dy_0 \]  

(3)

Where G is Green's function:

\[ G = \frac{e^{-ikR}}{2\pi R} \]

and

\[ R = \sqrt{(x - x_0)^2 + (y - y_0)^2} \]

The coordinate axes are defined in fig. 1. An exact solution of the pressure distribution in equation (3) is not possible and we can only obtain approximate results. Since there is a direct relation between the distribution-in-space of the scattered wave and the absorbed power (ref. 1) we can
obtain a second order solution for the absorbed power by inserting a trial function $p = A \pi \lambda_1$ into equation (3) and by adjusting the value of $\lambda$ so that the first variation of the far field amplitude is zero. We obtain:

$$p = \frac{Z}{Z_f + Z} \pi \lambda_1 (x, y, 0)$$  \hspace{1cm} (4)

where

$$Z_f = \frac{ik}{\psi} \left( \int_{-a}^{a} \int_{-b}^{b} \pi \lambda_1 (x, y, 0) dx dy \right) p(x, y, 0) dx dy$$  \hspace{1cm} (5)

$Z_f$ depends on the shape and size of the absorber, on the angle of incidence of the wave and on the frequency. For an infinitely large absorber $Z_f = 1/\cos\theta$.

The absorbed power in a diffuse field is given by:

$$W_a = \frac{1}{2} \frac{\text{Re}(Z)}{\rho c_0} \int_{-a}^{a} \int_{-b}^{b} |p|^2 dx dy$$  \hspace{1cm} (6)

where $\rho$ is the density and $c_0$ the sound speed in air. With equation (4) this becomes:

$$W_a = \frac{2SA^2 \text{Re}(Z)}{\rho c_0 (Z + Z_f)^2}$$  \hspace{1cm} (7)

with $A$ is the amplitude of the incident wave. The incident power on the sample in a diffuse sound field can be taken from:

$$W_1 = \frac{SA^2}{2 \rho c_0}$$  \hspace{1cm} (8)

Hence, the statistical absorption coefficient is equal to

$$\alpha_{stat} = \frac{2}{\pi} \int_{0}^{\pi} \int_{0}^{2\pi} W_1 d\phi d\theta = \frac{4}{\pi} \int_{0}^{\pi} \int_{0}^{2\pi} \frac{\text{Re}(Z) \sin\theta d\phi d\theta}{|Z + Z_f|^2}$$  \hspace{1cm} (9)

3. Measurements

3.1. Pressure distribution

The material used was a 5 cm thick polyurethane foam with a flow resistivity $\xi$ of about 5 x 10^3 Ns/m^4. The impedance of the material was measured on a sample of dimensions 1.5m x 2.0m as a function of angle of incidence in a free field with a method described in ref. 3 and 4. The results of these measurements are reported in ref. 4. The pressure distribution was measured on a smaller sample of dimensions a=b=0.5m, build in a hard plate (Figure 1). The angle $\theta$ was kept constant and set equal to $0^\circ$. Different angles of incidence $\theta$ were used.

The figures 2a and 2b show the calculated pressure distribution at the sample surface, for the frequency 2000 Hz and $\theta=0^\circ$, for $\theta = 0^\circ$ and $\theta = 45^\circ$, together with the measuring results. It can be concluded that the calculated sound pressures are accurate enough to predict the diffraction effects at the edges of the sample. Measurements and calculations done at other
frequencies and different $Y$ values give the same promising results. It is clear that the pressure at the sample surface is considerably influenced by the diffraction effects at the edges.

3.2. Statistical absorption coefficient

In the calculation of the statistical absorption coefficient from equation 9 the characteristic impedance $Z_C$ and the propagation constant $k_0$ were calculated with the Delany and Bazley formulae (3). The angle dependence of the impedance is given by:

$$Z = \frac{Z_C}{\cos \theta} \coth(-ik_0 \cos \theta)$$

where $\theta$ is the angle of refraction, given by Snell's law. The statistical absorption coefficient was calculated for different sample sizes and compared with measured results in the reverberation chamber obtained with the Eyring formula. The results obtained for one sample size are presented in figure 3. With other sample sizes the same good agreement occurs.

4. Conclusions

It is shown that the variational technique is able to predict the diffraction effects at the edges of a finite sample of finite impedance placed on a hard floor. This technique is a useful tool for the interpretation of the statistical absorption coefficient in the reverberation room.

References

1. P.M. Morse, K.U. Ingard 'Theoretical Acoustics'

2. S.I. Thomasson 'On the absorption coefficient'

3. J.F. Allard, Y. Champoux, J. Nicolas 'Pressure variation above a layer of absorbing material and impedance measurements at oblique incidence and low frequencies'.
   To be published

4. W. Lauriks, A. Cops, J.F. Allard 'Acoustical characteristics of foam materials'

5. M.E. Delany, E.N. Bazley 'Acoustical characteristics of fibrous absorbent materials NPL AERO'
   Report Ac37, March 1969.
Fig. 1. Definition of the measuring set-up.

Fig. 2. Sound pressure distribution above the sample surface at a frequency of 2000 Hz. Y = 0.0 m.

(a) $\Theta = 0^\circ$
(b) $\Theta = 45^\circ$

Fig. 3. Statistical absorption coefficient of a 5 cm thick foam material with sample area $S = (2.18 \times 1.22) m^2$

- $x-x$: measurement and calculated with Eyring's formula
- $o-o$: calculated with equation 9.
ON THE SOUND REFLECTION OF THE AUDITORIUM SEATS

Koji Ishida, Kiyoshi Sugino, and Isao Masuda

Tecnological Research Center, Tokyo Construction Co., Ltd.
2-13-9, Miyazaki, Miyamae-ku, Kawasaki-shi, Japan

In an auditorium, the first sound to reach the audience comes directly from the stage. The second sound heard is that which is reflected at the rows of seats. The direct sound and the reflected sound interfere with each other. It is well known that the sound waves propagating over rows of seats are subject to selective attenuation at low frequencies,1 but the characteristics of this interference are not well understood quantitatively. It is furthermore interesting that it has been reported that the surface reflection of audience seats at the grazing angle is negative.2-3 The paper describes the behavior of the reflection coefficient in the time and frequency domains using 1/10 scale model seat rows. Reflection caused by the existence of parallel barriers which model rows of seats between a sound source and a receiving point is also discussed in this paper. Estimates which were obtained by a formula governed by Kirchoff's boundary conditions4 are compared in the time domain with measured values. It was found that the estimates corresponded closely to measured values.

1. Measurement of reflection coefficient for 1/10 scale model seat rows

Reflection coefficients in the time and frequency domains for 10 rows of 12 seats in Fig.1 are measured by impulsive (13/μsec. recutangular wave) spherical wave incidence. Spherical waves were generated with a tweeter. Direct and reflected waves were received by a 1/16 in. condenser microphone. It was difficult to separate the direct wave and the reflected waves, so the direct wave was measured without setting the test model. The method of obtaining the impulse response and transfer function is shown in Fig.2. It is as follows. Direct wave (a), and the combination of direct wave and waves reflected from the rows of seats (b) were transformed into frequency functions by Fourier transform (A). The transfer function was obtained by their ratio (B). Impulse response (c) is defined in this paper by the inverse Fourier transform of the filtered transfer function. The transfer function of the filter, and its inverse Fourier transform are shown in Figs.2(d),(C), respectively. The amplitude of impulse responses was normalized by the amplitude of the direct sound from a point source. Sound pressure level on the ordinate of a transfer function was also referred to that of the direct sound. The scale model seats were made of polystyrene and grass wool. The scale models of the human body were made of polystyrene covered with felt, as shown in Fig.3. The absorption coefficients of the scale model when the seats were empty and when they were occupied are shown in Fig.4. The frequency and the time in all figures are multiplied by factor 0.1 and 10, respectively.
1.1 Comparison of the reflection coefficient by different vertical incident angle when the seats are empty and when they were occupied

Reflection coefficients for rows of seats at different vertical incident angles were measured at the receiving point (P0) when seats were empty, and when they were occupied. These coefficients are shown in Fig.5. In the time domain, when seats are empty (a) - (c), they had negative reflections immediately after direct wave was received. The closer incident angle \( \theta \) was to the grazing angle, the larger the negative wave, and the shorter the time difference between the direct and the negative wave. In the frequency domain, this negative reflection could decrease the direct sound because they canceled each other out under the frequency which corresponded to reciprocal number of the time period between a direct wave and negative wave. Conversely, they emphasized each other above that frequency. The transfer functions when the vertical angles of incidence were 87° and 78° are shown in Fig.5(a). The phenomenon of reversal occurred around 1.6kHz. After negative reflection, the low amplitude positive waves reached the receiving point 2ms to 3ms after the direct wave. The incident angle \( \theta \) was closer to 90°, the duration of the positive reflection was longer, amplitude was lower, and time delay after direct wave was shorter. In the frequency domain, attenuation was observed in the frequency range of 100-250Hz [Fig.5(b)]. Attenuation was reduced by decreasing the incident angle \( \theta \), and the dip frequency shifted toward the lower end at the same time. The low amplitude negative waves reached the receiving point 8ms to 14ms after the direct wave. The incident angle \( \theta \) was closer to 90°, the duration of the negative reflections was shorter, and the amplitude was lower. The highest peack of the transfer functions was observed at around 50Hz [Fig.5(b)], and it was decreased by increasing the incident angle \( \theta \).

Concerning the reflection coefficient in the time domain when the seats were occupied, the tendency of the positive waves following negative wave was almost the same as it was when the seats were empty (Fig.6). However, the time difference between direct wave and negative wave was shorter than the case with empty seats. This was because the distance travelled was shorter due to reflection of the upper part of the human models. The transfer functions obtained when the seats were empty and when they were occupied are compared in Fig.7. A sound pressure level of 500Hz or greater was larger by about 5dB for unoccupied seats than it was for occupied seat. When the seats were occupied, the direct sound and the negative reflection canceled each other out, and the sound pressure level decreased over a broad frequency range.
1.2 Comparison of the reflection coefficient by different horizontal incident angles when the seats were empty

The reflection coefficients for rows of the seats at different horizontal incident angles when the point source fixed at S1 was measured at receiving points, P4, P5 and P6 are compared in Fig.8. In the frequency domain, the attenuation was observed in the frequency range of 100~2500Hz [Fig.8(A)]. Attenuation was reduced by increasing the incident angle φ. At the same time, the dip frequency shifted toward the lower end.

1.3 Effects of floor absorption

Through theoretical calculation of the condition of the incident plane wave, it was found that sound-absorbent floor structures reduced attenuation at low frequency. The reflection coefficients for 10 rows of 12 seats on a floor covered with rigid, perforated panels (Fig.9) were compared with those of the same rows of seats mounted on a rigid floor (Fig.11), when the seats were empty (A) and when they were occupied (B). Fig.10 shows absorption coefficient of the perforated panel (A=0.6m, 42.5mm, 30mm air space). When the seats, both empty and occupied, were placed on a floor covered with rigid, perforated strip panels, the negative reflections reached the receiving point 8ms to 14ms after the direct wave disappeared. The positive reflections prior to negative one were roughly the same. The result of this measurement is estimated as in the following. Sound absorption by a perforated panel floor affects the latter negative reflections. Judging from the delayed time, it was assumed that the waves reached receiving point 8ms to 14ms delayed from direct wave were occurred due to multiple reflections between the floor and the lower part of the seats. In this case, it was estimated that the vertical incident angle of each reflection upon the floor surface was close to 0°, that is, the absorbent floor caused these reflections to disappeared.

On the other hand, the positive reflections weren’t absorbed by the surface of the floor because the vertical incident angle of each reflection between the floor and seats was closer to the grazing angle and there was little reflection between them. In the frequency domain, when using perforated, rigid strip floors, the sound pressure level up to 1000Hz was greatly reduced. When the seats were empty, the attenuation around 120Hz was larger compared with that resulting from the use of rigid floor.

1.4 Effects of seat underpass

The reflection coefficients for 3 rows of 5 seats (Fig.12) with underpass blocked and with underpass open were measured as shown in Fig.13. The seats were mounted on a rigid floor. When the underpass was open, the reflection coefficient in the time domain had positive reflections 2ms to 8ms after the
direct wave was received. When the underpass was closed, these reflections didn’t appear until 3.5s after the direct wave was received, and the latter negative reflections did not appear. In the frequency domain when the underpass was blocked, the dip frequency shifted toward the lower end and at the same time the sound pressure level of around 50Hz decreased.

2. Reflection of parallel barriers between a sound source and a receiving point which model seat rows.

In the previous section, it was shown that negative reflections appeared the rows of seats immediately after a direct wave. It was known that the reflection from the free edge of the panel appeared on the negative side. Seat rows can be regarded as a succession of parallel barriers (Fig.14). The values calculated under Kirchhoff’s boundary condition (Fig.15) in the time domain were compared with the measured values shown in Fig.16. Measured and calculated values show good agreement.

It is interesting that the second reflected boundary wave appeared on the positive side, and it is suggested that, in the case of multiple parallel barriers, the total reflections from the edge of barrier did not show unlimited increase toward negative side because of the positive waves. It is thought that the same phenomenon occurred in the case of the reflections of the rows of seats.

3. Conclusion

It is explained that the sound reflections of the rows of seats can be classified into three distinctive waves throughout the various measurements. Firstly, the acute negative wave reached the receiving point immediately after direct sound. This wave is considered to be the reflection of the upper edge of the seat back and the upper part of the human body. Secondly, the positive, low amplitude reflections reached the receiving point 2ms to 5ms after the direct wave. These waves are considered to be the early reflections between the lower part of the seats and the floor. Thirdly, the negative, low amplitude reflections reached the receiving point 5ms to 14ms after the direct wave. These waves are considered to reflect repeatedly at an obtuse vertical angle of wave incidence.

In the frequency domain, it is proved that the first negative reflection caused decreasing sound pressure level at the broad frequency range, the positive reflections caused attenuation in the frequency range of 100~250Hz, the later negative reflections caused increasing sound pressure level around 50Hz.

Acknowledgements

The authors wish to express their appreciation to Dr. Y. Sakurai, Assistant Professor at Kansei University, for his continuous encouragement and suggestion.

References


160
PREFERENCE ANALYSIS DERIVED FROM PERCEPTIVE TESTS IN ROOM ACOUSTICS

C. Lavantier (*), O. Warusfel (**), J.P. Jullien (***)

(*) IRCAM, 31 rue Saint Merri, 75004 Paris, FRANCE
(***) CNET, route de Trégastel, 22301 Lannion, FRANCE

INTRODUCTION

The topic of that paper is a synthesis of the main results given by perceptive tests. They have been made in the laboratory of IRCAM, in order to validate a set of objective criteria used to characterize a room acoustic quality. Seventeen tests have been proceeded, with an average of 14 subjects and 8 stimuli per test. They have been analysed with INDSCAL multidimensional scaling. This method is presented in an associated paper [Julien & col.]. Each test needed three or four factors to be explained so 58 factors have been extracted from the all study. The strong factor concept allows to find the same perceptive factor among several tests. Finally, a set of 13 factors could be selected. To be understandable by a large public and not only by acousticians, these factors have been named by descriptive terms. These words have been proposed by a group of selected persons after a large number of tests in order that they could focus on each perceptive factor. Preferences have not been enough analysed to allow us to give general results. Some punctually interesting elements can only be presented here.

PERCEPTIVE FACTORS DERIVED FROM SUBJECTIVE TESTS

Temporal criteria

Four basic criteria that deal with the temporal distribution of the energy were selected to structure the tests stimuli: LEV, Dir, C80 and RT. The feeling due to a variation of total energy needs two factors to be explained. When the LEV value is weak, the auditors feel outside the room until the direct sound reaches a certain threshold. Then they perceive the rising of sound strength. The first factor could be named remote present, the second one which is generally high correlated with LEV criterion could be named weak strong. These names come from a translation of French adjectives (distant présent) and (faible énergique).

Variations of LEV criterion are perceptively independant of RT and C80 variations. Variations of RT are well discriminated on a specific factor but C80 variations are completely substituted on that factor when C80 values are very weak. The RT correlated factor was named by several couple of terms. Dry reverberant is a possible description like small room large room (petite salle grande salle). Progressively, variations of C80 are perceived on a peculiar factor named confused distinct (branillé net). The substitution is still 50% on the RT factor for a C80 variation from -3 dB to 0 dB. It disappears when C80 value is large enough. Then, a C80 variation is no
more perceived as a clarity effect but as a feeling of proximity. A substitution appears on a factor named faraway nearby (lointain proche), correlated to a new criterion ZDIR where the frontal early energy is integrated to the strictly direct sound Dir. The figure 1 represents the perception of C80 criterion's variation.

![Figure 1](image1.png)

**Figure 1** Perception of C80 variation from very low value to very high value.

![Figure 2](image2.png)

**Figure 2** Individual vector preferences for speech message in conference rooms. There is here a consensus on first factor.

Considering the perceptive distance between two configurations where only one criterion varies, it is possible to quantify the relationship between an objective criterion's variation and a perceptive measure. The following table 1 presents this relation for the LEV, RT, C80 and ZDIR variations. Moreover, if we suppose the relative difference limen for the reverberation time to be approximatively 10% and the just noticeable difference in level and clarity to be about 1 dB, the correspondance in the perceptive field is for all thresholds roughly equal to 80 units.

<table>
<thead>
<tr>
<th>Factor names</th>
<th>Obj. crit</th>
<th>Obj. delta</th>
<th>Subjective delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>remote present</td>
<td>LEV</td>
<td>1 dB</td>
<td>from 75 to 95 units</td>
</tr>
<tr>
<td>weak strong</td>
<td>LEV</td>
<td>1 dB</td>
<td></td>
</tr>
<tr>
<td>dry reverberant</td>
<td>C80</td>
<td>1 dB</td>
<td>35 units</td>
</tr>
<tr>
<td>confused distinct</td>
<td>C80</td>
<td>1 dB</td>
<td>80 units</td>
</tr>
<tr>
<td>faraway nearby</td>
<td>C80</td>
<td>1 dB</td>
<td>35 units</td>
</tr>
<tr>
<td>faraway nearby</td>
<td>ZDIR</td>
<td>1 dB</td>
<td>85 units</td>
</tr>
<tr>
<td>dry reverberant</td>
<td>Ln RT</td>
<td>1</td>
<td>from 60 to 75 units</td>
</tr>
</tbody>
</table>

**Table 1** Perceptive quantification and description for temporal criteria variations.
Two models try to determine the preference behaviour for a set of stimuli. Since individual preferences vary widely, it is important to perform preference analyses for each subject. The vector model characterizes the following assumption: the more, the better. The ideal point model characterizes this one: some amount is ideal. Until now, we have worked with the first model to extract a possible consensus over the population. Figure 2 presents individual preference vectors for a speech message listened in conference rooms. The preferred rooms are the clearest and the less reverberant ones except for one subject. It may be said here there is a consensus on the first factor. The same consensus is not found when music is listened. The ideal point model seems then to be better adapted. We don’t have yet analysed the preference data with that model neither the effect of music style. It has been observed that, for wind chamber music, when clarity is higher than 4 dB, a reverberation time superior to 1.7 s is appreciated. When clarity is not sufficient, a low reverberation time improves preference. When examples are listened through headphones, the total preferred level seems to be lower than the one demanded when messages are listened through a reproducing system in an anechoic room.

Spatial criterion

Spatial distribution of first reflections have been presented in an associated paper [Waurasfel&col.]. Two factors appear from four spatial tests. The first one is best correlated with the early lateral energy fraction LE. It represents the feeling of being inside the music. That spatial impression might be bothered by the fact that lateral energy comes from one side. The dependency between spatial impression and listening level is not obvious. The second spatial factor is sensitive to the spread of the source. Therefore, a new criterion ANG has been calculated and correlated to that factor.

There is no consensus on the LE correlated factor. Some people prefer to locate the sound in front of them but a spread source seems to be preferred to a punctual source even if the music is performed by only one instrument. LE criterion’s variations are equivalent in term of quantification to the C80’s ones. An increase of 1 dB is evaluated to 80 perceptive units. The discrimination threshold estimated to 80 perceptive units corresponds to an ANG variation of 8°.

Coloration criteria

Three new criteria were introduce to explain three perceptive factors. The fraction between the direct sound and reflections Dir/REF seems to translate in the objective field the fact that if REF is much more important than Dir, the room’s coloration lacks of contrast. The second new criteria is the early central time of the relevant reflections ECT. And the last one is the standard deviation in time of these reflections SDF. Table 2 shows that these coloration effects are secondary regarding to their perceptive quantification.
The expected factors due to the RT and LEV values in octave bands were effectively found. The perception of RT variation for treble octave bands is two and half times more important than the perception for basses variations. A room with an excessive RT value in treble frequencies is rejected by preferences and is perceived as too sharp (acide). The perception of LEV is more balanced, but it is an excessive value in basses which is rejected. The correspondent room quality is considered as too heavy (lourd). Quantification is presented in table 3.

<table>
<thead>
<tr>
<th>factor names</th>
<th>obj. crit</th>
<th>sum of obj. delta</th>
<th>subj. delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>neutral contrasted</td>
<td>RT 125 Hz 250 Hz</td>
<td>1.0</td>
<td>120 units</td>
</tr>
<tr>
<td>dry live</td>
<td>RT 4 kHz 8 kHz</td>
<td>1.0</td>
<td>300 units</td>
</tr>
<tr>
<td>hollow warmth</td>
<td>LEV 125 Hz 250 Hz</td>
<td>1 dB</td>
<td>40 units</td>
</tr>
<tr>
<td>poor brilliant</td>
<td>LEV 4 kHz 8 kHz</td>
<td>1 dB</td>
<td>40 units</td>
</tr>
</tbody>
</table>

CONCLUSION

All the presented results establish a correspondance between perceptive aspects described by descriptive terms, and an objective caracterisation in terms of criteria values. This could be integrated to a software whose one of the applications could be to simplify the control of variable either mechanical or electronical acoustic system. The preference analyses must be pursued in order to influence choices among several possible solutions.

REFERENCES

[Julien&col.] J.P. Julien, C. Lavandier, O. Warusfel; "Compered analysis of the clarity and the reverberation in room acoustics"; ICA 1989
LA CARACTERISATION PERCEPTIVE DES SALLES DE CONCERT : INSUFFISANCE DES CRITERES TRADITIONNELS ET PROPOSITION D'UN MODELE BASE SUR LA DISTRIBUTION DES PREMIERES REFLEXIONS

A. Fischetti†, J. Jouhanneau++, Y. Hemim++

†Conservatoire National des Arts et Métiers
Laboratoire d'Acoustique
292, rue Saint-Martin, 75141 Paris Cedex 03, France

++Laboratoire de Statistique de l'Université de Paris VI
1, place Jussieu, 75005 Paris

PROCEDURE EXPERIMENTALE

L'espace de Projection de l'ICAM est une salle parallélépipédique dont les parois intérieures sont constituées de triédrès orientables qui comportent trois faces : absorbante, réfléchissante, et diffusante. De plus, la hauteur de plafond est variable, de 0 à 10 m.

Une étude de la salle a été conduite sur la base de 5 paramètres géométriques :
H : hauteur de plafond. Trois valeurs ont été sélectionnées : 5 m, 7,5 m, et 10 m.
A / D : rapport du nombre de panneaux réfléchissants au nombre de panneaux diffusants. (trois valeurs 30 %, 60 %, 90 %).
R : pourcentage de panneaux réfléchissants (trois valeurs : 0 %, 15 %, 30 %)
G : gradient de répartition des panneaux absorbants (panneaux groupés, faiblement répartis, uniformément répartis)

P : position des microphones dans la salle. Trois valeurs sont sélectionnées :
P1 : à 5 m, au centre de la salle, P2 : à 10 m, près d'un mur, et P3 : à 15 m, au centre de la salle.

Une source musicale a été enregistrée en chambre source, puis rediffusée dans l'Espace de Projection par l'intermédiaire de deux enceintes. On a effectué ensuite un enregistrement en utilisant une tête artificielle pour chacune des trois positions dans la salle. Cette procédure a été répétée pour différentes combinaisons des paramètres précités, et on mesure à chaque fois, la réponse impulsionnelle ainsi que 15 des critères acoustiques les plus utilisés.

ÉTUDE DE LA SPATIALISATION

Protocole de test

Le test d'écoute est la seule méthode adaptée à l'étude des salles, mais les déformations apportées par le système d'écoute sont mal définies.

Les enregistrements étant effectués par tête artificielle, l'écoute devrait être faite sur casque, mais le choix présenté deux inconvénients : l'allègement des caractéristiques spatiales, et la substitution d'une perception latéralisée à une écoute localisée (la plupart des enregistrements sonores sont écoutes sur enceintes).

Comme la spatialisation est le paramètre perceptif le plus perturbé par le système d'écoute, une étude préliminaire a été conduite.

31 échantillons sonores ont été proposés à 10 sujets chargés de les classer sur une échelle subjective relative à 5 paramètres : dimension apparente de la salle, distance apparente par rapport à la source, localisation latérale de la source, répartition, ou plénitude spatiale (sensation d'être enveloppé par le son), temps de réverbération subjectif, qualité sonore. Le test est effectué sur casque et sur enceintes.

Résultats

Les réponses, traitées par analyse des correspondances multiples, permettent de mettre en évidence les résultats suivants :

Differences casque / enceintes

Les résultats montrent que la distance apparente joue un rôle plus important dans l'écoute sur enceintes que sur casque, tandis que l'inverse se produit pour la distribution spatiale.
Par ailleurs, une analyse plus détaillée a mis en évidence certaines variations dans les relations entre critères acoustiques et perceptifs, selon le type d'écoutée.

Les différences entre l'écoute sur casque et sur enceintes sont nécessairement dues à une perception différente des premières réflexions. En effet, les différences concernent particulièrement la perception de la distance apparente, autrement dit du rapport entre l'énergie directe et l'énergie réverbérée. Or, ce sont les propriétés de cohérence spatiale des signaux qui déterminent le caractère direct ou réverbéré du son. Le caractère inhérent du champ diffus régnant dans la salle est respecté dans l'écoute sur casque, par le fait que les signaux parvenant aux deux oreilles de l'auditeur sont identiques à ceux captés par les deux micros de la tête artificielle. Ce qui n'est pas le cas dans l'écoute sur enceintes où la présence de deux sources impose un doublement perceptif (oreille gauche - oreille droite pour le haut-parleur de gauche, et oreille droite - oreille gauche pour le haut-parleur de droite) des deux composantes du front d'onde (dont le déphasage croît avec l'angle d'incidence).

Relation entre paramètres perceptifs et critères acoustiques

Après analyse, on trouve que:
- l'énergie comprise dans les 40 premières ms de la réponse impulsionnelle intervient dans la sensation de temps de réverbération subjectif, et dans celle de distance apparente;
- la plénitude spatiale est corrélée à l'amplitude et la densité des premières réflexions;
- pour un temps de réverbération moyen, la qualité subjective du son croît avec le niveau des premières réflexions, et le phénomène est d'autant plus net que ces réflexions sont, jusqu'à un certain point, étroites et espacées.

Conclusions

L'importance du rôle joué par les premières réflexions de la réponse impulsionnelle apparaît clairement dans chacun des deux aspects de l'expérience:
- elles semblent être la clé des différences de perception casque / enceintes;
- elles modifient de façon sensible tous les critères perceptifs étudiés.

De plus, pour expliquer certains résultats, comme ceux concernant la plénitude spatiale et la qualité globale, il devient indispensable de prendre en compte la distribution des premières réflexions. Tous les résultats confirment le fait que le modèle de représentation de la réponse impulsionnelle à partir des critères habituels ne permet pas d'expliquer toutes les différences perceptives. Certaines caractéristiques, visibles sur le relevé des réponses impulsionnelles, et parfaitement perçues par la plupart des sujets, ne sont pas traduites par les critères classiques. Parmi ces caractéristiques, la largeur, l'emplacement, et le regroupement de certaines réflexions semble jouer un rôle particulièrement prépondérant.

ÉTUDE DE LA CLARTE

Afin de mieux cerner l'incidence de la distribution des premières réflexions sur la perception, une étude plus spécifique du rôle des paramètres géométriques et acoustiques de la salle a été conduite.

Les échantillons sonores ont été regroupés par paires et proposés à 10 sujets chargés de les discriminer et de les classer préférentiellement selon le critère de clarté. Les résultats sont comparés à la valeur objective de Reichardt exprimée sous la forme de C80.

Le test est effectué sur des enregistrements de batterie.

Les résultats montrent nettement que le critère C80 est très insuffisant pour décrire les différences effectivement perçues. Ceci peut être interprété comme une absence d'intégration linéaire des processus perceptifs. Le résultat le plus démonstratif est à cet égard l'augmentation de la sensation de clarté pour des sons présentant de grands écarts entre certaines des premières réflexions, alors que ces écarts ne modifient en rien la valeur du C80.

La dernière partie de cette étude se propose de décrire les différences entre la clarté perçue et son évaluation par le critère C80.

ÉTUDE DE LA DISTRIBUTION DES PREMIÈRES RÉFLEXIONS

Procédure

Dans cette dernière partie, on effectue une distribution synthétique des premières réflexions, à l'aide d'un simulateur de salle à réponse impulsionnelle totalement contrôlée par découpage de bruit blanc. Les enregistrements numériques de la batterie sont donc filtrés et associés par paires, et les sujets doivent à nouveau évaluer et quantifier les différences de clarté. Les principaux paramètres étudiés dans la distribution artificielle des premières réflexions sont:
Résultats
Toutes les configurations montrent qu’une variation d’écart entre deux réflexions entraîne une variation de la clarté, parce que le sens dépend de la position relative de l’écart et des intervalles relatifs des réflexions (voir figures 1 et 2).

Le modèle le mieux approprié pour décrire ce phénomène semble être l’histogramme des écarts entre réflexions.

On peut en abaisser les valeurs des écarts, en ordonnant le nombre d’écarts qui ont cette valeur. La méthode d’évaluation de la clarté prend en compte la présence d’un grand écart relatif et la position de cet écart.

Une application simplifiée est certaine en sélectionnant des échos impulsionnels comprenant quatre réflexions précoces. En faisant varier leur temps d’arrivée, on peut créer ainsi huit distributions qui ont la même valeur en terme de C80. Le critère de distribution permet de décrire les différences de clarté parce que les enregistrements ainsi obtenus.

Complément sur le rôle de la largeur de réflexion
Le rôle de largeur d’une réflexion a été testé pour deux types de réflexions : une réflexion proche de l’onde directe et un réflexion proche du champ diffus. Pour une réflexion associée au son direct, la sensibilité de la clarté est en raison inverse de la largeur de la réflexion, tandis qu’on observe le phénomène inverse pour une réflexion proche du champ diffus.

Interprétations concernant la clarté
Le son direct est intégré dans un ensemble de réflexions, que l’on sépare habituellement en deux groupes : l’énergie "utile" qui comprend le son direct et les premières réflexions, et l’énergie diffuse, qui intègre les réflexions suivantes et le champ diffus.

Or, la présente expérience démontre que l’énergie n’est pas intégrée de manière globale dans les 80 premières ms, mais que la quantité intégrée avec le son direct, que nous appellerons paquet d’onde direct, dépend de la distribution des premières réflexions.

Les résultats montrent qu’une bonne clarté est liée à un paquet d’onde direct relativement étroit. Cette condition assure donc une bonne résolution temporelle du message, ce qui signifie que les pires directs étant plus étroits sont plus facilement discriminés.

En adoptant un point de vue fréquentiel, on peut dire qu’un paquet d’onde direct étroit entraine un spectre étendu vers les hautes fréquences, donc une bonne perception des attaques, ce qui n’est plus le cas si le paquet d’onde direct est large.

Plus précisément, on peut considérer que l’on a deux types de discrimination du signal musical dans la réverbération :
- une discrimination en niveau : si l’énergie précède est importante, on peut séparer le son direct du champ réverbéré grâce à la différence de niveau
- une discrimination temporelle : si on a un écart proche de l’onde directe, on peut séparer le son direct du champ réverbéré sur l’axe du temps (voir figures 3 et 4).

Conclusions
Les expériences ont montré que les critères habituels de caractérisation des salles de concert ne décrivaient pas complètement la perception. La mesure globale de l’énergie des premières réflexions est un critère très insuffisant, et il est nécessaire de prendre en compte leur distribution. Notamment l’existence de grands écarts entre deux réflexions successives augmente la sensibilité de la clarté dans certains cas et la diminue dans d’autres. La clarté subjective est d’autant plus accentuée que les réflexions précoces sont plus étroites, tandis que les réflexions tardives s’intègrent d’autant mieux dans le champ diffus si elles sont plus larges.

D’une manière générale, il est donc raisonnable de considérer que l’appareil auditif intègre globalement l’énergie comprise dans les 80 premières ms de la réponse impulsionnelle. Le rôle de la distribution des réflexions sur les autres paramètres perceptifs, notamment sur le timbre, reste à approfondir.
Fig 1 et fig 2 : on déplace la réflexion indiquée d'une flèche par pas de 5 ms. L'écartement entre cette réflexion et la précédente est indiqué en abscisse, et la variation de clarté perçue est placée en ordonnée (échelle arbitraire).

Fig 1 : Les 80 premières ms de la réponse impulsionnelle

Variation de la sensation de clarté
(N=3)

Histogramme des écarts entre réflexions.

Pour les histogrammes, on place les valeurs des écarts numérotés, et l'écart fléché est celui que l'on fait varier. L'augmentation de clarté peut être traduite par la présence de l'écart relatif entre les réflexions.

Fig 2 :
Les 80 premières ms de
la réponse impulsionnelle

Variation de la
sensation de clarté
(N=3)

La clarté diminue car il n'y a pas d'écart relatif important entre les réflexions.

Fig 3 :
Distribution jugée "claire" par les sujets : le deuxième écart est plus important que les autres.

Fig 4 :
Distribution jugée "peu claire" par les sujets : il n'y a pas de grand écart relatif.
*13th INTERNATIONAL CONGRESS ON ACOUSTICS*
*YUGOSLAVIA* *1989*

COMPARED ANALYSIS OF THE CLARITY INDEX AND THE REVERBERATION TIME

J.P. Jullien, C. Lavandier*, O. Warusfel

CNET, route de Trégastel, 22301 Lannion, France
IRCAM, 31 rue Saint-Méry, 75004 Paris, France

Introduction
The use of measured criteria to characterize the room acoustic quality has become very common and only very few people, among scientists, would reject this way of doing. Nevertheless, the choice of the "good" criteria remains an openned question considering that the links between the measurements and our perception are not completely understood. That was the motivation of the acoustic laboratory of IRCAM for starting a series of perceptive tests dedicated to it. This paper introduces the method which has been used and comments its utility to establish structured relations between acoustic criteria and the derived perceptive factors. Some results are given to illustrate the method, they concern mainly two criteria: Reverberation Time (RT) and Clarity Index at 80 ms (C80). A second associated paper [Warusfel et al.] develops the results dealing with coloration and spatial effects; a third one [Lavandier et al.] gives a synthesis of all the results now available.

Basic statements
Previous researches, [Reichardt & Schmidt], [Lehmann & Wilken], [Schroeder & Gottlob & Siebrasse], have already proven some basic statements which were considered as assumptions in our study: the number of the important perceptive factors related to room acoustic quality is limited and they should be accessible by experimental psychology methods; the relations between those factors and the acoustic criteria are not easy to establish; the accuracy of the correlation depends on the range of the criteria values; the preference judgments are extremely varying with subjects and musical styles. However, numerous results are now available and were a great help to interpret our results.

Elements of the method
At the starting point, we wanted to get rid of the semantics: that was to avoid overevaluating a perceptive aspect (too many terms devoted to it) or neglecting another one (no term proposed for it). Moreover, we were not completely confident that a qualitative term would be understood the same way by different subjects: we intended to solicitate musicians, students, scientists. As for preference judgments, it is above-mentioned that the intrinsic variability is very high. Then, distance judgments seemed to be the most neutral approach.

As a consequence, most of the preceding information contained in a test consists on the choice of the tested configurations of room acoustics: it explains the very special attention we gave that choice. The redundancy among the acoustic criteria is well known and they were selected so that, for all the tested configurations, the values of the selected criteria were controlled [Jullien]. The main selected criteria are: acoustical level (LEV), late reverberation time (RT), clarity index over 80 ms (C80), direct sound level (Dir), lateral efficiency (LE). A computer controlled sound field simulator (headphones or loudspeaker array in an anechoic room) was especially designed.

As for the study of coloration and spatial effects, we could not do the same way since we do not dispose of such a set of criteria. Then, it has been chosen to discretize the pattern of the first reflections: temporal distribution in two intervals (10-40ms; 40-80ms); spatial distribution in four directions (front-up, front-side, back-side, back). The late reverberation was completely diffused and controlled in each octave band (level and reverberation time).

Now, how to analyse distance judgments in order to get a structured representation of our perception? David Weassel had introduced at IRCAM the INDSCAL procedure and we could have the advantage of the knowhow he had brought. This procedure needs a distance judgment for each subject and for each couple of the tested configurations. Then, for practical reasons, the number of the tested configurations in one session is limited (not more than 9) and multiple tests must be proceeded to
overcome the limitation. INDSCAL procedure takes the distance judgments and constructs a common perceptive space where the points represent the tested configurations; the cartesian distances among the points are supposed to fit the distance judgments; INDSCAL allows a subject to give his own weight to each aspect of the perception; his individual space is derived from the common one by dilating it on each coordinate axis by his corresponding weight. Here, it can be seen that INDSCAL gives a very strong definition of a perceptive factor: only a coordinate axis is a perceptive factor and no rotation is allowed to interpret the positions of the points in the space.

**Interpretation of the coordinate axes**

As INDSCAL gives a very strong interpretation of the coordinate axes, one must determine accurately the dimension of the perceptive space. In most of the multidimensional scaling methods, the dimension is derived from the analysis of the explained variance: the dimension is increased until the explained variance reaches a maximum. This is not completely satisfying because it links an important parameter of the analysis to the accuracy of the subject's answers. So we used an other index, based on the stability of the perceptive space when one configuration or one subject is omitted in the analysis. This index is called the jackknife and was suggested by Suzan Winsberg at IRCAM. The good dimension is the one of the most stable space.

![Diagram](https://example.com/diagram.png)

**Figure 1**: basic structure of the chosen configurations in two different tests

Figure 1 shows the chosen configurations of two tests. For both of them, two criteria are varied over three values; the other controlled criteria are kept constant. The resulting perceptive space appears very clearly in the first, see figure 2: only two factors are needed and the correlations with the criteria are excellent. Here is an example where one can consider that the criteria (LEV, RT) characterize two perceptive factors in their utilization range. On the contrary, the second test needs 4 dimensions and the chosen criteria do not correspond directly to perceptive factors. Figure 3 shows the projection of the points on the two first perceptive factors. Other tests are necessary to understand the first factor; they have shown that it is characterised by the ratio Dir/REF (Direct sound over the early reflections), the correlation with the factor is -0.9. Then, if we trace the vectors between the point-configurations where C80 is varying and Dir/REF constant, the vectors are seen parallel to the second axis. It suggests that C80 characterises a factor when its values are around 0 dB. We shall see that it is no longer true for extreme values (see incoherent position of point 1). In all the cases, the selection of the factors, derived from INDSCAL analysis was confirmed by further unformal listening tests.
Linearity of the perception

As already mentioned, the position of the extreme points are not well explained. It was tried to analyse directly the perceived distances: they are averaged (after a normalization) over all the subjects; it gives a «primary» distance for all the couples of configurations. Three configurations form a triangle on which one can calculate the angles. If only one factor is varying among the 3 configurations, and if the perception of the distance were linear, the angle would be equal to 180°. It was observed that the angle is factor dependant (it gives a confirmation that individual scaling is necessary to reveal factors): the linearity is good for the factor linked to the criterion LEV but rather bad (the angle is around 110°) for factors linked to the criteria RT and C80. It shows clearly that the perception is not linear: the perceptual space given by INDSCAL might be seen as the tangential space of the real perceptual space; only a local analysis can be made and extreme points should be avoided; wider explorations have to be done step by step.

Substitution effect

For each test, the choice of the tested configurations is based on a simple structure related to few chosen criteria (see figure 1 for instance). This structure is preserved in the perceptual space if and only if each chosen criterion is related to one factor; if not, the criterion is said to be substituted on specific factors which must be explained with other criteria. Figure 4 shows the basic structure of a test: the related perceptual space is three-dimensional; figure 5 shows the plane of the two first factors, clearly characterized by RT and C80. The correlation of those two factors is not negligible (+0.55) which suggests a substitution effect of RT on C80 or the inverse. The observation of the perceptual space shows (confirmed by the other tests) that C80 is a complicated criterion: for low values (< -4 dB) and for important reverberation time (> 1.5 s) it is completely substituted on RT which is always a perceptive factor; for values around 0 dB, it is half substituted on RT and a specific factor, well characterized by C80 itself (see the parallelism of the vectors when only C80 is varying); for higher values, C80 tends to be substituted on an other factor related to the criterion Dir, the very early energy. Sometimes, the substitution is due to a perceptive mechanism (substitution of C80 on RT), but sometimes it can be deduced from objective considerations when the criteria themselves are not independent (like C80 and Dir). This last point stresses the importance to use uncorrelated criteria to construct the tested configurations so that the perceptive aspects are not masked.
To quantify the substitution effect, two different methods were used. First, the analysis on the points coordinates given by INDSCAL; second, the analysis of the above-mentioned primary distances. Since distance perception seems not to be a linear process, when points are distributed over a perceptive surface (indeed INDSCAL has given two factors); the mean curvature of the surface is estimated from the polygons formed by the configurations; then, the angles are corrected in order to measure the true angle on the curved surface. The substitution of C80 on RT could be precisely estimated by this method: in the interval [-6 dB, 0 dB], the angle with RT increases from nearly 0° (completely substituted) to 45° (half substituted); for values greater than 4 dB, the angle is 90° (complete independance).

Conclusion

It has been possible to overcome the major difficulties encountered in this study: the dimension of the perceptive space is unknown; numerous substitution effects occur and the revealed factors are also unknown; distance judgment is not a linear process. The used method gives a coherent representation of the perception and allows to establish structured relations between the factors and the objective criteria. The found structure is documented in the two associated papers [Warusfel & col.][Lavandier & col.].

References

PERCEPTION OF COLORATION AND SPATIAL EFFECTS IN ROOM ACOUSTICS

O. Waurusel (*), C. Lavandier (**), J.P. Julien (***)

(*IRCAM, 31 rue Saint-Merri, 75004 Paris, France
(**)CNET, route de Trégastel, 22301 Lannion, France

INTRODUCTION

This paper relates part of a research program that IRCAM has proceeded on the perceptive characterization of acoustical quality. The experimental procedure and analysis methods are presented in associated papers [Julien89][Lavandier89]. It is just reminded here that in order to sort out different perceptive effects, an artificial acoustic field simulator was used, where objective parameters could be independently controlled. Several tests, consisting each of a set of listening configurations, were analysed through dissimilarity judgments and led to the derivation of perceptive sub-spaces.

The present paper focuses on some tests dealing with coloration and spatial effects. The analysis of relations between the corresponding perceptive spaces and objective data provides useful elements for the choice of the most relevant objective criteria.

COLORATION EFFECTS

Under this concept are both included the global frequency behaviour of the room response, due mainly to absorptive characteristics of materials, and the particular coloration effects due to the time distribution of first strong reflections. It is of daily experience that the choice of materials in concert halls modulates the energy and duration of the reverberation field in function of frequency. Besides, the use of panels or overhanging reflectors close to the musicians, so as to improve the intelligibility, provides simultaneously a comb filtering effect. The following experiments deal with these different aspects.

Total Energy/Octave Bands dependence

The eight configurations constituting this test present different slopes of the total energy in function of frequency. As described in Fig.1.a two regions are considered: above and under 1kHz. For low frequencies one positive and one negative (+1.5dB/oct.) slopes are used while for high frequencies the choice is limited to a null or negative (-1.5dB/oct.) slopes. Each of these configurations can be presented with two values of reverberation time (1.7 and 2.3 sec). As total level was kept constant negative offsets can be noticed on configurations A,B,D and E,F,H.

The analysis of error criteria (percentage of explained variance, and jackknife [Julien89]) leads to an optimal perceptive space with four dimensions. As shown by Fig.1.b, apart from configuration H, pairs where high slope is varied are parallel and high correlated with the first axis. The angle that is noticed between 'low frequency pairs' and 'high frequency pairs' is objectively explained by the offset used to keep total level constant. Therefore the two first perceptive factors related with timbre perception are respectively governed by level in bass and treble frequencies.
Looking at plan 2-3 (Fig. 1.c) shows that variations of low frequencies level require two dimensions. In fact this third dimension is introduced only by the extreme configuration H (positive slope for low freq. and negative slope for high freq. at high reverberation time). Moreover Fig. 1.d, presenting the projection of individual preference, leads to interpret this dimension as a rejection factor.

The last factor (not presented on the picture) was high correlated with reverberation time.

Fig 1. Perceptive space derived from variation of energy in function of frequency. (a) Objective description of configurations: slope of energy in function of frequency. (b) (c) Plan projections of the configurations in the correspondent perceptive space. 

\[ \text{denotes increasing level at high frequencies} \]
\[ \text{denotes increasing level at low frequencies} \]
(d) Plan projection of individual preferences.

Reverberation Time/Octave Bands dependence

A similar experiment was conducted with reverberation time slopes in function of frequency: positive and negative slopes for low frequencies under 1kHz, null and negative slopes for high frequencies, two central reverberation time (1.5sec and 2.5sec).

The analysis of this test led to a four dimension perceptive space. Two axes were needed to explain variations of reverberation time in high frequencies, one of which was introduced by an extreme configuration: high central reverberation time with negative slope for low frequencies and null slope for high frequencies. As in previous experiment this configuration, which enhances high frequency reverberation time, was rejected by preference judgments.

Moreover the high frequency factor showed substitution effect with the central reverberation time factor. Although this confusion disappeared when the analysis was restricted to introduced subjects (acousticians and authors), this result tends to prove the leading effect of high frequencies in reverberation time perception.

On the contrary the last factor, corresponding to low frequency reverberation time, was independant of all the others, and thus mainly denotes a timbre effect.

First Reflections distribution

In next experiment the time distribution and relative level of reflections will affect both clarity and coloration. The control parameters of that distribution were the Early Center Time (ECT) - computed as usual center time but on first reflections only), and the spreading of reflections (measured by the standard deviation linked to ECT).
In a first experiment, still displayed on headphones, these variations were mixed with the level of the reflections distribution and the level of the reverberation process. As the original aim of this project was to validate usual objective criteria, those levels were adjusted so as to obtain specific values of Clarity Index (C80), Dir/Rev (level ratio between direct sound and reverberation: denoted H) (Fig. 2.a).

The plan 1-3 (Fig. 2.b) of the corresponding four dimensional space clearly separates the two factors related to time distribution: the standard deviation (axis 1), and the early central time (axis 2). However the importance of standard deviation must be reduced in regard to the amount of spreading that was needed to be perceived. As a matter of fact limen difference analysis which may be derived from the study of INDSCAL, perceptive space gives, for this criterion a value of 20 ms which can be, from an objective point of view, hardly exceeded.

The plan 2-4 related to energy distribution (Fig. 2.c.), shows that C80 variations will not be perceived on the same factor whether early energy time distribution is preserved (BC, HI) or not (AB, GH). Although the configurations D and F present paradox linear combinations of these factors - due perhaps to masking of reverberance by a louder reflection close to the direct sound - this test proves that when C80 is governed by increasing of the early reflections level a new factor will appear and is correlated with Dir/REF (level ratio between DirectSound and Reflections). When looking for semantic correspondence this factor was interpreted as a timbre effect.

**Fig. 2.a**

**Fig. 2.b**

**Fig. 2.c**

**Spatial Effects**

It is well known that spatial distribution of reflections are involved in space perception [Barron81]. Several tests were proceeded with a multichannel reproducing system, installed in an anechoic room, in which we could control the direction of the different signals [Jullien89]. This system allowed to verify the previous results derived from headphone displayed tests, and to approach spatial aspects of perception.

In the present test the control parameters are the choice of direction(s) from which reflections are displayed. For configurations ACEG the reflection signals are displayed on upper front and front side loudspeakers, while for configurations BDFH they are displayed on front side or back side loudspeakers. In order to study the relations between space and time distributions the reflection process could be separated into two
intervals: one between 10 and 40ms and one between 40ms and 80ms, each constituted of a single reflection or a regular distribution with same total level (Fig 3.a).

In the associated perceptive space the first axis is strongly correlated with variation of Lateral Efficiency (ED, EF, EB...), while the second is related to a criterion which is a modified early center time (Fig 3.b). The spatialisation of reflections leads to introduce time and direction weighting in order to reduce the importance of late back reflections. This tends to prove that back reflections are sooner integrated to the reverberation perception: as for configuration H.

On plan 1-3 (Fig 3.c) a new factor linked to the time and spatial distribution appears. This factor has been interpreted as perceptive width of source: adding first reflections close in time and space to the direct sound will enlarge the width impression of source area. This effect can be denoted with good correlation by a criterion that extends the direct sound notion to first reflections with delay and angle weighting. That notion of direct sound area also governs the perception when clarity reaches high values as direct sound did in headphone experiments dealing with clarity perception [Jullien89].

**Fig.3 Perceptive space derived from variation of time and spatial distribution of reflections. (a) objective description. (b)(c) Plan projections of perceptive space**

---

**CONCLUSION**

These experiments and methodology revealed efficient analysis of the relation between objective parameters and perceptive factors, even when complex perceptive paths were involved by control parameters variations. It has led to modulate some usual criteria or to introduce new ones so as to achieve a closer description of this perceptive space. Being able to control these perceptive factors, further listening experiments were conducted so as to propose a semantic correspondance with each of them [Lavandier89].

**REFERENCE**


PHYSICAL AND PSYCHOACOUSTIC EVALUATION OF FM SIGNALS PROPAGATING IN A ROOM

E. Ozimek

Institute of Acoustics, J. Mickiewicz University
ul. Matejki 43/69, Poznan, Poland

Introduction

The investigations pertain to the physical analysis and a psychoacoustic evaluation of changes in the frequency structure of FM signals propagating in a room. The physical part of the investigations consisted in the measurements of changes in the deviation of the frequency modulated signal propagating in a room, with respect to the modulating frequency, carrier frequency and the deviation value of the modulated signal transmitted to the room.

The psychoacoustic part of the investigations consisted in the determination of the perception threshold of changes in the deviation of FM signals with respect to the modulating frequency, carrier frequency and the deviation of the reference signal.

Changes in the Deviation of a Frequency Modulated Sound Propagating in a Room

The transfer of frequency modulated sounds by a room is a very complex and a comparatively new problem. So far only a few works have concerned this problem [1,5,4]. This is an important issue not only because it provides insight into the matter but because it is probably closely related to the evaluation of the intelligibility of speech sounds and the evaluation of the musical sounds in a room. Some effects resulting from the deformation of FM signals in a room are presented in papers [1,4]. According to these papers, a frequency modulated signal by a harmonic signal can be described as:

\[ \frac{U}{t} = U \cos(\Omega t + \phi) + B \sin(\Omega t + \Theta) \]

where:

\[ \Omega, \phi \] - carrier frequency
\[ \Omega, \Theta \] - modulating frequency
\[ \Phi = \frac{\Delta\omega}{\Omega} \] - modulation index
\[ \Delta\omega \] - signal deviation

After a superposition of the two waves described by expression /1/ i.e. the direct and the reflected wave delayed by \( \Delta t \)
and the simple transformation of the obtained results we get the resultant carrier frequency in form

$$\omega_{c} = \omega_{b} + \Delta \omega k_{2} \cos[\Omega / t - t_{n} - \frac{\Delta t}{2}]$$

where:

$$k_{2} = \cos \frac{\Omega \Delta t}{2}$$

Thus the resultant change of deviation is

$$\Delta \omega_{c} = \Delta \omega \cos \frac{\Omega \Delta t}{2}$$

It follows from expression /3/ that the resultant deviation changes in accordance with the modulating function. The maximal value of the deviation depends on the modulating frequency and the delay time of the reflected wave, hence only for small modulating frequency /\Omega \rightarrow 0/ and short echo delay times /\Delta t \rightarrow 0/ the deviation of the received and transmitted signals is almost equal.

The investigations were conducted in two acoustically different rooms /denoted as S and K/ with the reverberation time equal respectively, 1 and 4 seconds. The modulating noise band were generated by the microcomputer and analog-digital converter system.

The signals used were characterized by the following range of changes of the basic parameters:

- mid-frequency of the modulating signal: 1; 2; 4; 8; 16; 32; 64; 125 Hz
- carrier frequency: 0.25; 0.5; 1; 2 KHz
- deviation of the signal transmitted to the room: 12, 3; 25; 50; 100; 200; 400 Hz

Measurements of deviation changes were taken in a few selected measurement points of the rooms under study. The measurements consisted in the determination of quantity R which defined the difference between the RMS deviation value of a signal recorded in a given point of the field /\Delta f_{n}/ and the RMS deviation value of a signal transmitted to the room /\Delta f_{n}/

$$R = \Delta f_{o} - \Delta f_{n}$$

In Fig.1 a dependence of deviation changes in a frequency modulated signal, expressed by quantity R, on the modulating frequency has been shown, for room S. The deviation value of the signal transmitted to the room is the parameter of this dependence. On the basis of Fig.1 we can generally state that the difference between the deviation of signals transmitted to and received in a room increases with an increase of the modulating frequency /f_{m}/. Moreover for a fixed modulating frequency, the greater the deviation of the transmitted signal, the greater the value of the parameter R.

The dependence of deviation changes of an FM signal in the function of deviation value of a signal transmitted to
the room is also very interesting. This dependence, determined in rooms S and K, for modulating frequency $f_m = 8 Hz$ is shown in Fig.2.

![Graph](image)

**Fig.1.** A dependence of deviation changes in an FM signal propagating in a room on: modulating frequency $f_m$ /Fig.1/ and on deviation of the transmitted signal /Fig.2/.

**Perception of Deviation Changes in an FM Signal Propagating in a Room**

Deviation changes occurring in an FM signal propagating in a room are interesting from the psychoacoustic viewpoint. It can be assumed that if the deviation changes measured are smaller than their perception threshold, the changes have no significance for the evaluation of the acoustic properties of the room. However, if the perception threshold of deviation changes is much smaller than deviation changes really occurring in a room, the changes can be significant for the evaluation of the acoustic properties of the room.

In this investigation, the dependence of the difference limens of the frequency modulation on the modulating frequency was determined. This dependence is shown in Fig.7. Threshold curves indicate an averaged course for three subjects.

Having the results of the physical investigation pertaining to deviation changes in FM signal in a room and difference limens of the perception of these changes, we compared some of these data, which is shown in Fig.4. This comparison presents to some extent a monographic evaluation of the subject's perception of deviation changes in FM signal, really occurring, in a room /3,2/.

The $x$-axis of diagram presents the measure of the values of the absolute threshold /--A--/ and the difference limen
of the perception of deviation changes in FM signal. It also represents the difference between the deviation of the signal transmitted to and received in room S and K, values of the modulating frequency have been marked on the X-axis.

As can be seen in the figure, for the FM signal with parameters of $f_0 = 1000 \text{ Hz}$ and $\Delta f = 50 \text{ Hz}$, the values of deviation changes measured in rooms S and K are above their perception thresholds. This means that these changes should be easily perceived by the subjects in these rooms. This fact can be significant for the subjective evaluation of the acoustic properties of rooms, mainly for the evaluation of the intelligibility of speech and music sounds.

### Literature

1. E. Ozimek, L. Rutkowski
   - Deformation of Frequency Modulated FM/ Signals Propagating in a Room
   - Applied Acoustics 26, 1989

2. E. Ozimek, A. Säk
   - Perception of Irregular Frequency Changes of Sinusoidal Signals
   - Acustica Vol. 65/66 /148-152/, 1988

3. E. Ozimek, A. Säk, R. Piotrowski
   - Determination of Deviation Changes in FM Signals Propagating in a Room
   - Poznań - 1987 /unpublished materials/

4. L. Rutkowski, E. Ozimek
   - Changes of the Deviation of FM Signals in a Room
   - Architectural Acoustics, Strefskie Płosko 202-203/, 1985
3. C
ORCHESTRA ENCLOSURES AND STAGE DESIGN IN MULTIPURPOSE HALLS
USED FOR CONCERTS.

S. Strøm

Acoustics Research Center
The Norwegian Institute of Technology
N - 7034 Trondheim , Norway

The new civic centers in Tynset, Lørenskog, Nord-Odal and
Namsos (all finished 1988) are typical for many similar
centers built in Norway for the last years. Here we often find
a library, a theater workshop, exhibition halls, meeting rooms
and a multipurpose hall for cinema, theater, opera and concert.
The acoustical design of such a hall will always be difficult.
Especially concerts with a symphony orchestra on a theater
stage, surrounded by highly absorbing curtains, will be doubt-
full as to the acoustical design (figure 1). The use of stage
reflectors in front of the curtains will be absolutely
necessary. Such reflectors do not have to be heavy. Our
experience from many civic center projects have shown that a
1 mm thick plastic foil gives satisfactory strong reflections.

Orchestra enclosures for a theater/opera stage.

Marshall (1) has shown that to obtain a good ensemble for the
musicians it is the early reflected energy in the
higher frequency range that are most important. And
Schultz (2) concluded that
early sound arriving in the
audience do not need to have
a full frequency content. Most listeners will not
notice the missing low
frequencies in the early
sound. This may explain the
successfull use of light-
weight plastic foil for the
stage reflectors. The
physical data for such a
material are shown in figure
2-5. Figure 6-7 show how
such reflectors are used in
an orchestra enclosure.

--- without theater stage
curtains (concert hall)
----- with curtains (theater)

Figure 1. Measured reverberation time in the 500 seats
theater/concert hall in NAMSOS civic center (without an audience).
Figure 2. Measured transmission loss for a 1 mm thick plastic foil (glassfiber reinforced PVC-foil). The foil was hung vertically in front of the test opening for doors in the reverberation room.

Figure 3. Measured reflection and absorption factors for the 1 mm thick plastic foil (see also figure 1). The foil (2m x 2m) was hung vertically in the anechoic chamber, and we used an intensity measuring technique.

Figure 4. Calculated reflection factor for the 1 mm plastic foil (see also figure 2-3). Notice that these calculated data will be more accurate in the lower frequency range than the measurement data.
Figure 5. Measured sound transmission loss and calculated reflection factor for a 2 mm thick fire curtain to be used as stage enclosure in the new TRONDHEIM civic center concert/opera hall (finished 1989).

Figure 6. The orchestra enclosure design in SKENE theater/concert hall (Sweden), using the 1 mm plastic foil reflectors.

Figure 7. The orchestra enclosure design in the new concert/opera hall OLAVSRALLEN in Trondheim. Notice the use of a 2 mm thick foil (see figure 5) and the extra plywood enclosure inside the foil enclosure.
Stage reflector for sports arenas used as concert halls.

When large sports arenas are used as concert halls for symphony orchestras, we have to consider:

- the orchestra need a stage for the performance
- the roof height above the stage will often be too large, and the roof is seldom designed as a sound reflector.
Therefore we need a specially designed sound reflector hanging above the stage (figure 8).
- with an audience of more than 1000 we must expect that the reverberation time might be too short. Therefore we have to consider the use of an electroacoustical reverberation system.

Acoustics Research Center in Trondheim has made the acoustical design in many large sports arenas for concert arrangements:
- Norrköping, Sweden (1984)
- Steinkjer, Norway (1986)
- Trondheim, Norway (1986)
- Kinna, Sweden (1987)

Musicians from Moscow Symphony Orchestra expressed great enthusiasm about the acoustics in Norrköping and Steinkjer. In Steinkjer and Trondheim we used a temporary Multi-Channel-Reverberation system (MCR-system) to increase the reverberation time (figure 9). The 20 channel system was manually operated, and therefore prof. Krokatad could adjust the system closer to uncontrolled feedback. Despite the few channels he obtained a satisfactory increase of reverberation time. Besides, the distribution of loudspeakers above the audience gave a more even sound distribution in the hall (correcting the curved roof).

References:

Figure 8. The acoustical design of a plastic foil reflector hanging above the orchestra stage in a sports arena.

Figure 9. Calculated reverberation time in the sports arena in Steinkjer when used as a concert hall. The calculated values for an audience of 1300 are based on measurements in the empty hall.
On the Acoustic Feedback Model for Solo Singers on Concert Hall Stages.

S. Shirauna and S. Nakamura

Tokyo National University of Fine Arts and Music

12-8 Ueno Koen, Taito-Ku, Tokyo-To, 110 Japan

When singing on stage, as pointed out in our previous presentation [1], soloists require their surroundings to have certain reflective properties in order that they may feel acoustically supported by the hall. After the subsequent experiments since 1986, the contribution of the individual on-stage acoustic components became clear, and then, finally a model of an acoustic feedback circuit has been assumed based on the results of a series of our fundamental studies started in 1982.

Procedure of Experiments

Among post-1986 subjective experiments, the following one outlined in Fig.1 and 2, is especially important. It is based on the idea that on-stage acoustic conditions were considered to be composed of the reverberant sound and the frontal reflection.

Subjective responses of ten singers were determined under eighteen different conditions made by changing the relative levels of those two variables and the reverberation time of the reverberant sound.

![Fig.1 Outline of experimental set-up.]

![Fig.2 Time pattern model of sound pressure level.]

<table>
<thead>
<tr>
<th>Table 1 Results of correlation coefficient analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>①</td>
</tr>
<tr>
<td>②</td>
</tr>
<tr>
<td>③</td>
</tr>
<tr>
<td>④</td>
</tr>
<tr>
<td>⑤</td>
</tr>
<tr>
<td>⑥</td>
</tr>
<tr>
<td>⑦</td>
</tr>
<tr>
<td>⑧</td>
</tr>
<tr>
<td>⑨</td>
</tr>
<tr>
<td>⑩</td>
</tr>
<tr>
<td>⑪</td>
</tr>
<tr>
<td>⑫</td>
</tr>
<tr>
<td>⑬</td>
</tr>
<tr>
<td>⑭</td>
</tr>
<tr>
<td>⑮</td>
</tr>
<tr>
<td>⑯</td>
</tr>
<tr>
<td>⑰</td>
</tr>
<tr>
<td>⑱</td>
</tr>
<tr>
<td>⑲</td>
</tr>
<tr>
<td>⑳</td>
</tr>
<tr>
<td>㉑</td>
</tr>
<tr>
<td>㉒</td>
</tr>
<tr>
<td>㉓</td>
</tr>
<tr>
<td>㉔</td>
</tr>
<tr>
<td>㉕</td>
</tr>
</tbody>
</table>

Results and Considerations

As shown in Table 1, subjective responses can hardly be related to either acoustic components in a simple way, except for the case of ①. Even in the case of ②, the influence of the reverberant sound cannot be denied.
Such complicated relationships between subjective responses and acoustics can be recognized much more clearly by looking at Fig.3.

Using even still more different analyses, results like these in Fig.4 have been obtained. These kind of illustrations would be most helpful for further consideration.

They show that the combining conditions of those two components are, in general, essentially important as the background of the acoustic perception, and that there are some typical coupling patterns such as {Seasoning Effect} (Trade-off Effect) and the others peculiar to each item.

Fig.3a Relationship between Fig.3b - and Detection of [Acoustic Conditions] and [Frontal Reflection]. [Detection of Reverberant Sound].

Fig.3b and [Detection of Fullness of Room Resonance].

Fig.3c and [Ease of Singing].

Fig.4 Subjective response distribution.

Fig.5 Time pattern model of sound pressure level for equivalent conditions.

Fig.6 Equivalent conditions of reverberant sound for [Fullness] and [Quantity].
Perceived Characteristics of Room Resonance

Three different kinds of subjective acoustic factors should be noticed for room resonance from the viewpoint of solo singers on stage.

1. Fullness of Room Resonance: The original Japanese used in the experiments was "tegatæ" [response to hands]. This factor was found to be related primarily to the sound-pressure level of reflections in a stationary state. It hardly matters whether the level is of the reverberant sound or of the frontal reflection. Accordingly, reverberation time has very little effect on it. The relationship of the levels between two acoustic components like this can be classified as a (Trade-off Effect).

2. Quantity of Room Resonance: It is, quite naturally, controlled by the reverberant sound. The equivalent condition in perception is found to be obtained under certain combinations of the values of the reverberation time and the stationary sound-pressure level.

The combining conditions of (Trade-off type) are found within those two factors as shown in Fig.5. This means that all of the decay curves which give the same impression in (Quantity), cross each other at a certain time point as shown in Fig.5.b. Those characteristics have been confirmed by other subjective experiments taking quite different and strict procedures[2].

3. Direction of Room Resonance: The obtained results show that the reverberant sound can be heard coming from the audience side, as long as there is a frontal reflection of a certain, not necessarily high, level. This kind of combination effect should be called a (Seasoning Effect of Frontal Reflection).

Fig. 7 Subjective response distribution for Musical Interpretations of Room Acoustic Conditions.

Background of Musical Interpretations of Room Acoustics

As shown by such complex patterns in Fig.7, from the viewpoint of singer, neither of those impressions can relate in a simple way either to the acoustic components or to those items of the room resonance impressions. Continuous changes can be imagined for each item with the change of reverberation time between those extreme patterns shown here.
Model of Acoustic Feedback Around a Solo Singer on Hall Stages

Considering the situation of actual concert stages, the role of the stage enclosure has been taken into account in this model. Namely, as shown in Fig. 8 and 9, it gives a localized reverberant sound field of relatively higher stationary level and of shorter reverberation time, than those related to the whole volume of the hall.

In addition, it should be noticed that the direct sound rather negatively affects on detecting reflections, as it gives the forward masking effect[3].

Fig. 8 On-stage acoustic model in a large hall.

Fig. 9 Time pattern model of sound pressure level for on-stage acoustics in a large hall.

Fig. 10 Acoustic feedback model around a solo singer on hall stages.

Acknowledgements
We are grateful for the assistance of many students as the subjects of our tedious experiments.

References
THE ACOUSTICS OF THE NEW "OPERA DE LA BASTILLE" IN PARIS

H.A. MÜLLER, J.P. VIAN

MÜLLER BBM, Robert Koch Strasse 11 8033 PLANEgg B. MUNCHEN/GERMANY
CSfB, 24, rue Joseph Fourier - 38400 SAINT MARTIN D'HERES/FRANCE

INTRODUCTION

Six years ago the French President de la République has decided to build in Paris a new modern and popular opera house, fitting as well as possible the contemporary requirements for lyric art.

The two basic points of the project which have influenced deeply the programme are the words "modern" and "popular".

Modern is easy to understand. But popular means different things. First of all, it means a large number of seats at a reasonable price, which in turn has a lot of consequences about the number of performances, alternation, and therefore large spaces to rehearse to store decor etc... Second, it means that the very classical and a little bit aristocratic organisation of the horse shoe italian theater has to be replaced by a more democratic approach, giving the best and the most equal view and acoustical conditions to the largest number of people as possible. This last point corresponds to an actual acoustical challenge for a large lyric theater.

The building contains four main halls:
- the large hall for operas and concerts - 2700 seats
- the modular room for operas, concerts, dramas which can be transformed from a classical concert hall to an italian theater or an experimental hall with seating capacity ranging from 300 to 1300 seats
- the amphitheater (semi-circular) for small, simple and varied performances - 800 seats
- the studio - 300 seats, already opened in April 87 in an attached building called Tour d'Argent, for conferences, master classes, recitals and small musical performances.

The tremendous large stage system is composed of 7 stages of the same size as the main stage, with elevators, turning table and corresponding storing area at the basement of the building.

A lot of rehearsal rooms are also provided for stage, orchestras, chorus, ballet and so on...

The acoustical program specified a reverberation time of 1.1 s to 1.4 s, an EDT of 1.3 s, and a strength index greater than -29 dB. The background noise had to fulfill the NR 20 curve, and the sound insulation had to be less than 70 dB(A) between sensitive or noisy rooms. Wishes about early reflections from the auditorium to the singers was also expressed along with some general rules such as angles of the walls, ratios of underbalconies spaces to the main space, and maximum width of the hall.
It is also important to point out that all the proscenium area around and above the stage opening and the orchestra pit are movable to achieve various relations between audience and stage, depending on the kind of opera performed and the stage design. Consequently the iron curtain must be placed between the orchestra pit and the audience and all the very important surfaces around the singers' place are light weighted and might be not perfectly connected or oriented.

The client has committed the CSTB together with Müller BBM to do the acoustical planning of the whole building, including sound insulation, noise from scenographic or technical equipments, and vibrations from the metro.

ACOUSTICS OF THE MAIN HALL

In harmony with the client, we have proposed for the main hall, beyong the classical high clarity for speech intelligibility in a lyric theater, a rich acoustics a little bit more reverberant than the usual lyric theaters (1.6 s), with strong lateral reflections, and the loudest sound as possible to take up the large distances of the hall.

To achieve these goals a lot of attention has been payed to the proscenium area which has to give a reflection for every seat and good acoustical coupling between the stage and the pit.

The lateral walls have been chosen parallel to the axis in order to create good first order lateral reflections, and the distance between the farthest seat and the stage has been reduced as much as possible.

The ceiling was adjusted to give an even diffused reflection for every seat and the disturbing effect of the light bridges was decided to be the smallest possible. The volume per seat was adjusted to be close to 8 m³.

1) Computer simulation

As usual the first step of the acoustical planning was to discuss the gross shape of the hall with the architect. Without going into details we can tell that we have proposed several modifications on the competition sketches, mainly to change the radial orientation of the walls, to adjust the volume and to improve the functioning of the ceiling.

To support these demands and to evaluate their actual influence we have run computer tests using the Epidauré programme at different steps of the design.

As already published, the Epidauré programme is based on a cone tracing method to easily determine the image sources up to a very high order, whatever the shape of the hall. It calculates impulse responses for different receiving points in a hall, maps of criteria over the audience area, and it allows to listen to the sound of the hall.

Up to 10 different numerical models were tested and for each most of the usual criteria (SPL, RT, EDT, C80, D50, LE) were analyzed. Echoes' occurance were also investigated.

Two important questions were strongly discussed with the architect, one about the width of the hall and the second about the shape of the proscenium (side walls and ceiling). From an acoustical point of view the proposed width of 38 meters seemed much too large but the large number of seats of the hall made difficult to reduce it. When considering the SPL map and the lateral efficiency map over the orchestra seats calculated with 38 m between lateral walls and with only 32 m, it has been decided to
reconsider the positioning of the seats and to place the walls at only 32 m. The differences between the SPL at the front seats and at the back seat has been reduced from 12 dB(A) with radial walls, to 6 dB(A) with the parallel walls (38 m), and then to 4 dB(A) with parallel walls (32 m).

A similar procedure happened about the proscenium area. The decision was made to design the movable vertical and horizontal elements in such a way that it cannot be possible to misorient them or to make them too much absorbing by letting open the light windows, because of the loss of level at the farthest seats.

2) 1/20 scale model test

As the client had ordered a scale model test of the final design, we didn't make computer simulation till the end of the studies, and we ended the project with the acoustical model.

The mean dimensions of the hall are: – main floor length 36 m – width 32 m – height 23 m – maximum distance from the singer at the first balcony 41 m – at the second balcony 48 m. The volume is 21000 m³ and the volume per seat is 7.7 m³. The walls in the audience area are covered by granit stones, the floor is made out of thick wood, the ceiling is made out of 6 mm glass tiles rigidly clamped together, the under-face of the balconies are staff and all the remaining surfaces around the stage are wood. The seats have a wooden reflecting back. They are designed to have the upholstered part completely covered by the spectator when occupied and to be the most absorbing as possible when unoccupied.

The acoustical model was built with dense polystyrene board or lacquered wood for all the reflecting surfaces. The seats rows were simulated by small vertical partitions of the scaled down shoulders' height, partly covered by a strip of velvet. The seats’ absorption was adjusted to the known values of occupied seats, using a 1/20 scale reverberant room. The model reproduced the architectural details greater than around 10 cm full scale and was filled up with dry air (relative humidity lower than 3%) to ensure the right high frequency air absorption.

The microphones used were the B & K 1/8" with nose cone. For binaural measurements we have used a small and simple dummy head with two microphones on each side. The source was a spark generator specially designed for model test measurements. 15 measurement points were spread over the audience, and the source positions tested were either on the stage or in the pit. Several measurements were done to analyze the stage-pit relation and the acoustical feedback from the hall to the singers and orchestra.

For each receiving point the impulse response was digitally recorded, and all the following criteria were derived by computer (RT, EDT, C60, D50, S/N, G). The strength coefficient G was directly calculated from the impulse response for integration times of 30 ms and 500 ms by the formula:

$$ G = 10 \log \left( \frac{S_0 \cdot E_{500}}{\frac{1}{4} \pi r^2 \cdot E_{ DIR}} \right) $$

with $S_0 = 1 \text{ m}^2$, $r = \text{distance to the source}$, $E_{500} = \text{energy of the impulse response before 500 ms}$, $E_{DIR} = \text{energy of the direct sound measured on the first peak}$, if the source is visible from the receiver.

193
Echograms for each receiving points were also calculated and integrated with an exponential time constant of 35 ms to test the possibility of echoes. In addition every recorded impulse response was listened after scaling down the frequency to detect echoes. The sound of the hall was also investigated by listening to anechoic music convolved with binaural impulse responses in order to feel the difference between good seats and far seats or seats below overhang balcony.

The results of the tests were quite satisfactory. With small adjustments of the first part of the ceiling it was possible to maintain the strength coefficient between -31 dB(A) at 10 m from the singer and -36 dB(A) at the worse seats. The reverberation time was 1.6 s at 500 Hz and the clarity 80 ms ranged at 500 Hz between 2 and 6 dB.

Due to an architectural demand to simplify the appearance of the fixed proscenium surface we have proposed to replace the random diffuse reflecting elements by 2D primitive roots diffuser. No change in reverberation time could be observed, but a slight improvement in homogeneity of most of the criteria was sensible.

ACOUSTICAL MEASUREMENT

At the date we are writing this text the acoustical results are not available because the hall is not already finished. Acoustical measurement and musical test have to be carried out during the two months before the first opening on July 14th 1989. We will report on the measurement results and the acoustical quality of the completed hall during the presentation of this paper.
ANALYSE ACOUSTIQUE DU STADE OLYMPIQUE DE MONTREAL

Jean-Gabriel Migneret
Laboratoire d'acoustique, Centre de recherches en aménagement et en développement (CRAD), 1624 pavillon Félix-Antoine Savard, Université Laval, Ste-Foy, Québec, G1K 7P4

1. INTRODUCTION

Conçu par l'architecte Roger Taillibert pour accueillir les Jeux olympiques de 1976, le Stade olympique de Montréal fut l'un des ouvrages d'art les plus controversés du Canada. C'est l'un des plus beaux monuments de Montréal et certainement, le chef-d'œuvre technique le plus audacieux. La Régie des installations olympiques (RIO), maintenant en charge du complexe, doit en assurer l'entretien et développer une utilisation optimale du Stade pour différentes activités sportives comme pour des spectacles à grand déploiement, tels récemment le groupe "Pink Floyd" ou l'opéra "Aida". Pour cerner les dimensions impressionnantes de l'ouvrage, ainsi que ses principales caractéristiques, on peut se reporter au tableau suivant:

<table>
<thead>
<tr>
<th>Principales caractéristiques physiques et acoustiques du Stade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume ouvert: 1,948,587 m³</td>
</tr>
<tr>
<td>Volume avec doubleur d'hiver: 2,115,003 m³</td>
</tr>
<tr>
<td>Volume avec toile déployée: 2,275,013 m³</td>
</tr>
<tr>
<td>Surface occupée par les sièges: 27,440 m²</td>
</tr>
<tr>
<td>Volume par siège (avec toile déployée): 42,8 m³/pers.</td>
</tr>
<tr>
<td>Surface par siège (grands seulement): 0,517 m²/pers.</td>
</tr>
<tr>
<td>Surface totale déployée (toile déployée): 182,505 m²</td>
</tr>
<tr>
<td>Longueur et largeur totales du terrain: 207,9 m et 140,8 m</td>
</tr>
<tr>
<td>Hauteur minimale et maximale de l'anneau technique: 41,5 et 50,8 m</td>
</tr>
<tr>
<td>Point le plus haut de la toile: 70,7 m</td>
</tr>
<tr>
<td>Point le plus haut du stade: 168,4 m</td>
</tr>
<tr>
<td>Temps de réverbération moyen avec toile déployé (vole 500 et 1000 Hz): 15,05 s</td>
</tr>
<tr>
<td>Temps de réverbération moyen avec doubleur d'hiver (vole 500 et 1000 Hz): 10,02 s</td>
</tr>
<tr>
<td>Plus long temps de réverbération mesuré avec une source ponctuelle: 23,89 s</td>
</tr>
</tbody>
</table>

2. MESURE DU TEMPS DE REVERBERATION ET CALCUL DES PARAMETRES ACOUSTIQUES

Les temps de réverbération ont été mesurés à l'aide d'un analyseur 2131 de "Brüel & Kjaer", interfacé avec un ordinateur "Hewlett Packard" 9816. Le logiciel utilisé a été développé au Laboratoire d'acoustique de l'Université Laval [1]. Dans tous les cas, cinq déclenchements au minimum ont été enregistrés et ce sont les régressions moyennes qui ont été calculées pour toutes les bandes de haute correspondances. Ces temps de réverbération ont été relevés selon deux conditions principales du Stade et avec deux types de sources, soit avec l'ensemble du système de sonorisation ou bien avec un 'cluster' ponctuel de haut-parleurs d'environ 400 Watts placé à l'une des extrémités de l'axe longitudinal. Les résultats moyens pour la situation avec la toile de couverture (position normale d'état) et ceux pour la situation avec la doubleur d'hiver (vole avec laine minérale tendu directement sur le dessus de l'anneau technique), sont l'objet de la Figure no 1.

Il est très important pour l'étude du Stade, étant donné son volume de plus de 2 millions de m³ avec sa toile de couverture d'état, de tenir compte de l'absorption moléculaire de l'air [2]. Les relations utilisées ensuite pour calculer les coefficients moyens d'absorption de toutes les parois du Stade tiennent ainsi compte de la température et de l'humidité relative nées lors des mesures de réverbération. Ces relations sont dérivées des formules classiques de SABINE et d'EYRING, avec la proposition de JAÉGER [3] pour la longueur du litre parcouru moyen Lm, soit par exemple la formulation de LECLERcq [4] pour les grands locaux:

\[ TR = \frac{6 \cdot Lm}{\log (1 - \alpha) \cdot \alpha = \text{m} \cdot \text{Lm}} \]

Nous avons ainsi calculé les coefficients moyens d'absorption pour les surfaces déployées du Stade, selon les deux conditions de la toile de couverture. Ensuite nous avons posé le raisonnement suivant, soit qu'essentiellement en position actuelle l'absorption aux basses fréquences était dictée par la présence de la toile de couverture, alors qu'aussi hautes fréquences, cette toile lisse de Kevelar de poids de 3 mm d'épaisseur ne pouvait présenter qu'une absorption résiduelle minime. Une régression d'équation:

\[ TR(\text{théorique}) = 24.86 - 1.88 \cdot 10^{-2} f + 7.65 \cdot 10^{-6} f^2 + 1.67 \cdot 10^{-9} f^3 + \ldots \]

donne le temps de réverbération théorique du Stade (avec seulement sa toile d'état), sans l'effet des basses fréquences en dessous de 315 Hz. On peut donc décrire à partir de l'absorption correspondante:
3. ANALYSE FINE DE LA REVERBERATION: ECHOS ET REPONSE IMPUSIONNELLE

Des mesures en bruit rose continu, réalisées à l'aide d'un "clustering" de haut-parleurs localisés à environ 5 m au-dessus du sol, sur le mur du côté et dans la salle du Stade, nous ont donné des résultats relativement homogènes dans les hauts des sections 338, 703, 301, 372 et 338, soit un peu à gauche de la grande partie du public. Par contre, les deux mesures réalisées au centre du Stade, au sol et à 30 m de hauteur, laissent voir un maximum de réverbération dans les hautes (123 à 400 Hz), avec une valeur de 23.89 sec à 315 Hz et 30 m de hauteur. Ce modèle de réverbération préalable des toiles de couverture correspond, comme vu ci-dessus, à la réponse impulsionnelle, à un phénomène d'échos flottants entre la toile et le plancher du Stade. Ce phénomène est alors décrit par Cremer et Muller [6] pour des plafonds de différentes courbures.

Lors de des mesures impulsionnelles, nous avons employé un enregistreur conventionnel pour les niveaux de pression (généralement avec une vitesse papier de 30 mm/sec) et un analyseur double FFT modèle 2032 de "Bruel & Kjaer", pour la pression acoustique instantanée ("trigger" à -10 ms et emploi simulé des deux canaux A et B, pour une durée totale de 460 ms). Dans les deux cas la source de bruit utilisée était un pistolet de départ de calibre 38.

En ce qui concerne l'identification des échos, l'enregistrement de la réverbération au centre du Stade démontre très clairement le mécanisme des échos flottants qui s'instaure entre la toile de couverture et le plancher du Stade. L'aller et retour correspond environ à 390 ms, alors que l'enregistrement montre de 9 à 10 échos distants de 383 ms, tout au long de la décroissance. Quant à la première poignée enregistrée, elle correspond à la réflexion sur le champ choré du champ central (environ 220 ms). Cette poignée s'atténue très rapidement, mais la microsonde est éloignée du centre du terrain, sans pour autant être, au centre de la toile, ce qui produit un phénomène de doubles échos flottants du fait de l'aller et retour entre la toile et le sol, avec des durées respectives de 733 ms et 167 ms.

L'énergie émise réalisée au bas de la section 601 laisse apparaître le plus long écho distinct, avec sa réflexion sur le mur du côté et dans le parcours de 431 ms, correspondant à une durée de 1,267 ms. Les deux premières poignées correspondent au son direct à 140 ms et à sa réflexion sur le sonorisation de 250 ms, quant à l'écho du mur du côté il est suivi.
d'un retour d'énergie diffuse provenant de l'ensemble du Salle. Finalement un dernier écho est notable à 1,750 ms du premier, soit encore un second retour longitudinal complémentaire (le plus long parcours retrouvé sur plusieurs graphiques correspond dans le sens opposé à 1,867 ms).

Le dépouillement des réponses impulsionnelles laisse apparaître deux modes, l'un, plutôt longitudinal, est commandé par le comportement de la voûte, l'autre, transversal à l'ensemble des gradins, est commandé par le toit permanent du Salle (au delà des sections 600 et 700). Pour le mode de réflexion régi par le toit permanent du Salle, comme le montre un coup de feu tiré de l'anneau technique vers la toûte, on n'enregistre pas de retour immédiat d'énergie avant 500 ms. Les premiers phénomènes de réflexion que l'on peut constater lors d'un tir vers les gradins (même cheminement que le son des haut-parleurs du système de sonorisation) décrivent du comportement acoustique localisé entre le toit et les gradins.

Par exemple, dans la section 131, on peut noter deux retours de réflexion assez diffus, l'un à 212 ms et l'autre à 287 ms, ils correspondent respectivement à la double réflexion sur les gradins des niveaux 600 et 700 puis sur le toit, suivi peu après du retour de l'énergie réfléchie par le niveau 100 puis par la partie de la toiture du Salle immédiatement au-dessus (alté et retour d'environ 250 ms). Les fonctionnements distincts du toit du Salle et de la toiture de couverture sont illustrés de façon schématicque, dans les Figures no 3 et 4.

**FIGURE No.3: Illustration de l'établissement des échos**
**FIGURE No.4: Illustration de la réponse impulsionnelle obtenue sur les gradins vides.**

### 4. Analyse de l'intelligibilité selon l'indice "RASTI" et modélisation

La méthode RASTI ("Rapid Speech Transmission Index") est une méthode objective de mesure de la qualité de la transmission de la parole qui tient compte tout spécialement de l'intelligibilité. Il s'agit d'une version condensée de la procédure STI ("Speech Transmission Index") mise au point par HOUTGAST et STEENeken en 1980 [7]. Le signal de test utilisé consiste en deux octaves de bruit blanc, centrées sur les bandes de 500 Hz et 2 KHz et choisies pour être égales au niveau moyen de la voix. Les modulations de bases fréquences présentes dans la voix humaine sont simulées dans le signal RASTI par sept fréquences discrètes de modulation semblables à celles trouvées dans une voix réelle.

Les mesures RASTI consistent en une analyse du signal résultant à la position d'écoute, de façon à calculer le facteur de réduction du niveau de modulation pour chacune des sept fréquences de modulation. Si le rapport signal sur bruit devient prédominant la fonction de transfert de modulation donne une réponse plate, c'est-à-dire que toutes les fréquences de modulation sont affectées de la même manière; par contre si c'est le temps de réverbération qui domine, la fonction de transfert de modulation prêente une pente négative parce que la réverbération affecte les fréquences de modulations les plus hautes. De façon précise, la fonction de transfert de modulation MTF sur laquelle se base l'indice RASTI est calculée, pour un temps de réverbération donné T et un signal sur bruit SN, à l'aide de la relation:

\[
\text{m}(\text{SN}) = \left[ 1 - \frac{1}{(2 \pi \text{T} / 135)^2} \right] \left[ 1 + \frac{(1 + 10 \times (\text{SN})/10)}{1} \right]
\]

équation dans laquelle \(m(\text{SN})\) est la réduction du niveau de modulation pour la fréquence de modulation et le temps T de réverbération se rapporte au débit de la décroissance, telle que mesurée entre le point d'émission et le point d'écoute. Les valeurs de m obtenues peuvent être interprétées alors comme un signal sur bruit apparent et moyennées pour calculer l'indice RASTI selon les équations:

\[
X_1 = 10 \log \left( m(\text{SN})/(1-m(\text{SN})) \right) \quad \text{et} \quad \text{RASTI} = \frac{X_1 + 15}{30}
\]

La moyenne arithmétique des ces sept \(X_1\) valeurs est donc obtenue et normalisée pour fournir l'indice final situé entre 0 et 100%.

A titre d'exemple, on montrera la cartographie détaillée de l'indice RASTI relevée sur 28 sections de sièges entre les colonnes 9A et 11A. Bien entendu une telle cartographie a été réalisée pour l'ensemble du Salle avec en plus le relevé détaillé des niveaux de bruit continu émis par le système de sonorisation (à partir d'un bruit rose, pour 500 et 2000 Hz, ainsi qu'en dB(A)).
L'étude ainsi complète va permettre d'éviter les plus graves erreurs lors de la sonorisation des grands spectacles et surtout d'optimiser les travaux acoustiques projetés, travaux qui devront se répartir entre la réduction de bruit, le contrôle de la réverbération et l'amélioration du système de sonorisation.

En nous appuyant sur la modélisation de l'indice RASTI, nous avons développé un modèle informatique tridimensionnel qui tient compte de l'acoustique locale des espaces compris entre la toiture permanente du Stade et les gradins. La définition des haut-parleurs utilisés comporte les éléments suivants (pour les basses fréquences de 500 et 2000 Hz): le niveau de pression de référence à 1 m et 1 W, la puissance acoustique moyenne (soit le réglage prévu), le facteur de directivité (10° degré pour les haut-parleurs de la scène), les angles de dispersion verticale et horizontale (angle limites pour Q = 1), les coordonnées de la source, l'inclinaison de l'axe dans le plan vertical et le glissement latéral (dans un plan perpendiculaire). Initialement le niveau de bruit de fond technique devait être modélisé à la grandeur du Stade, néanmoins il s'est avéré très difficile de caractériser la directivité des nombreuses sources (76 ventilateurs de puissances variables pour le Stade seulement), aussi est-elle plutôt un système de valeurs de référence en fonction de la position dans les gradins qui a été considéré pour le calcul du rapport signal/bruit (en tenant compte également de l'absorption moléculaire aux haute fréquences). De la même manière, pour la détermination du temps de réverbération initial, après des mesures précises à l'aide de l'analyseur en temps réel crée de "Briel et Kjaer", nous avons retenu sur trois événements de référence les valeurs moyennes suivantes:

<table>
<thead>
<tr>
<th>Événement (Hz)</th>
<th>bas 100</th>
<th>bas 100</th>
<th>bas 300</th>
<th>bas 400</th>
<th>haut 400</th>
<th>haut 400</th>
<th>bas 100</th>
<th>haut 700</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 Hz (-5 dB)</td>
<td>2,37</td>
<td>2,39</td>
<td>2,15</td>
<td>2,17</td>
<td>2,77</td>
<td>2,79</td>
<td>2,25</td>
<td>1,78</td>
</tr>
<tr>
<td>500 Hz (-15 dB)</td>
<td>3,95</td>
<td>3,68</td>
<td>2,79</td>
<td>2,67</td>
<td>2,79</td>
<td>2,68</td>
<td>2,97</td>
<td>2,97</td>
</tr>
<tr>
<td>2000 Hz (-5 dB)</td>
<td>1,88</td>
<td>3,04</td>
<td>2,05</td>
<td>1,76</td>
<td>1,93</td>
<td>1,95</td>
<td>2,94</td>
<td>2,68</td>
</tr>
<tr>
<td>2000 Hz (-15 dB)</td>
<td>2,72</td>
<td>2,90</td>
<td>2,43</td>
<td>2,96</td>
<td>2,11</td>
<td>2,04</td>
<td>2,34</td>
<td>2,12</td>
</tr>
</tbody>
</table>

Ces mesures ont été obtenues à partir de compteur de fumée tiré de l'anneau technique (point de suspension pratique pour les grappes de haut-parleurs les plus importantes), il faut en effet tenir compte que le temps de réverbération initial est affecté par la localisation de la source et par sa directivité.

A l'exception des sections les plus perturbées par les bruits de ventilation, comme les gradins des niveaux 100 et 200, dans les basses fréquences (avec les canons d'air de l'anneau technique) et les gradins du niveau 700 dans la bande de 2000 Hz (avec les conduites de distribution d'air en entrée des gradins), pour la plupart des localisations, le rapport signal sur bruit affecte très peu les valeurs de MTF, donc de RASTI (évidemment, avec un bon choix de haut-parleur et sans les cris de la foule). On voit donc combien la connaissance précise du EDT prend d'importance dans la qualité du modèle. Pour les grappes de haut-parleurs se trouvant plus proches des gradins, c'est-à-dire lorsque la modélisation doit se faire à l'intérieur du rayon acoustique, il faut en plus ajouter une correction de manière à obtenir un indice RASTI plus réaliste, ceci suivant une équation de la forme:

\[
\text{EDT corrigé} = 0,3 \times \log (\text{EDT A} - 0,3) / \log r
\]

La relation dans laquelle d est la distance et r le rayon acoustique local (précise en pratique égale au chemin minimum acoustique maximum pour un point d'écoute entre deux grappes de haut-parleurs). Cette correction permet de mieux prendre en compte la directivité de la source et sa proximité, pour une vitesse initiale de décroissance de l'énergie acoustique donnée (mais le EDT local ne peut être mesuré avec les haut-parleurs proposés).

On peut se poser la question à savoir si d'octrois indices tels que le "Clarity Index" (Ciga. de REICHARDT, 1981) auraient pu être employés, avec par exemple la modélisation de l'intelligibilité proposée par BRADLEY (1981), néanmoins, la valeur du EDT nous apparaît plus facile à obtenir et à modéliser dans un contexte de réverbération assez complexe comme celui du Stade olympique, puisqu'elle est liée à l'énergie totale des autres paramètres. D'autre part, les mesures RASTI réalisées avec le système de sonorisation actuel ont prouvé une juste harmonie entre le rapport signal/bruit et l'influence de la réverbération; de même, la cartographie globale du Stade suit minutieusement les nuances perçues mentionnées lors des tests d'audition (et aussi les nombreux commentaires formulés après douze années d'exploitation, tant par les spectateurs que par les spécialistes).

5. REFERENCES

HALL AVERAGE CHARACTERISTICS OF 10 HALLS

J.S. Bradley, Institute for Research in Construction, National Research Council, Ottawa, Canada, K1A 0R6

Introduction

There is a need for a large data base of modern acoustics measurements in a variety of concert halls to aid in interpreting these measurements. To this end, measurements were made in 7 European halls in collaboration with European colleagues to obtain data in well known halls and to permit comparisons between researchers. The present paper is an initial examination of results in these plus 3 Canadian halls with a focus on the variations of particular quantities within halls.

Measurements were made using the RAMSoft computer program in conjunction with a Norwegian Electronics real time analyser. Twelve different acoustical quantities are calculated in each of 6 octave bands, and each measurement result is obtained in seconds in-situ in the hall. Only five quantities are presented in this paper. Both the early decay time, EDT, and the reverberation time, RT, are calculated from straight line fits to either the first 10 or 30 dB of the integrated impulse response decay curve. CR0 is an early to late arriving sound energy ratio in dB with an 80 ms early time limit, related to clarity or the balance between clarity and reverberance. The overall strength or normalized level, G, is the ratio of the total level to the level of the same source in an anechoic field at 10 m. The lateral energy fraction, LF, is the linear fraction of the early arriving energy from lateral directions, within the first 80 ms. Only mid-frequency results, that are averages of the 500 and 1000 Hz octave results, are presented in this initial study.

Within-Hall Variations

The standard deviation, STD, of each parameter about the hall mean value is a simple indication of spatial variation. STD values are presented for CR, EDT, and G values in Figures 1 to 3. In each case the order of the halls along the horizontal axis has been sorted to place the halls with the largest spatial variations to the right. The three-letter codes for each hall are explained in Table 1. It is assumed that large within hall variations are not desirable and are not indicative of an even diffuse sound field. In some respects halls that appear to the right on these figures are therefore less satisfactory. While a number of halls seem to appear towards the right on all of these figures there are exceptions and one hall, SWP, appears at the left of Figure 1, but at the right of Figure 3.

| ODN Odense Koncerthuus | VNA Vienna Musikvereinsaal |
| AMS Amsterdam Concertgebouw | PLY Paris Salle Pleyel |
| STU Stuttgart Liederhalle | RTH Toronto Roy Thomson Hall |
| SLZ Salzburg Neue Festspielhaus | NAC Ottawa National Arts Centre |
| MUN Munich Gasteig Philharmonic | SWP Montréal Salle Wilfrid Pelletier |

Table 1. Titles and locations of the 10 halls.

Several other quantities can be considered as measures of diffusion or the evenness of the sound fields. In a previous paper2 we calculated the difference between measured CR0 values and calculated values assuming an ideal exponential decay with the measured RT. Such differences are shown for the present 10 halls in Figure 4. The halls to the right of this figure, where the differences exceed +1 dB, can be considered to have a more directed or less diffuse early sound field causing them to have higher CR0 values than might be expected from their reverberation times.

Figure 5 plots the differences between hall average RT and EDT values indicative of the linearity of the decays and the diffusion of the hall. For this figure, 7 of the halls, including some famous for their good acoustics, are all quite close to zero mean difference. There are then 3 halls at the extreme right of this figure that are clearly different from the others in having non-linear decays. These same three halls are to the right of Figure 4, but the order is not the same. Thus having EDT different to RT is related to having unexpectedly high CR0 values, but there is not a simple one to one relationship.
The slope of the G values, in dB per 10 m, are presented in Figure 6. In an ideal diffuse field reverberant sound levels are not expected to vary with distance. Barron has found that in British concert halls levels do vary with source-receiver distance, and that levels tend to decrease more rapidly with distance in halls with diffusing ceilings. The present data contains some more extreme cases than in Barron's data and the importance of the ceiling on the slope of G values becomes more clear. Initially it was thought that the slope of G values was related to the mean height of the ceiling. However in hall ODN with a relatively low ceiling, G values decreased most rapidly with distance. All four halls to the right of this figure have ceilings that cause energy to be scattered back towards the stage end of the hall. This may be due to the shape of the ceiling section, hall ODN, or to objects suspended below the reflecting ceiling as in halls NAC, SWP, and RTH.

The 5 halls labelled 'intermediate' on Figure 6 all have hard reflecting ceilings with no scattering material hung below the reflecting ceilings. Many of these halls have approximately flat ceilings. They include halls famous for their acoustics, and have G slopes of approximately 1 dB per 10 m. Thus even good halls do not agree with simple diffuse field theory.

Hall SLZ to the left of this figure has a reflecting ceiling that in section is shaped to direct energy towards the rear of the hall. As a result, the G values at the rear of the hall are approximately the same as near the stage. The ceilings of the Alberta Jubilee auditoria produce this same effect.

Figure 7 compares the hall mean slopes of EDT values. One hall, VNA, is unusual in that EDT values increase a little towards the rear of the hall. Most halls have slopes of EDT values with distance that are very close to zero. Three halls to the right of this figure show quite large decreases in EDT values with distance. In hall SLZ the strong early arriving ceiling reflections directed to the rear of the hall cause more rapid initial decays towards the rear of the hall has we have observed in the Alberta Jubilee auditoria. The large EDT slopes for halls Ply and NAC are again thought to be indicative of a lack of diffuse conditions in these halls.

Discussion

There is considerable overlap in the rank ordering of the 10 halls in terms of the various indicators of spatial variation and non-diffuse conditions. In particular, halls VNA and AMS which are generally assumed to have good acoustics, have very even diffuse conditions from these measurements. Thus these parameters seem to be important in evaluating the overall acoustical quality of a hall.

The systematic variation of some parameters from the front to the rear of halls is seen to explain a significant part of the overall spatial STD values in Figures 1 to 3. The ceiling appears to be particularly important, because it is usually the largest reflecting surface. Where there is material below the reflective ceiling, sound energy will be scattered back towards the source and sound levels at the rear of the hall will be 2 to 3 dB lower than in halls without this back-scattering material. Where ceiling sections are shaped to direct sound energy to the rear of the hall, sound levels will not vary much with distance but EDT values can be considerably decreased at rear seats.

Lateral Energy

Higher values of LF are normally expected to occur in narrower halls. Indeed, Gade found a linear relationship between LF values and the mean width of 21 Danish halls. Figure 8, where the present results are compared with Gade's earlier regression line, indicates that this same trend was not found for the present halls. Hall VNA which has been supposed to be an example of a hall with high spatial impression has no larger LF values than other halls. This was examined further by considering early lateral and non-lateral sound levels separately. Figure 9 shows the mean lateral energies versus width and there is an apparent relationship with the mean hall widths. The slope of the regression line shown on Figure 9 is similar to that of Gade's regression line shown of Figure 8. Thus in the present data the lateral component of the energy is related to the width of the halls. However, the early non-lateral energy was also found to decrease with increasing mean width. Thus the ratio of these, LF, does not vary greatly with width. It is interesting to note that the two halls furthest below the regression line of Figure 9, have audience seating in
boxes on the side walls. These boxes prevent strong side wall reflections to the rear of the hall and would cause some energy to scattered back towards the front of the hall similar to some ceilings.

One is tempted to speculate that the early lateral level is more important than LF values. The present measurements do not support the popular belief that halls such as VNA and AMS have particularly high LF values, but they do have the highest early lateral sound levels. In a recent subjective study, Barron found the early lateral level to be more strongly correlated to spatial impression judgments than LF values. We know that spatial impression is influenced by both the portion of lateral energy and the overall level. Perhaps it is the early lateral level that incorporates both the lateral and energy level aspects into one quantity that is more closely related to spatial impression judgments.

Conclusions
Evidence has been presented here that better halls have smaller within hall variations of principal acoustical quantities and therefore tend to have more even diffuse conditions throughout the hall. The details of the ceiling and to a lesser extent of the side walls can cause systematic variations of parameters with distance from the source. Measurements of the early lateral energy suggest that further subjective studies are necessary to verify whether it is the absolute or the relative level of early lateral reflections that is important.

References

Figure 1: Spatial standard deviations of C60 values by hall.

Figure 2: Spatial standard deviations of EDT values by hall.

Figure 3: Spatial standard deviations of G values by hall.
Figure 4. Measured - calculated C80 differences by hall.

Figure 5. Differences of RT-EDT values by hall.

Figure 6. Slope of G values versus source-receiver distance.  Figure 7. Slope of EDT values versus source-receiver distance.

Figure 8. Hall average LF values versus mean width.

Figure 9. Hall average early lateral levels versus mean width.
MONITOR-ROOM ACOUSTICS AND ITS ANALYSIS USING HEAD AND TORSO SIMULATOR

K. Okabe, T. Mizuuchi, H. Hamada and T. Miura

Tokyo Denki University
2-2, Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101, JAPAN

1. Introduction

Since 1933, a dummy head microphone (DHM) technique has been powerful tool to record sound fields and reproduce them to listeners, although it has defects such as the incorrect reproduction of the frontal sound events (so-called "front-back reversal") [1] or incorrect localization in median plane when using headphone reproduction. Even in case of loudspeaker reproduction by Schroeder's pre-filtering scheme [2], these phenomena sometimes arise. These defects, however, can be improved by a modified scheme [3]. Using recent dummy head or head and torso simulator (HATS) technique, we can reproduce the original sound field with high fidelity in aspects of localization, tone color and spatial impression.

Another possibility in using a DHM or a HATS is facility as acoustical measurement tool and the interpretation of the measured results from subjective point of view.

In this article, we report the acoustic measurement results of a HATS in room acoustics. We measured binaural impulse responses in the same monitor-room before and after making acoustical alterations. The room was modified according to the recent trends in which the reflections from the front of the room are eliminated and the strong and diffused reflections from rest of the room are introduced. The effects of this approach called "LEDE:Live-End-Dead-End" [4] were measured by a HATS. The head part of the HATS used in the measurements was constructed by CAD (Computer Aided Design) system and NC (Numerical Controlled) cutting machine. It has average dimensions of Japanese male adults and complete symmetrical configurations.

2. Theory

The measured impulse response \( h(t) \) was processed using analytic signal theory. We can calculate the analytic signal \( \tilde{h}(t) \) of the impulse response \( h(t) \)

\[
\tilde{h}(t) = h(t) + j \tilde{h}(t)
\]

where

\[
\tilde{h}(\tau) = \mathcal{H} \{ h(t) \} = \frac{1}{\pi} \int_{-\infty}^{\infty} h(\tau') (t-\tau') \, d\tau'
\]

203
The imaginary part $\hat{h}(t)$ corresponds to Hilbert transform of impulse response $h(t)$. Energy-time curve (ETC) i.e. Amplitude of the analytic signal $h(t)$ was calculated as function of time. This approach called time delay spectrometry (TDS) is relatively new method of measurements for linear time invariant system and it illustrates the time response of the system. The method provides better understanding of time domain phenomena such as reflections behavior in rooms.

3. Measurements

The monitor-room to be measured has dimensions of about 7m in width and length (Fig. 1). The alteration points before and after the reconstruction were as follow:

1. loudspeaker system
2. reflective front wall to absorptive one
3. absorptive rear wall to reflective one
4. diffusive treatment of rear wall

Fig. 1 Configuration in floor plane of the monitor room and measuring point.
The HATS was set at the mixing engineer's position, 2.7m from loudspeakers, and its ear position was 1.2m above the floor. Electrical impulse with duration of 10 micro sec was fed into loudspeakers and resulting response was picked up by the microphones of the HATS. The impulse responses were averaged in the micro-computer to increase signal to noise ratio. The HATS was free field equalized, i.e. when a sound source locates in front of the HATS in free field, its transfer function from the source to the output of the HATS has flat spectrum.

The same measuring procedure were repeated before and after the acoustical alteration of the monitor-room. The calculated FTC are shown in Fig.2, both of them were measured by the HATS's left ear.

Fig. 2 Energy-time curves in the monitor room (a)before and (b)after acoustical alterations. Both of them were measured by the left ear of the HATS.
4. Discussion

Two ETCs clearly show the effects of the alteration as follows. Initial parts of the ETC indicate phase linearity of the loudspeaker system. New system shows good linearity which results in sharp attack and decay in the initial part of the ETC, on the other hand old loudspeaker system shows the dissipated initial energy response that is caused by the different group delay distortion of the each loudspeaker unit.

Lateral reflections in the new monitor room are suppressed relative to those in the former room. Instead of this, reflections from rear wall are prominent in the new room. They are appearing around 20 ms later from the direct sound and are reported to be subjectively important for good monitor room acoustics [4].

These features in the ETCs can be correlated with subjective assessments for the new monitor room.

In this article, we concerned with monaural aspects of the measured impulse response by the HATS. Next step to be come up is analysis of the binaural impulse responses of the HATS. Binaural model approach [5], especially including contra-lateral suppression [6], will provide interpretations from physical characteristics of the rooms to subjective features such as localization and spatial impression which are also very important for monitor room acoustics.

Acknowledgment

We are grateful to Mr. T. Anazawa of Nippon Columbia Co. Ltd., Mr. T. Shigeta and Mr. Y. Sakiyama of Nittobo Acoustic Engineering Co. Ltd. for their kind supports to carry out the measurements.

Reference

ACOUSTIC IMPROVEMENT OF CLASSROOMS MEASURED WITH THE RASTI-METHOD
E. Pekkarinen a,b, E. Mänttä a, C. V. Viljanen a, H. Koskela a and J. Honka a
a Turku Regional Institute of Occupational Health, Hämeenkatu 10, SF-20500 Turku, Finland
b Department of Otorhinolaryngology, Helsinki University Hospital, SF-00290 Helsinki, Finland
c Department of Physical Sciences, University of Turku, SF-20500 Turku, Finland

Introduction
The most important information channel in teaching work and thus in classrooms is speaking and hearing. Therefore acoustic design of classrooms should be based on achieving the highest possible degree of speech intelligibility for all the pupils. Classrooms are, however, noisy (1) and reverberant (2) environments, for speech communication. Noise and reverberation cause smoothing effect on the temporal envelope of the speech signal and thus decrease speech intelligibility (3,4).

In the classroom the distance between teacher and pupils may be several meters. An acceptable speech transmission between them need good reverberation conditions and low noise levels. The typical shape of a classroom is rectangle and the surfaces are usually hard. One possibility to affect reverberation conditions is to cover ceiling and walls with sound absorbing material, i.e., mineral wool. The amount and the location of the material may affect the speech transmission properties of the room. The reverberation time measurement does not measure speech intelligibility and the psychoacoustic method of speech recognition do not show minor changes in speech transmission between various location in a room. Recently, a new objective measure, the speech transmission index (STI) has been proposed. It is based on modulation transfer functions and it takes into account the effects of interfering reverberation and background noise (5). If a reduced number of octave bands and modulation frequencies are used in the measuring procedure the method is Rapid Speech Transmission Index (RASTI).

The aim of this study was to find out good acoustic conditions for speech intelligibility, measured with the RASTI-method, using minimum amount and proper placement of mineral wool in a quiet test classroom.

Materials and Methods
The classroom was built in an acoustic laboratory. The dimensions were 7.6 x 10 m² with the height of 3.3 m. The reverberation time of the empty classroom was reduced with 50 mm thick mineral wool panels. In the first experiment the back wall was covered with increasing amount of absorption material and the walls were kept hard. In the second phase mineral wool was added on the ceiling alone and the other surfaces were without absorption. Finally the absorption material was added both on the back wall and on the ceiling.

The background noise levels were measured with the B&K Sound Intensity Analyser 4433, while the RASTI-measurements were conducted with the B&K Speech Transmission Meter 3361. The signal transmitter was placed in front of the hard wall (7.6 m) and the receiver in
eight different locations of the classroom (Fig 1). The RASTI-values presented are means of three successive measurements and the measurement time for each value was 16 s. The subjective intelligibility scale used is: excellent (1.00-0.75), good (0.75-0.60), fair (0.60-0.45), poor (0.45-0.30) and bad (0.30-0) (6).

Fig. 1. The measurement situation in the test classroom. T is the location of the transmitter (B&K 4225) and the locations of the receivers (B&K 4419) are numbered from 1 to 8.

Results
The background noise levels did not exceed 35 dB(A). The RASTI-value for the empty room with hard walls and ceiling was mean 0.38 (poor). The addition of absorption on the back wall increased the RASTI-value up to 0.61 (good), when ceiling and other walls were hard. The addition of mineral wool only on the ceiling resulted in the RASTI-value 0.60 (good) at its best. The only way to get RASTI-value higher than 0.75 (excellent) was to use absorption material at least on two different surfaces, e.g., on the ceiling and on the back wall. If 100% coating of the ceiling and back wall was used, RASTI-value of 0.82 was reached. The results in more details will be shown during the presentation.

Conclusions
The use of mineral wool on one surface was not enough to reach excellent acoustic environment, the use of mineral wool on two surfaces was necessary. The total quantity needed was about 30% of the surface area of all the walls and the ceiling. The RASTI-method was sensitive to show differences in intelligibility also when the location of the mineral wool was changed but the quantity of it was kept constant.

References
MESSUNGEN DER SCHALLAUSBREITUNG IN FLACHEN RÄUMEN

G. Elbert

Institut für Technische Akustik
Rheinisch-Westfälische Technische Hochschule Aachen
Templergraben 55, 5100 Aachen, Bundesrepublik Deutschland

1. Einleitung


2. Beschreibung der Halle

Die Halle weist einen Grundriss von 56 x 29 m² auf. Ihre Grundfläche beträgt 1597 m², es ist an einer Längsseite des Raumes ein kleines Treppenhaus abgebaut. Die Halle verfügt über eine von Pfeilern und Unterzügen gestützte Betondecke in 3,80 m Höhe. Der Boden besteht aus Beton, die Außenwände aus verputztem Mauerwerk mit Isolierverglasung.

Die Halle war nach ihrer Errichtung, bis auf geringe Mengen Baumaterial, zunächst leer. Sie ist dann mit Maschinen und Material zur Produktion von Schuhen ausgerüstet worden. Bei den Maschinen handelt es sich vornehmlich um Pressen und Stanzen der Abmessungen von etwa 3 x 2 x 2 m³, das Material besteht aus Leder und Kunststoffteilen.

3. Messungen

Die Messungen erfolgten an zwei verschiedenen Terminen, zuerst in der leeren und dann in der eingerichteten Halle. Es wurde der stationäre Schalldruckpegel in verschiedenen Abständen von der Schallquelle, sowie die Nachhallverläufe in Oktavbändern gemessen. Als Messsignal wurde rosa Rauschen verwendet. Zur Erzielung einer möglichst ungerichteten Abstrahlung diente ein Codedakoderlautsprecher. Die Höhe von Lautsprecher und Mikrofon war jeweils gleich, sie betrug in der leeren Halle 1,20 m und in der eingerichteten Halle 1,50 m. Es sind für die Messungen in den beiden Einrichtungszuständen der Halle möglichst die gleichen oder, falls dies aufgrund im Weg stehender Einrichtungsgegenstände nicht möglich war, zumindest ähnliche Meßpfade benutzt worden. Hier wird sich jedoch auf die Darstellung der Ergebnisse, die auf einem Meßpfad parallel zu den Seitenwänden in der Mitte des Raumes gewonnen wurden, beschränkt. Dar gestellt wird der auf die Leistung der Schallquelle bezogene Schallpegel:
L = L_p - L_w  mit  L_p : gemessener Schalldruckpegel  
L_w : Leistung des Lautsprechers 

Die Leistung des Lautsprechers wurde im Hallraum in den verwendeten Oktavbändern bestimmt. In den Abbildungen 3.1 und 3.2 sind die in den beiden Ausrüstungszuständen der Halle gemessenen Schallpegel über den auf die Rauminhöhe h bezogenen Meßebstand r aufgetragen.

Abb.3.1: leere Halle  
Abb.3.2: eingerichtete Halle 

Es ist in den Abbildungen ein deutlich stärkerer Abfall des stationären Schallpegels der eingerichteten Halle zu erkennen.

4. Vergleich mit berechneten Schallpegeln 

4.1 Ray-Tracing


zwischen 0,6 und 0,8 beträgt. Bei der Berechnung der leeren Halle wurde ein Diffusitätsgrad von 0,7 für die strukturierte Decke (Unterzüge), und ein solcher von 0 für alle übrigen Flächen angenommen. In der eingerichteten Halle betrug der Diffusitätsgrad 0,7 für alle Flächen. In den Abbildungen 4.1 und 4.2 lässt sich sowohl für die leere als auch die eingerichtete Halle eine gute Übereinstimmung von gemessenen und gerechneten Schallpegeln erkennen.

Abb. 4.1: leere Halle

Abb. 4.2: eingerichtete Halle

4.2 Durch Spiegelquellen erweitertes Flachraummodell nach Kuttruff [3]


\[
u_d(r) = \frac{P}{\pi c h^2} \left( \frac{\rho_a}{1 + \frac{r^2}{h^2}} \right) + \frac{\rho_g}{1 - \rho_g} \left( \frac{b}{b^2 + \frac{r^2}{h^2}} \right)
\]

mit 
- \( u_d(r) \): Energiedichte des gestreuten Schallfeldes
- \( P \): Leistung der Quelle
- \( h \): Raumhöhe
- \( r \): Abstand zwischen Sender und Empfänger
- \( \rho \): Reflexionsgrad (\( \rho = 1 - \alpha \) mit \( \alpha \): Absorptionsgrad)
- \( \rho_a, \rho_g \): arithmetischer bzw. geometrischer Reflexionsgrad von Decke und Boden


In den Berechnungen ist das kleine Treppenhaus an einer Seitenwand der Halle vernachlässigt worden. Berechnet wurde also ein Rechteck mit den Abmessungen des Raumes. Aus der Abbildung 4.4 ist wiederum eine gute Übereinstimmung der gemessenen mit den berechneten Schallpegeln zu er-
kennen. Die Abweichung bei der kürzesten Meßentfernung ist letztlich auf Mängel im Abstrahlverhalten des Dodekaeders zurückzuführen. Die Übereinstimmung der Pegel bei der leeren Halle in Abbildung 4.3 ist im Mittel zwar ebenfalls nicht schlecht, die errechneten Verläufe weichen jedoch offensichtlich von den gemessenen ab. Als Ursache muß die mit diesem Verfahren verbundene Annahme vollständig diffuser Streuung von Decke und Boden angenommen werden, was ja für die leere Halle sicher nicht zutreffend ist.

![Graphik zur Darstellung der gemessenen und errechneten Pegelverläufe für leere Halle und eingerichtete Halle.]

**Abb. 4.3** : leere Halle  
**Abb. 4.4** : eingerichtete Halle

5. Zusammenfassung und Ausblick


Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft dankenswerterweise durch eine Sachbeihilfe unterstützt.

6. Literatur

/1/ Vorländer, M. Ein Strahlverfolgungsverfahren zur Berechnung von Schallfeldern in Räumen, Acustica 69 [1988], 138

/2/ Vorländer, M. Messung von Schallstreuung an rauhen Wänden Fortschritte der Akustik DAGA 88, 761

/3/ Kuttruff, H. Stationäre Schallausbreitung in Flachräumen Acustica [1985], 62

212
3. D
AUDITORY SPACIOUSNESS AND ENVELOPEMENT

M. Morimoto, Z. Mackawa

Environmental Acoustics Laboratory, Faculty of Engineering, Kobo University, Nada, Kobe 657 Japan

1. INTRODUCTION

The spatial attribute of an auditory event is an important factor in the subjective evaluation of concert halls. The spatial attribute comprises the characteristics of position and broadening of an auditory event. It is well known that broadening is closely related to good acoustics in concert halls. In the past, many different terms were used to describe broadening as listed Blauert[1]. However, most of them were not distinctly defined. Therefore, it was not clear whether or not, all of them describe a characteristic of broadening of an auditory event. The authors hypothesize that broadening comprises two characteristics at least. One is auditory spaciousness and another is feeling of envelopment. Figure 1 illustrates the concept of two characteristics of broadening. Spaciousness is defined as the width of an auditory event perceived temporally and spatially to be fused with the auditory event of a direct sound. Envelopement is defined as the fullness of auditory event around a listener, excluding auditory event relating to spaciousness.

In this study, we performed two experiments, in which we varied respectively the interaural cross-correlation coefficient (IACC) of the early part (direct sound + early reflections) and of the reverberation part of a sound field, as a first trial, to answer to a fundamental question: "Can a listener discriminate between spaciousness and envelopment?"

2. SOUND FIELDS USED IN EXPERIMENTS

The source signal used for the experiments was a solo performance of violin recorded in an anechoic chamber. The musical motif was a 14s section (bars 7-12) of "Introduction et Rondo Capriccioso" by Saint-Saens.

Figure 2 shows the signal configuration of all sound fields used in experiments. They consist of a direct sound, two discrete early reflections and three reverberations. Time delays and levels of early reflections and reverberations were set, considering echo disturbance, coloration, split of auditory event and excess reverberation. The reverberation time was constant at 2.0s and its frequency characteristic was flat. The IACC was adjusted

Fig.1 Concept of spaciousness and envelopement.

//: spaciousness
////////////: envelopment

215
by controlling the direction of loudspeakers radiating early reflections and reverberations. Figure 3 shows the arrangement of loudspeakers.

Nine kinds of sound fields used in experiments are listed in Table 1. IACC of sound fields were measured by using KEMAR dummy head without ear simulators(HK, DB-100). The level of sound field was 78.8-79.6dBA RMS SLOW PEAK measured at the left ear of KEMAR mentioned above. The energy ratio of the early part to the reverberation part was 0.44-0.47 in D value.

![Fig.2 Signal configuration of sound field](image)

![Fig.3 Arrangement of loudspeakers](image)

<table>
<thead>
<tr>
<th>Sound Field</th>
<th>Early Part</th>
<th>Reverberation Part</th>
<th>Whole Part</th>
<th>Loudspeaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td>L1, R1</td>
</tr>
<tr>
<td>2</td>
<td>0.90</td>
<td>0.91</td>
<td>0.79</td>
<td>L1, R1</td>
</tr>
<tr>
<td>3</td>
<td>0.90</td>
<td>0.51</td>
<td>0.66</td>
<td>L1, R3</td>
</tr>
<tr>
<td>4</td>
<td>0.69</td>
<td>0.91</td>
<td>0.82</td>
<td>L2, R2</td>
</tr>
<tr>
<td>5</td>
<td>0.69</td>
<td>0.71</td>
<td>0.71</td>
<td>L2, R3</td>
</tr>
<tr>
<td>6</td>
<td>0.69</td>
<td>0.51</td>
<td>0.58</td>
<td>L3, R1</td>
</tr>
<tr>
<td>7</td>
<td>0.47</td>
<td>0.91</td>
<td>0.71</td>
<td>L3, R3</td>
</tr>
<tr>
<td>8</td>
<td>0.47</td>
<td>0.71</td>
<td>0.62</td>
<td>L3, R2</td>
</tr>
<tr>
<td>9</td>
<td>0.47</td>
<td>0.51</td>
<td>0.50</td>
<td>L3, R3</td>
</tr>
</tbody>
</table>

*ER: Early reflection, REV: Reverberation*
3. DISSIMILARITY JUDGEMENT CONCERNING SPATIAL ATTRIBUTE

In the first experiment, the psychological configuration governing spatial attribute of an auditory event was investigated. The paired comparison test was carried out. The interval between two sound fields was 2s. All pairs were arranged in random order and separated by an interval of 5s.

The subject was requested to judge the degree of dissimilarity concerning spatial attribute of auditory event for all pairs on a five-point scale from "the same" to "extremely different." Each subject was twice tested, while seated, with head fixed in a partially darkened anechoic chamber. Before the test, the subject was instructed the concepts of spaciousness and envelopement, with Fig.1. Eight male students were served as subjects.

Responses of four subjects, which were regarded as being reliable and reproducible, were analyzed by using Kruskal's MDS. The analysis indicated that the psychological configuration governing spatial attribute of auditory event was two-dimensional (stress=0.078).

4. JUDGEMENT OF SIX CHARACTERISTICS OF SPATIAL ATTRIBUTE

In the second experiment, six characteristics of spatial attribute of an auditory event for nine sound fields used in the first experiment were investigated by using the paired comparison test, in order to interpret the psychological configuration obtained in the first experiment. They were spaciousness, envelopement, depth, distance, distinction of contour, and volume of listening room. Four subjects whose responses were reproducible and reliable in the first test served as subjects. Each subject judged twice each characteristic separately. The other conditions were the same as in the first test.

The psychological scales of six characteristics for nine sound fields were obtained from all responses of four subjects by using the Thurston Case V model. Here, let us observe "spaciousness" and "envelopement" which are main subjects in this paper. Table 2 shows the correlation coefficients between the two characteristics and three kinds of IACC. Spaciousness has a highly negative correlation with IACC of the whole sound field (early part + reverberation part). But, envelopement has a highly negative correlation with IACC of not only the whole sound field, but also of the reverberation part.

<table>
<thead>
<tr>
<th></th>
<th>Early Part</th>
<th>Reverberation Part</th>
<th>Whole Part</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spaciousness</td>
<td>-0.59</td>
<td>-0.54</td>
<td>-0.82**</td>
</tr>
<tr>
<td>Envelopement</td>
<td>-0.12</td>
<td>-0.86**</td>
<td>-0.78*</td>
</tr>
</tbody>
</table>

** Highly significant (p<0.01), * Significant (p<0.05)
5. DISCUSSION: INTERPRETATION OF THE PSYCHOLOGICAL CONFIGURATION

The psychological configuration obtained in the first experiment was interpreted by multiple regression analysis. In analyzing, a dependent variable was the psychological scale value of each characteristic for a sound field obtained in the 2nd test and an independent variable was the coordinate of a sound field obtained in the 1st experiment. Figure 4 shows the psychological configuration of nine sound fields and axes of six characteristics. Physical factors, three kinds of IACC, were indicated together in the figure. The direction of each arrow indicates "more spaciousness", "more envelopement" and so on.

Here, let us observe "spaciousness" and "envelopment", again. Two axes of them intersect at an angle of 45 degree. This means that they are not perfectly independent each other. However, let us observe the relation between them and IACC. The axis of spaciousness is almost in the opposite to that of IACC of the whole sound field. This coincides with the result of the second experiment. Namely, spaciousness depends upon IACC of the whole sound field. This agrees with the previous result[21]. On the other hand, the axis of envelopement is almost in the opposite to that of IACC of the reverberation part. This indicates that envelopement depends more strongly upon IACC of the reverberation part rather than that of the whole sound field, though envelopement has a highly negative correlation with IACC of both of the whole sound field and the reverberation part in the second experiment.

Summarizing, the results of two experiments seem to suggest that a listener can discriminate between spaciousness and envelopement. Spaciousness and envelopement depend upon IACC of the whole part and the reverberation part of a sound field, respectively, under the experimental conditions in this study.

REFERENCES


---

Fig.4 Psychological configuration of sound fields and axes of six characteristics of spatial attribute and of physical factors. Numeric indicates the kind of sound field listed in Table 1. IACC 1, 2 and 3 indicates IACC of the whole, the early part and the reverberation part of sound field, respectively.
AN EXPERIMENTAL SYSTEM
FOR KNOWLEDGE-BASED ARCHITECTURAL ACOUSTICS

Masuzo YANAGIDA
Communications Research Laboratory
4-2-1 Nukui-Kitamachi, Koganei, Tokyo, 184 Japan
Takayuki HIDAKA
Takenaka Technical Research Laboratory
2-5-14 Minami-suna, Koto-ku. Tokyo. 136 Japan
and Takahira YAMAGUCHI
ISIR, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, 567 Japan

INTRODUCTION

Several indexes to evaluate sound fields have been proposed so far, particularly for concert halls[1,2], and the interrelation between those indexes and subjective measures has been discussed enthusiastically[3]. However, acoustic design is still carried out based on designers' intuition and experiences. As design by intuition and experiences is in most cases very inefficient and gives results close to the optimum, there could be some intelligent supporting system for design engineers, provided that there can be found means to express, store and manage the domain knowledge together with a tool to simulate sound fields. This paper describes an expert system named ARDEX (Architectural acoustics Design EXPert)[4], a cooperative system composed of an expert subsystem and a field simulator.

ARDEX

Brief Description

ARDEX is designed to be used as a tool for room acoustic designing and for studying room acoustics. ARDEX is an expert system of a synthesis type regarded most difficult to be constructed. At present, the basic conceptual design of the system is almost fixed and the initial configuration generator and a part of the field simulator are implemented on a mini computer.

Conceptual Design

The block diagram of the system is depicted in Fig.1. The system is composed of two major subsystems, an expert subsystem (ES) and a field simulator (FS). The ES is composed of an inference engine, a scheduler, user interface, knowledge bases to get initial configuration, to eliminate acoustical defects and to modify room configuration and a working memory to hold the current configuration with its score evaluated with certain evaluation measures.

The field simulator is designed to support both the ray-tracing method and the image method. The ray-tracing method (Field Simulator I) is used for detection of global acoustic defects, for allocation of listening points for field evaluation, and also for preprocessing for the image method to speed up the calculation. The image method (Field Simulator II) is the tool to obtain the scores by the evaluation measures.

At present, the initial design strategy in the ES and the image method in the FS are implemented on Data General MV/7800XP by Prolog and FORTRAN, respectively.

The Design Strategy in ARDEX

The design strategy takes a form of optimization with a multiple loop to obtain a satisfactory room configuration though it is not guaranteed to converge.

1)Discove fatal acoustic defects such as acoustic shadows, concentration of sound, reflection with long delay time etc.

2)Improve evaluation scores at each specified listening point.
3) Improve the integrated evaluation score for the specified set of source-listening point pairs.

4) Search for other possibilities to get out of local optima.

Features of the Proposed System

a) Type of Inference in the Expert Subsystem:
Forward reasoning of the type "if [conditions], then [execute]" is employed for generation of initial configuration and also for the modification process to improve the evaluation score that represents acoustic characteristics of the room.

b) Optimization Strategies:
The process is controlled by a scheduler referring to meta knowledge of optimization strategies. After eliminating acoustical defects, the proposed system enters into a multiple loop to modify the room configuration to make the evaluation score satisfy some basic requirements and to optimize the evaluation score for the specified listening points, and then for the specified source positions.

c) Field Simulator:
The proposed system requires a means to evaluate the current design and an inference engine together with a knowledge base. The ES is closely connected to a field simulator, a tool for evaluation of the acoustical field obtained by definite change on the current room configuration.

d) Uncertainty in Evaluation Measures:
Several evaluation measures for room acoustics have been proposed so far, but the necessary and sufficient conditions for desiring good halls under specified conditions are not known. The proposed system uses some of those measures merely as the tentative candidates for numerical evaluation of the current design stored in the working memory. As the system has the difficulty that it has no other way but use unreliable evaluation measures, the proposed system has the facility of listening to the simulation output to evaluate the validity of the evaluation measures employed.

e) Design Strategy and Evaluation Measures:
Design strategy and evaluation measures are decided by user specification, or by the default procedure. The design strategy is controlled according to the following factors:

(1) Evaluation Scores:
X_n : n-th evaluation measure.
Y_1 : integrated evaluation score at listening position R_1.
Z : total evaluation of the room.

(2) Weights
b_n : weight for X_n;
w_1 : weight for Y_1 to evaluate whole the room.

The evaluation measures take the form
X_n = a_n | x_n |^{3/2}
where x_n denotes n-th evaluation measure and the weights a_n(n=1,...,N) can be a function of x_n.

The following are tentatively employed as the evaluation measures keeping possibility to be replaced by more reasonable ones if found.
1. Sound pressure level
2. Time delay of the first reflection wave
3. Reverberation time at 500 Hz
4. Interaural Cross Correlation (IACC) [2]

The evaluation score for a listening point would be a weighted sum of the scores of evaluation measures, and the total evaluation for a room would be also a weighted sum of evaluation scores of representative seats.
f) The Rule Base

(1) Examples of rules for deciding initial configuration
   If usage = music, then suggested volume/seat = 7.8 m³.
   If principal program = Baroque, then $T_e = 35$ ms.
   (2) An example of rules to dissolve fatal acoustic defects
       If (incidental energy)/(unit area) > Avp,
       then use less reflective material for the wall on which the ray
       or change the angle of the wall,
       or increase the distance between source and the wall.
       (3) An example of rules to improve an evaluation score
       If (time delay of the first reflection) > $T_e$, then move the reflection wall toward inside by
       $d = c \times (T_1 - T_e)/2$,
       or create a new wall to make $T_1 = T_e$.
       Where $c$ denotes the sound velocity.
       (4) An example of rules to modify evaluation measures
       If evaluation score $X_n$ increases after modification,
       and perceptual score does not increase,
       then reduce the weight for that measure
       and send a message to the user.

PERFORMANCE EXAMPLES OF FIELD SIMULATOR II

Image method was used to a room depicted in Fig. 2 assuming a source
position denoted as S and a listening point R. The sound rays which reach
at R directly or via one reflection are also drawn. The pattern of
propagation delay and decay of each path from S to R can be drawn on the
screen and the impulse response from S to both ears of the listener
seated at R can be calculated. Given a source sound, the system produces
the sounds which the listener seated in R should hear by convolving the
input source sound and the impulse response. Design process goes either
with auditory feedback or completely automatically.

CONCLUSIONS

A knowledge based system to support acoustic design is described.
Although the conceptual design of the total system is almost fixed, the
system is under development completing only the initial configuration
generator subsystem and a prototype of Field Simulator II. Implementation
of the Expert Subsystem is scheduled after refining the Field Simulator
together with developing an effective interface between Field Simulator
and Expert Subsystem.

Although the domain knowledge is unreliable and unclear at present,
the proposed system is expected to contribute not only to design process
itself but also to basic research to make the domain knowledge in aural
perception rich.

Acknowledgement The authors express their sincere thanks to Mr. O. Takizawa,
CRL, for his help in transplanting the system for the new machine.

References
   the Performing Arts", August, 4-6, 1986.
Fig. 1 A block diagram of the system.

Fig. 2 An example of room configurations on which Field Simulator II is operated.
S: Sound source, R: Listening position, Lines represent sound rays.
RAUMAKUSTISCHE VERBESSERUNG BEIM UMBAU DER PAULSKIRCHE IN FRANKFURT AM MAIN

W. Westphal* und H. Kuttruff**

* Planungsbüro für Akustik, D 6706 Wachenheim ** Institut für Technische Akustik, RWTH Aachen, D 5100 Aachen

1. Einleitung


2. Der Raum

Der Grundriß und der Längsschnitt des Raumes sind in Bild 1 wiedergegeben. Es handelt sich also im wesentlichen um einen elliptischen Zylinder, über dem sich eine schwach konkav geformte Decke wölbt mit einer laternenartigen Überhöhung in der Mitte. Die Hauptsachsen der Ellipse betragen 36,0 m und 28,7 m. Bei einer mittleren Höhe von 22 m ergibt sich das Volumen zu 17 800 m³. Für das Publikum ist eine feste Bestuhlung von 933 Plätzen vorhanden. Der Saal ist bei Festveranstaltungen immer voll besetzt, es ergibt sich dann aber immer noch das sehr große spezifische Volumen von 19 m³ pro Person.

3. Umbaumaßnahmen

Durch die Vorgabe, daß das Erscheinungsbild des Raumes aus Gründen des Denkmalschutzes nicht verändert werden durfte, waren die Möglichkeiten der raumakustischen Gestaltung stark eingeschränkt. Die an sich wünschenswerten Einbauten zur Schall-Übertragung und zur Schall-Zerstreuung waren uns verwehrt. Somit blieb als letzte Möglichkeit nur, die Wände und die Decke stärker schallabsorbierend auszustatten.


Die bislang bei anderen Gebäuden verwendeten Schallschluß-Reduktionsputze waren entweder bei guter Absorption mechanisch zu weich oder bei ausreichender Härte zu wenig absorbierend. Nach unserer Kenntnis wurde hier erstmals ein Putz verwendet, der beide geforderte Eigenschaften gleichzeitig aufweist, nämlich hinreichende breitbandige Schallabsorption und gute mechanische Festigkeit. Sein im Kondensator Rohr, also bei senkrechtem Schalleinfall gemessener Absorptionsgrad für 5 cm Dicke beträgt 0,55 für Frequenzen oberhalb von 400 Hz. Durch einen nachträglichen Farbanstrich verminderte er sich bei mittleren und hohen Frequenzen etwas, die verbleibende Absorption war aber immer noch gut. Dieser Putz wurde im ganzen unteren Bereich bis zu einem Zwischengesims in 13,5 m Höhe homogen auf die Wand aufgetragen. Im darüberliegenden Wandbereich wurde zwecks verstärkter Absorption der tiefen Frequenzen eine andere Konstruktion gewählt, nämlich der gleiche Putz, aber nur in 13 mm Stärke und aufgetragen auf einer gelochten Gipskartonplatte, welche 10 cm Abstand von der gemauerten Wand hatte. Der Absorptionsgrad dieser Konstruktion weist ein breites Maximum von 0,7 im Bereich um 180 Hz auf.

Als weitere akustisch relevante Ausstattungsgegenstände sind noch das leicht gepolsterte Gestühl zu erwähnen und 12 große Fahnentranen.
4. Akustische Situation nach dem Umbau

Bei den Messungen, die vor und nach dem Umbau zur Bestim- 
mung der wichtigsten Größen durchgeführt wurden, war der Raum 
zu einem Drittel mit Publikum besetzt, das gleichmäßig ver-
teilt war, d.h. nur jede dritte Reihe wurde besetzt. Zur aku-
stischen Anregung des Raumes bei den Messungen der Nachhall-
zeits, der Deutlichkeit und des Klarheitsmaßes diente jeweils 
ein Schuß aus einem 9 mm Revolver.

Bild 2 zeigt die 
Nachhallzeit vor und 
nach dem Umbau. 
Weg


den der unterschied-
lichen Krümmung ver-
schiedener Abkling-
kurven haben wir zur 
einheitlichen und 
definierten Darstel-
lung die Anfangsnachhallzeit $T_0$ aus-
gewertet. Sie ist 
durch die beschrie- 
benen Maßnahmen von 
3,6 sec auf 1,3 sec 
bei 500 Hz zurückge- 
gangen. Dies ist für 
das Raumvolumen von 
17 800 m$^3$ ein er-
staunlich niedriger 
Wert. - Aus den ge-
wonnenen Reflekt-
ogrammen wurden wei-
therin der Deutlichkeitsgrad $D$ und das Klarheitsmaß $C$ ermit-
telt. Beide Werte streuen natürlich erheblich je nach Lage 
des Hörerplatzes. In Bild 3 und 4 sind die Durchschnittswerte 
über sechs gleichmäßig im Raum verteilte Plätze wiedergege-
ben. Bei beiden Größen trat durch den Umbau eine deutliche 
Verbesserung ein.

\[ D = \frac{\int_0^{50} p^2 dt}{\int_0^{90} p^2 dt} \]

\[ C = 10 \log \frac{\int_0^{50} p^2 dt}{\int_0^{90} p^2 dt} \]

Bild 3: Deutlichkeitsmaß 
Bild 4: Klarheitsmaß

Bild 5: RASTI-Wert (obere Zahl: vor Umbau, untere Zahl: nach Umbau)

6. Zusammenfassung

ACOUSTIC DESIGN OF TOKYO DOME

D. Syo, K. Kageyama and H. Kimura

Takenaka Tech. Res. Lab. 5-14, Minamisuna 2-chome, Koto-ku, Tokyo 136, Japan
Nikken Sekkei Ltd. 4-27, Kouraku 1-Chome, Bunkyō-ku, Tokyo 112, Japan

Introduction

The Tokyo Dome, the first enclosed baseball stadium in Japan, was planned in an air supported structure. It was designed as a facility for multipurpose use and was completed 17, March, 1988. Since its opening, the Tokyo Dome has been used for not only baseball games but also concerts, exhibitions, football games, professional boxing, etc. The Tokyo Dome has room volume of 1,240,000 m³, total surface area of 84,180 m² with 56,000 seats in base building configuration and 13,000 movable seats around the field. The dimensions of Tokyo Dome is listed in Table 1. This paper describes an outline of the acoustic design of Tokyo Dome and its acoustic characteristics measured and these are compared with those of the existing air-supported dome in Canada and U.S.

Outline of Acoustic Design

A major acoustic requirement for a large space like Tokyo Dome is the excellent speech clarity. For this objective, how to cope with the acoustic problems such as excessive reverberation, generation of long pass echo, non-uniformity of sound pressure/tone quality, etc. were the themes in the acoustic design. Although there were many restrictions on the specifications of interior materials and the room shape, the acoustic design was carried out referring to the following three factors and the measurement results of the existing air-supported domes.

1) Shortening the reverberation time by covering a large area with sound absorbing materials.
2) Lowering the ambient noise level.
3) Controlling the interior sound field by fully utilizing electro-acoustic equipment.

Setting Reverberation Time and Selection of Interior Materials

The reverberation time at 500Hz under unoccupied condition was set 0.6 seconds considering the values for large spaces as shown in Fig.1. The ratio of the wall surface to the total surface of Tokyo Dome is large compared with the existing air-supported domes. Therefore, almost walls were effectively covered with sound absorbing materials to shorten the reverberation time and the reflected sound/long pass echo from the back of the seats. In the concourse which is united space with the arena, the ceiling are covered with sound absorbing materials to curb the reflection of crowd noise.

A new fabric material was developed for such an extraordinary large roof and it has similar sound absorption performance to the inner fabric roof used in the existing air-supported domes. Quality control to the acoustic performance of the inner fabric roof was applied to 35 rods (approx. 1,000 samples) delivered to the construction site by checking the air permeability and the sound absorbing coefficient using the tube method. Figure 2 shows the specifications and locations of sound absorbing materials.
Table 1: Dimensions of Tokyo Dome

| Room Volume | 1,240,000 m² |
| Total Surface Area | 84,100 m² |
| Field Area | 13,000 m² |
| Roof Area | 33,180 m² |
| Height (max.) | 61,690 m |
| Seating Area | 22,500 m² |
| Seating Capacity | 56,000 seats |

Fig. 1: Reverberation Time at 500 Hz in Large Spaces.

Fig. 2: Specifications and Locations of Sound Absorbing Materials.

**Setting Ambient Noise Level and Fan Noise Control**

The design target value for ambient noise level was set to be below the levels of NC-45 referring to the measurement results of the existing air-supported domes.

The machine rooms for the pressure fans and air-conditioners were planned to be immediately under the ring beam for various reasons. Therefore, the spontaneous attenuation of fan noise throughout the duct system was not expected. In order to reduce the fan noise, plenum chambers, elbow silencers and splitter type silencers were installed at the wall and ceiling surface of the machine rooms were finished with sound absorbing materials.

**Electro-Acoustic System Design**

The loudspeaker system was planned using semi-distribution system with the main cluster speakers serving the field and seating area and satellite cluster speakers serving only the seating area. The locations and specifications of the loudspeakers were decided as follows with a maximum sound pressure level of 95 dB(A) and with the following four factors taken into consideration:

1) Ensuring uniform sound levels throughout the field and the seating area.
2) Coordinating the direction of the sound with “aurora vision”.
3) Preventing the balls from hitting the loudspeakers.

228
4) Capability of meeting various events.

Total sound system is controlled by a computer system which has many back-up functions and 6 basic patterns. Serving level, delay time for each loudspeaker and loudspeaker on-off etc. are recorded in the disk and manual operation of the system is only for selection of patterns. Besides, interactive system is made possible for setting of option patterns.

**Measured Acoustic Characteristics**

**Reverberation time**

The reverberation time was measured under unoccupied condition radiating pink noise from the main cluster speaker. Figure 3 shows the reverberation times in Tokyo Dome and other existing air-supported domes. The reverberation time at 500Hz in the Tokyo Dome is 5.6 seconds, which meets the design target value and is the shortest among the main existing air-supported domes. The reverberation time at 125Hz with 56,000 persons is estimated 4 seconds. 3 seconds at 500Hz and 2.5 seconds at 2kHz from the values measured in the unoccupied dome. These reverberation times are excellent for a large space like that of Tokyo Dome.

**Ambient noise level**

The ambient noise level was measured operating 8 units of pressure fans and 36 units of air-conditioners. Figure 4 shows the ambient noise levels in Tokyo Dome and other existing air-supported domes. The ambient noise level in Tokyo Dome is NC-41, which meets the design target value and is the lowest among the main existing air-supported domes.

![Fig. 3 Reverberation Times measured in Tokyo Dome and the existing air-supported domes.](image)

![Fig. 4 Ambient Noise Levels measured in Tokyo Dome and the existing air-supported domes.](image)

**Percent articulation, RASTI and sound pressure level distribution**

The percent articulation, RASTI and the sound pressure level distribution were measured in the pattern of sound system of baseball configuration. Figure 5 shows the results measured. The percent articulation is 95.5% in average (good hearing condition), RASTI is 0.47 in average (fair) and the sound pressure level variation is within ± 2.5dB at 1kHz.
Fig. 5 Percent Articulation, RASTI and Sound Pressure Level Distribution measured. (Upper: Percent articulation, Middle: RASTI and Lower: Sound pressure level distribution at 1kHz)

Echo time pattern
The echo time pattern was measured with the pistol shot noise. Figure 6 shows the echo time patterns in the field of Tokyo Dome and the existing air-supported domes. Both the density and duration of reflected sounds in the echo time pattern of Tokyo Dome are sufficiently small in comparison with those of the exiting air-supported domes. It is regarded that these results are due to the difference of the room shape and the area of sound absorbing treatment.

Fig. 6 Echo Time Patterns measured in the field of Tokyo Dome and the existing Air-supported domes.

Summary
As can be seen from the discussions, the acoustic characteristics of Tokyo Dome meet the design targets. It is evaluated that the Tokyo Dome has short reverberation time, low ambient noise level and high speech clarity as a large space with capacity more than 1,000,000m².

Furthermore, sound services for the various events can be served utilizing the easy-to-operate computer controlled sound system.

References
ACOUSTICAL DESIGN OF THE "KAROLOS KOUN" THEATRE IN ATHENS, GREECE

Tsinikas N.\textsuperscript{1}, Tzekakis E.\textsuperscript{1}

\textsuperscript{1}Laboratory of Architectural Technology, School of Architecture, Aristotle University of Thessaloniki, 54006, Thessaloniki, Greece

Summary

The "Karlois Koun" Theatre is a new, small, almost cubical in shape theatre in Athens for an audience of 307 persons (187 in stalls and 120 in a balcony). Its volume is 2605 m\textsuperscript{3} and its volume/seat ratio of 9 m\textsuperscript{3} is above optimum. The placement of the audience is optimum for direct sound. Twelve roof reflectors were positioned inside the structural elements of the wooden roof and eight lateral reflectors were mounted in the columns between the corridors and the scene. The use of wood as a structural element as well as in large quantities in general, made the R.T. predictions in the lower frequencies problematic. However, R.T. predictions using a microcomputer program, done to define the amount of absorptive materials required and measurements made after the construction to confirm the prediction were compared satisfactory.

1. Introduction

The old Art Theatre "Karolos Koun" was established and housed in a small theatre on the basement of another 2000 seats theatre in the centre of Athens. For more than 30 years it functioned presenting high quality plays and many famous actors performed there under the guidance of Karolos Koun.

Fig. 1 Plan of the Theatre.

The lack of sufficient sound insulation between the two theatres created coexistence problems and finally with the help of the Ministry of Culture an old
warehouse was given in Plaka, the old city near Acropolis, to be transformed to house the new Art Theatre "Karolos Koun".

2. General description of the new Art Theatre

The new Art Theatre was designed by the Architect M. Perrakis. The plan, the longitudinal section and the cross section of the theatre are shown respectively in figures 1, 2 and 3. The dimensions of the new Art Theatre "Karolos Koun" are 15,5 by 13,5m and 10m high. The audience is in two levels, 187 in stalls with rake angle of 15° and 120 in balcony with rake angle 21°. The level of the balcony leads to two side corridors. Between the corridors and the scene there are columns dividing the space into three clites. The space is covered with a roof from which the central part is higher than the roof of the side corridors forming a kind of three clites basilica. Apart from the periphery walls and floor of the stalls all the other construction is made of wood including columns, beams, floor of the balcony, floor of the corridors and roof. The volume of the room is 2605m³ and the total capacity is 307 persons covering an area of 139,4m². The total internal surface is 1563m² and the mean free path is 7,64m. The volume per seat ratio of 9m³ is more than optimum but it doesn’t cause problems for the control of reverberation. The rake of the audience is quite good for direct sound. The use of wood in large surfaces can easily cause problems in low frequencies which can be avoided by increasing the stiffness of their supporting skeleton. From similar use of wood in such spaces in Europe no problem has been reported in low frequencies up to now.

Fig. 2 and 3 Longitudinal and Cross section of the Theatre. Reflections from the roof, from ceiling, from under balcony reflectors, from side wall reflectors and from side walls as reflectors.

3. Acoustic Design

The roof reflects sound to the audience from the two sides to almost all the area of the stalls and to every seat arrive two reflections from the roof. The disadvantage arose from the height of the part of the stalls under the balcony which is very small creating poor acoustic conditions to the rear seats from the roof leaving only direct sound and side wall reflections to reach them.

R.T. prediction using a microcomputer program (1) has been done to define the amount of absorptive materials required. The (A) and (C) curves in figure 4 are predicted R.T. of untreated (without absorbing materials) and treated room with seats.
The central part of the ceiling is a very good reflecting surface and is used for that purpose (fig. 2). Before and after that part of the roof, twelve reflectors are positioned, eight above the scene and four above the balcony. The angle of these reflectors is adjustable to reflect sound as required. The area receiving sound from the ceiling reflectors is also shown in the longitudinal section. [2]. For the part of the stalls, under the balcony which is an acoustically poor area an horizontal reflector under the balcony is provided and an inclined reflector on the rear wall in order to enhance sound in the rear seats, figure 2. Their effect on improvement of sound energy distribution including direct sound must be considered as significant. These constructions under the balcony improve the situation locally although not in a completely satisfactory way, forcing the designers to the improvement of the sound field by using a small electroacoustics system. Eight side wall reflectors, four on each side, were considered to be installed, mounted in the columns between the corridors and the scene. These side wall reflectors contribute to the audience and the receiving area, figure 3. The side walls are inclined to be used as reflecting surfaces and they also contribute to the enhancement of sound field, figure 3. Finally, provision for movable side reflectors in the scene as parts of the scenery is taken. Also, the rear wall of the scene is used as reflector.

The reverberation time prediction showed that in order to achieve an optimum reverberation time of 1.1 s with audience 110m² glasswool light slabs 5cm thick are needed, covered by wood slabs 2x4cm/10cm on a piece of cloth [3], and 200m² of resonators consisted of plywood 6mm on .10cm glasswool in frames 1x3m [4].

Fig.4 R.T. curves. (A) predicted, untreated room with seats, (B) measured, untreated room with seats, (C) predicted, treated room with seats, (D) measured, treated room with seats.

4. Reverberation time measurements

After the construction of the theatre, before and after the positioning of the absorbing materials, R.T. measurements were made to confirm the prediction. From the projected acoustical treatment some provisions were not constructed due to time problems from the contractor. The side wall reflectors finally were only four, the ones on the second level and the side walls were left vertical and not inclined. Also, all the wood surfaces were not painted by many hands of varnish
but just one. The R.T. measurement of the theatre with seats and without
absorbing materials is shown in (B) curve in figure 4 and is lower than expected in
low frequencies due of course to the enormous amount of internal wood surfaces.
The R.T. measurement of the theatre with seats and absorbing materials is shown
in (D) curve in figure 4 and is lower than expected in low frequencies, near
optimum with a slight difference in middle and middle to high frequencies. By
introducing a small amount of a porous absorber the curve of reverberation will be
level throughout the frequencies, although almost the same in all frequencies
which is adequate for speech, but slight lower than the optimum.

5. Conclusions

The new Art Theatre "Karolos Koun" in Athens is one of the first theatres
that has been constructed under acoustical project supervision, providing a
balance between architecture and acoustics. There is no sacrifice of acoustics to
the architectural solutions nor the opposite. The close, systematic collaboration
with the architect whose drawings were appreciated by the critics and the
newspapers as very successful gave a final result unexpected by actors, public,
officials and the founder of the theatre.

Fig. 5. Photograph of the interior of the theatre. Note the ceiling construction, and
the movable reflectors.

References

[1] Tzekakis E., Reverberation time prediction program, Proceedings FASE 5th
Symposium, Thessaloniki, 1985

[2] Tzekakis E., Tsinikas N., Two applications of Tosan 4 - Reverberation time
prediction program, Proceedings FASE 5th Symposium, Thessaloniki, 1985

Proceedings FASE 5th Symposium, Thessaloniki, 1985
ACOUSTIC DESIGN AND FINAL RESULTS OF RECONSTRUCTED THEATRES

KOTSCHY, András

ÁÉTV- Building Design Institute, Acoustic Lab.
7016 Budapest, Krisztina krt. 99. Hungary

History
In the past years 4 old Hungarian country theatres
(Szeged, Kaposvár, Veszprém, Kecskemét) have been restored. Besides this, a new up-to-date theatre that seats 700 was completed in Győr, in 1978. It seems a fairly good result in such a small country like Hungary.

Completed Reconstructions
The neo-Baroque building of the National Theatre in Szeged was completed in 1883, according to the plans of Viennese architects F. Fellner and H. Helmer. The theatre burnt down in 1885. During the quick restoration, some alterations were carried out, the seating capacity was reduced from the original 1760 persons to 1018 ones. The second formal opening took place in 1886, and the theatre which was of a high quality in every respect became one of the most successful buildings of its kind at the turn of the century.

Both the opera and the drama performances were regarded very good acoustically.

The reconstruction of the theatre started in 1981 and was completed in 1986. The most valuable part of the building is the impressive auditorium so there was no reconstruction here. Since the parts below the auditorium were rebuilt, the old roofing of the ground floor was opened, and the new one was completed with wooden flooring, with an arrangement of seats providing better visibility and with more comfortable, upholstered seats. The seating capacity was reduced to 750 (it was 856 the most before restoration). The most changes were carried out on the stage, the surface of which was increased by 150 sq metres and it was equipped with new steel construction and up-to-date stage techniques.

When planning the reconstructions of all the theatres, the main was that the reverberation time in the auditoriums won't change substantially, and that the expected values in the empty and occupied auditorium should come close to each other.

The Csiky Gergely Theatre, Kaposvár was built in 1911, according to the plans of the architects E.Magyar and J. Stahl. The building, which is richly decorated both on the outside and the inside in Art Nouveau styles, was meant to be a summer theatre, and at the time of its opening it had a seating capacity of 1500.

The reconstruction started in 1986 and took 18 months.
The formal opening was in March, 1988.

The interior architecture of the auditorium was kept in its original form. The new back wall for establishing technical rooms on the ground floor of the auditorium follows it very well. There is a new wooden flooring on the ground floor but the low gradient was not changed because of the low rows of boxes. New seats were made and they were arranged in a more spacious way. There are no seats any more on the second-floor side gallery (the spotlights of the stage lighting are placed there), thus the earlier seating capacity of 647 was reduced to 547.

The Art Nouveau building of Petőfi Theatre, Veszprém was built in 1908, according to the plans of T. Nagyaszay. Meetings and balls were also held there, and therefore, the auditorium had a flat floor with an inclined sectional platform.

The reconstruction, which started in 1985, was completed in 1988. During it a new inclined floor was made in the auditorium aiming at improving visibility and audition. The original decorative windows were bricked up to protect against the noise of the traffic, but their ornamentation can be still admired on the inner side walls and is an integral part of the sight of the auditorium. The walls in the auditorium have ceramic cover bearing the architect's original motives, and it also functions as a resonant absorbent surface in the lower frequency range (100 Hz). The walls of the boxes were covered with tapestry according to the original plans. We used it as an acoustically active surface too.

The seating capacity of the auditorium is 445.

In Table 1 the main technical data of the reconstructed theatres are given.

### Acoustic Measurements and their Results

**I. Objective parameters**

According to the possibilities, we have carried out acoustic measurements in the theatres before and after the reconstruction.

The measured objective parameters are the reverberation time, the rapid speech transmission index (RSTI) measured with the new Briel & Kjaer instruments and the intelligibility level.

In two theatres, we had the possibility to carry out speech intelligibility measurements, too. It was a good occasion to compare their results with those of the RSTI measurements.

The summarized measurements results are given in Table 2.

**II. Subjective parameters**

In all 3 theatres simple test-forms were distributed to the public before the performances chosen earlier in advance and we asked their opinion about the natural acoustic sound in the auditorium.

In the National Theatre, Szeged we had the possibility to realise subjective examinations by 41 test persons in empty...
and occupied hall, which were evaluated according to the 
recommencement of the CMA.
The test persons heard short music samples of concert and 
opera selections performed by an orchestra and singers and 
also of piano and violin recitals. We asked the subjective 
opinions of the musicians in the orchestra too.
The summarized subjective judgements are given in 
Table 3.

Conclusions
We can generally conclude that according to the public 
opinions, the room acoustic conditions of the examined 
thrares had been relatively good and remained so after the 
reconstruction too, or it was improved to a certain extent.
There was only a short time to carry out the measurements 
in the theatres, especially in the occupied auditoriums.
We can regard them only informing measurements but they seem 
to be suitable as a control of the acoustic efficiency of the 
reconstruction of the theatres.

We get a relatively consequent agreement between the 
results of the rapid speech transmission index (RASTI) and 
those of the speech intelligibility measurements. It means 
that we can obtain good results with the new measuring method 
more quickly and in a more simple way. As for the results of 
several similar measurements carried out in different lecture 
halls, we can claim that the results of the speech intelligi-
gibility measurements are more "optimistic" and the results of 
the RASTI measurements are more "pessimistic" than reality, 
but the results of the RASTI measurements are closer to 
reality.
The correlation between the objective and subjective 
parameters is quite good.

References
-CMA recommendation, NTISwF GOSbssZTROJ SzSzSzR, Moscow 1983
-V. L. Jordan: Acoustical Design of Concert Halls and Theatres, 
Applied Science Publishers LTD 1980
* 13th INTERNATIONAL CONGRESS ON ACOUSTICS *
* YUGOSLAVIA * 1989 *

<table>
<thead>
<tr>
<th>Theatre</th>
<th>seats</th>
<th>Volume (m³) auditorium</th>
<th>stage</th>
<th>m³/pers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szeged</td>
<td>750</td>
<td>4700</td>
<td>10500</td>
<td>6.3</td>
</tr>
<tr>
<td>Kaposvár</td>
<td>547</td>
<td>2450</td>
<td>4100</td>
<td>4.5</td>
</tr>
<tr>
<td>Veszprém</td>
<td>445</td>
<td>3300</td>
<td>2500</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Table 1
The main technical data of the reconstructed theatres

<p>| Theatre  | Tₘₜ₂₃₅₆₇₈₉₁₀₀₄₀₀₄₀₂₀₀₄₀₃₀₄₀₅₀₆₀₇₀₈₀₉₀₁₀₀₁₁₀₀₁₂₀₀₁₃₀₀₁₄₀₀₁₅₀₀₁₆₀₀₁₇₀₀₁₈₀₀₁₉₀₀₂₀₀₀₂₁₀₀₂₂₀₀₂₃₀₀₂₄₀₀₂₅₀₀₂₆₀₀₂₇₀₀₂₈₀₀₂₉₀₀₃₀₀₀₃₁₀₀₃₂₀₀₃₃₀₀₃₄₀₀₃₅₀₀₃₆₀₀₃₇₀₀₃₈₀₀₃₉₀₀₄₀₀₀₄₁₀₀₄₂₀₀₄₃₀₀₄₄₀₀₄₅₀₀₄₆₀₀₄₇₀₀₄₈₀₀₄₉₀₀₅₀₀₀₅₁₀₀₅₂₀₀₅₃₀₀₅₄₀₀₅₅₀₀₅₆₀₀₅₇₀₀₅₈₀₀₅₉₀₀₆₀₀₀₆₁₀₀₆₂₀₀₆₃₀₀₆₄₀₀₆₅₀₀₆₆₀₀₆₇₀₀₆₈₀₀₆₉₀₀₇₀₀₀₇₁₀₀₇₂₀₀₇₃₀₀₇₄₀₀₇₅₀₀₇₆₀₀₇₇₀₀₇₈₀₀₇₉₀₀₈₀₀₀₈₁₀₀₈₂₀₀₈₃₀₀₈₄₀₀₈₅₀₀₈₆₀₀₈₇₀₀₈₈₀₀₈₉₀₀₉₀₀₀₉₁₀₀₉₂₀₀₉₃₀₀₉₄₀₀₉₅₀₀₉₆₀₀₉₇₀₀₉₈₀₀₉₉₀₀₁₀₀₀₀ |
| Tₘₜ₂₃₅₆₇₈₉₁₀₀₄₀₀₄₀₂₀₀₄₀₃₀₄₀₅₀₆₀₇₀₈₀₉₀₁₀₀₁₁₀₀₁₂₀₀₁₃₀₀₁₄₀₀₁₅₀₀₁₆₀₀₁₇₀₀₁₈₀₀₁₉₀₀₂₀₀₀₂₁₀₀₂₂₀₀₂₃₀₀₂₄₀₀₂₅₀₀₂₆₀₀₂₇₀₀₂₈₀₀₂₉₀₀₃₀₀₀₃₁₀₀₃₂₀₀₃₃₀₀₃₄₀₀₃₅₀₀₃₆₀₀₃₇₀₀₃₈₀₀₃₉₀₀₄₀₀₀₄₁₀₀₄₂₀₀₄₃₀₀₄₄₀₀₄₅₀₀₄₆₀₀₄₇₀₀₄₈₀₀₄₉₀₀₅₀₀₀₅₁₀₀₅₂₀₀₅₃₀₀₅₄₀₀₅₅₀₀₅₆₀₀₅₇₀₀₅₈₀₀₅₉₀₀₆₀₀₀₆₁₀₀₆₂₀₀₆₃₀₀₆₄₀₀₆₅₀₀₆₆₀₀₆₇₀₀₆₈₀₀₆₉₀₀₇₀₀₀₇₁₀₀₇₂₀₀₇₃₀₀₇₄₀₀₇₅₀₀₇₆₀₀₇₇₀₀₇₈₀₀₇₉₀₀₈₀₀₀₈₁₀₀₈₂₀₀₈₃₀₀₈₄₀₀₈₅₀₀₈₆₀₀₈₇₀₀₈₈₀₀₈₉₀₀₉₀₀₀₉₁₀₀₉₂₀₀₉₃₀₀₉₄₀₀₉₅₀₀₉₆₀₀₉₇₀₀₉₈₀₀₉₉₀₀₁₀₀₀₀ |
| Tₘₜ₂₃₅₆₇₈₉₁₀₀₄₀₀₄₀₂₀₀₄₀₃₀₄₀₅₀₆₀₇₀₈₀₉₀₁₀₀₁₁₀₀₁₂₀₀₁₃₀₀₁₄₀₀₁₅₀₀₁₆₀₀₁₇₀₀₁₈₀₀₁₉₀₀₂₀₀₀₂₁₀₀₂₂₀₀₂₃₀₀₂₄₀₀₂₅₀₀₂₆₀₀₂₇₀₀₂₈₀₀₂₉₀₀₃₀₀₀₃₁₀₀₃₂₀₀₃₃₀₀₃₄₀₀₃₅₀₀₃₆₀₀₃₇₀₀₃₈₀₀₃₉₀₀₄₀₀₀₄₁₀₀₄₂₀₀₄₃₀₀₄₄₀₀₄₅₀₀₄₆₀₀₄₇₀₀₄₈₀₀₄₉₀₀₅₀₀₀₅₁₀₀₅₂₀₀₅₃₀₀₅₄₀₀₅₅₀₀₅₆₀₀₅₇₀₀₅₈₀₀₅₉₀₀₆₀₀₀₆₁₀₀₆₂₀₀₆₃₀₀₆₄₀₀₆₅₀₀₆₆₀₀₆₇₀₀₆₈₀₀₆₉₀₀₇₀₀₀₇₁₀₀₇₂₀₀₇₃₀₀₇₄₀₀₇₅₀₀₇₆₀₀₇₇₀₀₇₈₀₀₇₉₀₀₈₀₀₀₈₁₀₀₈₂₀₀₈₃₀₀₈₄₀₀₈₅₀₀₈₆₀₀₈₇₀₀₈₈₀₀₈₉₀₀₉₀₀₀₉₁₀₀₉₂₀₀₉₃₀₀₉₄₀₀₉₅₀₀₉₆₀₀₉₇₀₀₉₈₀₀₉₉₀₀₁₀₀₀₀ |
| Tₘₜ₂₃₅₆₇₈₉₁₀₀₄₀₀₄₀₂₀₀₄₀₃₀₄₀₅₀₆₀₇₀₈₀₉₀₁₀₀₁₁₀₀₁₂₀₀₁₃₀₀₁₄₀₀₁₅₀₀₁₆₀₀₁₇₀₀₁₈₀₀₁₉₀₀₂₀₀₀₂₁₀₀₂₂₀₀₂₃₀₀₂₄₀₀₂₅₀₀₂₆₀₀₂₇₀₀₂₈₀₀₂₉₀₀₃₀₀₀₃₁₀₀₃₂₀₀₃₃₀₀₃₄₀₀₃₅₀₀₃₆₀₀₃₇₀₀₃₈₀₀₃₉₀₀₄₀₀₀₄₁₀₀₄₂₀₀₄₃₀₀₄₄₀₀₄₅₀₀₄₆₀₀₄₇₀₀₄₈₀₀₄₉₀₀₅₀₀₀₅₁₀₀₅₂₀₀₅₃₀₀₅₄₀₀₅₅₀₀₅₆₀₀₅₇₀₀₅₈₀₀₅₉₀₀₆₀₀₀₆₁₀₀₆₂₀₀₆₃₀₀₆₄₀₀₆₅₀₀₆₆₀₀₆₇₀₀₆₈₀₀₆₉₀₀₇₀₀₀₇₁₀₀₇₂₀₀₇₃₀₀₇₄₀₀₇₅₀₀₇₆₀₀₇₇₀₀₇₈₀₀₇₉₀₀₈₀₀₀₈₁₀₀₈₂₀₀₈₃₀₀₈₄₀₀₈₅₀₀₈₆₀₀₈₇₀₀₈₈₀₀₈₉₀₀₉₀₀₀₉₁₀₀₉₂₀₀₉₃₀₀₉₄₀₀₉₅₀₀₉₆₀₀₉₇₀₀₉₈₀₀₉₉₀₀₁₀₀₀₀ |</p>
<table>
<thead>
<tr>
<th>intelligibility(%)</th>
<th>RASTI</th>
<th>C₅₀ (dB) empty</th>
<th>occup.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szeged</td>
<td>89.7</td>
<td>86</td>
<td>1.88</td>
</tr>
<tr>
<td>Kaposvár</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Veszprém</td>
<td>85.0</td>
<td>84.5</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.80</td>
</tr>
</tbody>
</table>

Table 2
The measured objective parameters
(b= before the reconstruction; r= after the reconstruction; e= empty; o= occupied; m= calculated)

<table>
<thead>
<tr>
<th>Theatre</th>
<th>very good</th>
<th>good</th>
<th>adequate</th>
<th>satisfactory</th>
<th>bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szeged, test persons</td>
<td>35</td>
<td>32</td>
<td>26</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>public</td>
<td>42</td>
<td>47</td>
<td>8</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Kaposvár, public</td>
<td>44</td>
<td>44</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Veszprém, public</td>
<td>57</td>
<td>31</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3
Summary of the subjective judgement of the public
and test persons (%)
VARIABLE ACOUSTICS, THE DESIGN OF DIFFERENT HALLS:

P.H. Heringa and V.M.A. Peutz
Peutz & Associés B.V.
P.O. box 407
6500 AR NIJMEGEN
The Netherlands

As shown in the companion paper the acoustical requirements for different uses of a hall and especially a concert hall and a theatre are quite different or even contradictory. However, in many cases it is possible to find a solution that is quite satisfactory for both purposes: Variable Acoustics. In general, one has to find an acoustical volume that can be satisfactory for the different purposes and that can rather easily be adapted to the use of the hall in terms of volume and absorption. Hereafter we give a few examples of already realised halls of very different basical sizes (and building cost) and use.

Figure 1 shows the main hall of the congress centre SAVA-Centar Belgrade. For this case a solution with movable ceiling-elements is applied. These elements can be lowered to reduce the volume, and it is possible to turn the elements in order to create gaps to a sound absorbing volume. For symphonical music an orchestra shell was designed. The figure shows the different possible uses of the hall.

The next three examples (fig. 2, 3 and 4) are based on the same principle but are of a different size. We use the maximum of the volume of the hall in the concert situation by a good coupling of the stage volume of the hall to the auditorium. The back and side stages with storage are screened in the concert situation by panels fixed to rails. The required variation in acoustics is obtained by removing these panels and installing the required theatre equipment: curtains, required stage opening etc. The largest example, with 700 seats, has the advantage that theatre equipment such as curtains don't have to be removed. In this case it was financially possible to create a stage tower, so curtains could be raised and a reflector could be installed underneath. In the halls shown in figure 2 and 3 it was also necessary to install retractable curtains in the auditorium part to achieve the required variation. The example in figure 4 is with respect to his volume just a small recital hall in its concert situation.

Figure 5 shows a newly built theatre in Rotterdam of which acoustics can be adapted to theatre and to opera by means of retractable curtains and a reflector.

Figure 6 is the example of the Espace de Projection, IRCAM, Paris a hall destined for concert. Its acoustics can, by will of the conductor, be adapted by means of rotatable three-sided elements. Each side has different acoustical properties. These elements cover walls and ceiling completely. Furthermore, as shown in figure 6, different parts of the ceiling can be raised and lowered.
1. Sava-Centar, Belgrade, Yugoslavia, 1979
Congress hall/concert hall/opera/theatre, 4200 seats.
Architect: S. Maksimovic, Belgrade
Acousticians: V. Peutz and Milosavljevic

<table>
<thead>
<tr>
<th>Volume</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concert</td>
<td>31000 m³</td>
</tr>
<tr>
<td>Theatre</td>
<td>22000 m³</td>
</tr>
</tbody>
</table>

2. De Maaspoort, Venlo, The Netherlands, 1984
Concert hall/opera/theatre, 700 seats.
Architect: Werkgroep Venlo, 's-Hertogenbosch

<table>
<thead>
<tr>
<th>Volume</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concert</td>
<td>10000 m³</td>
</tr>
<tr>
<td>Theatre</td>
<td>5000 m³</td>
</tr>
</tbody>
</table>
Small concert hall/theatre, 500 seats.
Architect: De Bever, Eindhoven

<table>
<thead>
<tr>
<th>Volume</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concert</td>
<td>5000 m$^3$</td>
</tr>
<tr>
<td>Theatre</td>
<td>2300 m$^3$</td>
</tr>
</tbody>
</table>

Swimming pool transformed into theatre/recital hall, 300 seats.
Architect: Town-architect Sikma, Zevenaar

<table>
<thead>
<tr>
<th>Volume</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concert</td>
<td>3500 m$^3$</td>
</tr>
<tr>
<td>Theatre</td>
<td>2000 m$^3$</td>
</tr>
</tbody>
</table>
5. Espace de Projection, IRCAM, Paris
Concert hall with variable acoustics
Architect: Piano and Rogers Paris/London.
Volume, RT (500, 1000 Hz)
Max. 4500 m³ 4.5 s
Min. 900 m³ 0.8 s

6. Schouwburg Rotterdam, 1988
Theatre/opera
Architect: Quist Rotterdam
Volume, RT (500, 1000 Hz)
Opera 5000 m³ 1.2
Theatre 4000 m³ 0.9
ACOUSTIC DESIGN OF VARIABLE REVERBERATION UNITS FOR MULTIPURPOSE HALLS

FURUKAWA Sen-ichi

Consultant in Acoustics
3-19-3-504, Kamishakujii, Nerima, Tokyo 177, Japan

INTRODUCTION

We have designed architectural acoustics of multipurpose halls. In order to use one hall for multipurpose, reverberation of a hall should be changed desirable acoustic conditions of each use from a concert to an assembly. Then, it is necessary that variable reverberation units are provided in a space of an auditorium.

In this report, we have described a concept of acoustic design for multipurpose halls, acoustic design of variable reverberation units and measured data after the completion of the halls.

CONCEPT OF ARCHITECTURAL ACOUSTIC DESIGN

I think that direct sound and early reflected sound are very important for audiences and performers, how sound to reach and to hear them about every events such as musics, dramas, shows and meetings. Because it is necessary for communication of informations that the early reflected sound increase the energy of the direct sound arriving at audiences.

By this reason, we have decided the room shape of a hall plane- and cross-sectionally. At first, in order that the sound reaches every seats directly without obstructors from a sound source on a stage, the floor slope of an auditorium is decided to be able to see the front part of a stage from audiences, and we have designed the shape of the walls and the ceiling which are useful to reflect the sound for the audiences in early time from sound sources on a stage.

To use for a concert, we make use of the late reflected sound to increase the reverberation of a hall. To use for an assembly, we make use of the late reflected sound to decrease the reverberation of a hall. Variable reverberation units have these functions.

DESIGN OF VARIABLE REVERBERATION UNITS

We have designed to arrange variable reverberation units on the wall and the ceiling in an auditorium, except that the wall and the ceiling reflect the sound in early time from sound sources on a stage. That is the rear wall, upper part of the rear side walls and the both side of the ceiling in an auditorium.

Variable reverberation units are divided into three groups.

(1) Rotary type: A rotary device is composed of a sound absorptive surface on one side and a sound reflective surface on the other side. In order to increase the difference between sound absorption and reflection at low frequency, absorbing finish of this device has a long air space like a cylindrical type has, as shown in Fig. 1 and 2.
(2) Opening and closing type: This device has sound reflectors to be moved in front of the sound absorptive wall as shown in Fig. 3, or sound absorbers to be moved in front of the sound reflective wall as shown in Fig. 4. The sound absorption of this device increases at low frequency according to the increase of an air space behind the sound absorbing surface. Simple method of this type uses drapery to drow over the sound reflective wall for absorbing the sound energy.

(3) Suspended type: This device consists of a sound absorber with a sound reflector at its bottom and suspend from the ceiling as shown in Fig. 5. The sound absorbers are lifted up to reflect the sound and are pulled down to absorb the sound for shortening the reverberation time.

Besides, there are clearances to move a sound absorber or a sound reflector on the wall and the ceiling. Therefore, it is necessary to take into account the sound absorption caused by the clearance for the device moving. An example of the sound absorption of the clearance is indicated in Fig. 6.

ACOUSTIC DATA OF VARIABLE REVERBERATION UNITS

An example of a hall with a rotary type and a suspended type is shown in Fig. 7. This auditorium (Matsue Plover Hall) has 14 rotary cylindrical devices and 12 suspended cylindrical absorbers with sound reflectors at their bottoms. Figure 8 shows an example of opening and closing type of variable reverberation units. This auditorium (Nakanida Bach Hall) has 80 sound reflectors moving up and down in front of the sound reflective wall. In Figure 9, the auditorium (Shindo Music Hall) has 8 sound absorbers moving up and down in front of the sound reflective wall. An example of variable reverberation unit used drapery on the side walls is indicated in Fig. 10.

CONCLUSION

Variable reverberation units are settled the rear wall, the upper part of the side walls and the ceiling which are not used for reflection of sounds in early time. In order to communicate the speech information, it is necessary to decrease the reverberation of a hall by changing the variable reverberation units from reflective to absorptive. Furthermore, it is important that the variable reverberation units are designed to absorb the sound at low frequency.

REFERENCES

Fig. 1 Cylindrical type of rotated variable reverberation unit.

Fig. 2 Rhomb shaped type of rotated variable reverberation unit.

Fig. 3 Sound reflector moving up and down in front of sound absorbing surface.

Fig. 4 Sound absorber moving up and down in front of sound reflective surface.

Fig. 5 Cylindrical sound absorber suspended from the ceiling.

Fig. 6 Sound absorption of the clearance between reflectors.
Fig. 7 Comparison of reverberation time between sound reflective and absorptive condition of variable reverberation units, and longitudinal section of Matsue Plover Hall.

Fig. 8 Comparison of reverberation time between sound reflective and absorptive condition of variable reverberation units, and a longitudinal section of Nakanida Bach Hall.

Fig. 9 Comparison of reverberation time between sound reflective and absorptive condition of variable reverberation units, and a longitudinal section of Shido Music Hall.

Fig. 10 Comparison of reverberation time between sound reflective and absorptive condition of variable reverberation units, and a longitudinal section of Suka Tomomasa Memorial Hall.
3. E
AN AUTOMATIC SUPPRESSION OF MICROPHONE-LOUDSPEAKER COUPLING
-CONCEPT AND PRELIMINARY SIMULATION-

Yasuhiro RIKO and Tadayori MAKINO
Corporate Engineering Division,
Matsushita Communication Industrial Co., Ltd.
600 Saedo-cho, Midori-ku, Yokohama, 226 Japan

ABSTRACT

A new concept for minimizing the loop gain between microphones and loudspeakers is described.

Contrary to the most conventional concept in which the highest sensitivity direction of microphone array is to be set to that of sound source, this new concept described here is to set the null or minimum sensitivity direction of microphone array to the loudspeaker by detecting direction indicating signal which is a kind of very low level random pulse mixed into sound signal and also fed to the output of microphone in order to separate it from the mixed signal.

Some preliminary simulations are also described.

1 INTRODUCTION

In sound reinforcement system, there were long histories of several methods to minimize the coupling between microphones and loudspeakers.

As a new concept for this purpose we propose a system which consists of adaptive microphone array, reference signals from loudspeaker for detection of the direction of those loudspeakers into microphone arrays and a controller of the directional characteristics of microphone array especially the direction of dips.

The dip direction of microphone array is adjusted by controller to the direction of loudspeaker, which is determined by the reference signal mixed in the sound signal as a spread spectrum mode of very low level. The reference signal is generated as a random pulse and mixed into sound signal by very low level which is not audible but has satisfactory SNR for demodulation. The same reference signal is fed to the microphone output to separate it from signal.

Fig.1 shows the principle difference from the most conventional methods in which the maximum sensitivity direction is adjusted to original sound source such as lecturer, mainly by manual. On the contrary, new system shown in Fig.1b, microphone array seeks
loudspeaker automatically by reference signal. This is an open loop (pulse generator-mixer-amplifier-loudspeaker-microphone array-demodulator-controller) separated from closed sound signal loop.

2 FUNDAMENTAL CONCEPT AND OPERATION
2-1 Control algorithm of directivity of microphone array

The fundamental concept of controlling microphone array is to set the null direction or minimum sensitivity direction in line with the direction of sound wave from the loudspeaker. Theoretically, as is well known, these direction can be set for plural number of arrayed microphone elements.

It means that we can design microphone system according to the requirements for the number of installing loudspeakers.

As the first step simulation, only one loudspeaker was considered which is shown in Fig.2, together with the microphone array.

2-2 Direction Sensitivity Controlled Microphone

The microphone array placed on a circular form. Fig.2 shows the simulation configuration of the microphone array.

The distance $d$ between microphones $M_0$ and $M_1$, and the difference $\Delta \xi$ of the propagation paths between the two microphone units from the loudspeaker make Equation (1).

$$\Delta \xi = d \times \cos(\Theta) \quad (1)$$

where $\Theta$ stands for the angle between the loudspeaker direction and the central axis of microphones.

The sound arrival time difference $\Delta t$ between the two microphone units is expressed by:

$$\Delta t = \Delta \xi / c = d \times \cos(\Theta) / c \quad (2)$$

where $c$: sound velocity.

Subtracting the output signal of the microphone $M_1$, whose output is electrically
2.3 Detection of Loudspeaker Direction

The direction of the loudspeaker is detected by the reference signal from the loudspeaker to the microphone.

The sound level of reference signal is low enough not to be audible.

The spread spectrum concept makes transmission frequency range wider to 50Hz to 20KHz, using a phase modulation by M sequence random pulse of 10KHz clock, and emitted from the loudspeaker as the reference signal mixed with the sound signal.

The process gain of the spread spectrum mode is

$$\Delta L = \Delta f / \Delta f$$

(4)

where

$\Delta F$: frequency range spreaded

$\Delta f$: frequency band needed.

The direction of the loudspeaker is derived from sound propagation time by measurement of three microphone units.

By changing the pulse patterns of M sequence random pulse, each direction of the plural loudspeakers can be detected.

3 RESULT OF SIMULATION

Fig. 3 shows the polar patterns of the directional sensitivity controlled microphone.

Fig. 4 shows the power spectrums of mixture of reference signal and sound signal.

Fig. 5 shows the power spectrums of demodulated signal.

delayed by $\Delta t$, from the output signal of the microphone unit $M_0$, cancels the sensitivity of direction $\Theta$.

The composite output signal $V$ of those microphone units is expressed as:

$$V = V_0 \times (kd)^2 \{1 - \cos(\Theta - \theta)\}$$

(3)

where

$k = w / c$

$\theta$: angle between the axis of $M_0$ and $M_1$ and arbitrary point.

In this case, number of dip directions is three which means this system can respond three speakers.

In Fig. 2, there are two sets of microphone arrays of different radius circles to cover two frequency ranges.
4 CONCLUSION

A preliminary investigation which is intending to minimize the sensitivity of a microphone system to the direction of the loudspeaker in order to reduce coupling of them was performed.

To detect the direction of the loudspeaker, a reference signal from the loudspeaker is added to the sound signal.

As the result of the simulation where one loudspeaker is applied, the sound coupling between the microphone and loudspeaker was decreased by 40 dB.

Higher degree of directivity of microphone using more elements and divided frequency range to more bands, and experiments in the actual sound field with sound reverberation should be the next step.

The authors consider that one of the advantages of this system is its capability of responding to the change of sound field condition such as unstable temperature gradient on propagation path etc, at real time automatically.

Fig. 3 Simulation Results of Direction Controlled Microphone Frequency: 1 KHz

Fig. 4 Power Spectrum of mixture of reference and sound signal.

Fig. 5 Power Spectrum of demodulation signal
BINAURAL SOUND SIMULATION OF CONCERT HALLS BY A BEAM TRACING METHOD

J. MARTIN, J.P. VIAN

Centre Scientifique et Technique du Bâtiment
24 Rue Joseph Fourier, 38400 Saint-Martin-d’Hères, France

INTRODUCTION

These recent years, many works have been carried out on generation of binaural signals and auditory spaciousness [1], [2], [3]. Auditory spatial impression is the concept of the size and volume that a listener attributes to the hall when he is exposed to a sound field in this hall. This is a very important factor in predicting the acoustical quality of halls. Here we propose a new method to calculate the binaural impulse responses at the two ears of a subject seated in a concert hall. This approach is based on the geometrical cone method developed by the C.S.T.B. in the EPIDAURE program [4] which have been proved to be very valuable in predicting the objective room acoustics criteria for speech or music. This extension of EPIDAURE provides the two impulse responses between the source and the ear’s canal.

Listening tests of halls, even not yet existing can be organised by convoing these responses with dry music.

DESCRIPTION OF THE METHOD

The generation algorithm of the binaural response must take into account many parameters such as:
- the hall geometry and the acoustical characteristics of the different walls materials,
- the different paths which arrive at the two ears including the starting angles from the source, the surfaces on which the beam is reflected, the arrival angles on the head and their associated angles,
- the impulse responses of left and right ears versus the arrival angles of the beam.

The principle used here consists in expressing these physical operators in terms of elementary linear filters by mean of signal processing tools. Then we can express the binaural response from the monophonic one predicted by EPIDAURE program.

Figure 1 gives the general block diagram of the proposed algorithm.

![Figure 1 - General block diagram of the algorithm](image)

253
The different parts of the algorithm are:

1. Epidaurë Program

   The Epidaurë program is composed of several modules which allow to calculate the impulse response of a hall and related criteria [4]. Here two of these modules are used:

   The first permits the entry of geometrical data set describing the room, its materials, source and auditor positions. Then a cone tracing program produces simplified impulse responses at specific auditor locations so that individual reflection paths can be identified with all their propagation characteristics. The maximum reflection order which can be used is equal to eight and the program usually runs with an number of beam between 10 000 and 40 000.

2. Ebinaur Program

   As seen in figure 1, two kinds of filters must be used. One like source, materials and air-absorption filters are computed, the others, like ears filters are measured.

   Source and material filters are built in the same way. We usually know their spectral characteristics by octave bands. The first step consists in interpolating these values to obtain a function for all discrete frequencies. We use for that cubic–spline interpolation. Let \( h(f) \) denotes the interpolated function, where \( f \) denotes the frequency. \( h(f) \) is a real function without any phase component. That corresponds in the time domain to a filter which is not real and causal (the impulse response exhibits non-zero values for negative time). So \( h(f) \) must be modified such that spectral characteristics remain the same but the corresponding impulse becomes real and causal. In other way, it means that a phase \( \phi(f) \) must be found to verify this objective. Then, the result will be written as follows:

   \[
   \hat{h}(f) = h(f) \ e^{i\phi(f)}
   \]  

   (1)

   We assign \( h(f) \) to be the transfer–function of a minimum phase filter which guarantees the causality constraint [6]. This can be obtained by the logarithmic formula of BAYARD–BODE relations which provide the following phase component:

   \[
   \phi(f) = \text{HT} \left[ \log |h(f)| \right]
   \]  

   (2)

   where \( \text{HT}[] \) denotes the HILBERT transform:

   \[
   \text{HT} \left[ x(f) \right] = \frac{1}{\pi} \int \frac{x(f')}{f-f'} \, df'
   \]  

   (3)

   This expression becomes very simple in the time domain:

   \[
   \text{HT} \left[ x(f) \right] \xrightarrow{\text{FOURIER TRANSFORM}} -i X(t) \ \text{sgn}(t)
   \]  

   (4)

   where \( X(t) \) denotes the FOURIER transform of \( x(f) \) and \( \text{sgn}(t) \) the sign pseudo–function. These relations remain true for discrete time system. So, having \( \phi(f) \), \( h(f) \) is deduced from equation (1):

   \[
   \hat{h}(f) = \hat{h}(f) \ e^{-i\phi(f)}
   \]  

   (5)

   We use the same technique to build the air absorption filter but \( h(f) \) is directly obtained by the formula given in [6].

   As previously stated the ears transfer function filters are measured. This measure has been performed by a Time Delay Spectrometry (T.D.S.) method implemented on a personal computer. The T.D.S. is based on frequency sweeping [7] which allows to separate one peculiar path (here
between the loud-speaker and the ears) and to obtain the corresponding transfer function. Figure 2 shows the measuring system:

**FIGURE 2**: measuring system of ears transfer functions

The transfer functions are estimated versus several angles of incidence (azimuth and elevation) of the direct sound on the auditor head. Figure 3 shows the angular sampling step that we have chosen.

**FIGURE 3**: Successive locations of the loud-speaker on a quarter of hemisphere. The whole hemisphere have been explored

The receiver was composed of two electret microphones capsules placed in the ear canal entrance. The sample frequency rate was 20 kHz and the number of time samples 512.
For each path \( i \), with delay \( \tau_i \), the product of the successive transfer function (source) \( x \) (reflection on surfaces) \( x \) (air absorption) \( x \) (left ear : right ear) is calculated and an inverse FOURIER transform gives the two impulse responses of path \( i \) \( L_i(t) \) and \( R_i(t) \). The global responses are the summation over all of the paths contributions:

\[
L(t) = \sum_i L_i(t - \tau_i) \\
R(t) = \sum_i R_i(t - \tau_i)
\]

LISTENING is obtained by convolving (overlap-add-sectionning technique) a dry recording with the two impulse responses.

RESULTS AND CONCLUSION

Giving written conclusions on something devoted to listening is very hard to do. Meanwhile we present here the result of some investigations performed with the program.

First, we simulated an anechoic chamber with a moving source for a fixed auditor location. From this test, it appears that the azimuthal localization is quite good whatever the elevation angle.

The elevation localization is good except when the azimuthal angle becomes close to 0° or 180° as it usually happens with dummy head recording.

Then the auditor is placed in a semi–reverberant room. Several listening tests have been performed when changing the room size. The spatial impression is quite coherent with the volume of the room or the distance between the source and the auditor.

The last tests are still under analysis. They consist in comparing listening in a real hall with predicted listening of the same hall. The results promise to be very positive.

This program gives a very tractable method to assist conception or modification of a hall. In addition, it is a very valuable tool to study the auditory spaciousness effect and to create artificial sound field.

BIBLIOGRAPHY

Generation of binaural signals for research and home entertainment* ICA 86

[2] H. Eil and J. Blauert "A measuring system for acoustic scale models" ICA 86 (Toronto)


256
GEOMETRICAL ACOUSTICS DESIGN SOFTWARE

Tzekakis E.*

*Laboratory of Architectural Technology, School of Architecture, Aristotle University of Thessaloniki, 54006, Thessaloniki, Greece

Summary

The geometry of a room is a complex room acoustical parameter whose control may well be the subject of a suitable microcomputer software application.

Using a machine with advanced graphics tools, an application has been developed that enables the easy drawing of room sections, vertical or horizontal ones. Once drawn, such a section may be used to study the reflection pattern of a ceiling or a side wall using rays and to change the position and/or the inclination of reflecting surfaces. In addition, the application calculates and draws a reflected energy distribution diagram that serves as a qualitative criterion.

In this paper, the development of such an application on an Apple machine is described.

1. Introduction

The geometrical design of a room is a very significant and complex factor, that plays an important role in its acoustical performance. Although geometrical acoustics considerations are important only to the middle and higher frequencies, the directness of the geometrical design acts as a positive offset and makes it a very commonly used tool.

A room is a three-dimensional object with an equivalent sound field. Although

Figure 1: Example of the results to be expected from applying the program to a typical case of a cross section of a small auditorium.

the final acoustical design of any significant room is performed in three dimen-
Figure 2: Example of a typical screen layout in the drawing mode, showing the menu bar above, and the room's section with reflectors and source position.

sections, a large part of the actual work is prepared in two dimensional drawings, mostly two dimensional cross sections of the room. In addition, for many smaller rooms and simpler cases, the consideration in two dimensional sections is simply enough.

By using computers to implement geometric acoustical design, besides making the whole process easier to perform, one can use the same data to check on other acoustical parameters (echoes, excess time delays), to control the distribution of the reflections across the audience plane, or to analytically describe the final design (reflector's coordinates). It is evident that such an application must be implemented in a machine with an advanced graphics environment, using an elaborate and easy user interface.

The aim of geometric acoustical design is to ensure that the room's shape facilitates the even distribution of reflected sound energy. This is mostly done through the suitable positioning of sound reflectors. The usual way to do this, is to draw a section of the room, mark the audience area, introduce reflectors in places that are probably suitable,
select a source position and draw the 'sound rays' that start from the source, and after their reflection on the reflector surface reach the audience area. In most cases a small number of reflectors is studied concurrently and a series of changes in the dimensions, the shape or the angle of inclination of the reflectors, is introduced in order to distribute reflected energy in all parts of the audience area, or to make the distribution more even. And this is exactly what the discussed application is about.

2. Program description.

In order to make these design steps possible to implement in a Computer Aided Design environment, the application has two main parts, the drawing part and the executing part. Of course both parts are implemented using a third hidden part that covers the presentation of the room while it is being drew, as well as the presentation of the results of the program execution. The second, executing part of the program is always accessible from the first part, as long as one source and at least one reflecting surface have been defined.

The first part of the program has two modes, the drawing mode and the selecting mode. Both help in the preparation of the drawing of the room, that is equivalent to the introduction of the data.

In the drawing mode a line drawing tool is available that gives the current coordinates in meters. A hidden autogrid provides an easy way to produce a drawing with all its lines connected. The drawing is performed in a drawing surface equivalent to DIN A3. Only part of this surface is visible at all times and the movement of the surface is done through scrolling horizontally as well as vertically. A different scale may be selected to help either viewing the whole drawing (x0.5) or one particular detail (x2 or x5). Line position, line thickness as well as coordinates are correct in all scales. In this mode a source may be positioned or repositioned at all times. The program translates this graphic input into real coordinates and feeds them into a room record structure. This record can hold information for up to 500 lines, and can be saved at any time under any name or can be used to update an earlier saved version. Saved records may be also recalled at any time.
The other mode needed to complete the design of the room is the selecting mode. Using a pointing device (arrow), a line can be selected in order to be either changed or better defined. A selected line may be moved in all directions, made longer or shorter, or turned in either direction, until it conforms to the needs of the design. In addition a selected line may be defined as an absorber, meaning that any sound reaching it will be stopped, or as a reflector, meaning that this line must be sprayed with sound rays in order to show the resulting reflections. Finally a selected line may be deleted.

Equipped with lines (sections of surfaces) with suitable absorbing or reflecting attributes in a complete drawing, and a source, the application may be executed using commands that define the sound rays' density (from trace only, to dense). Working with the application is mostly seeing the result of the execution and returning to the drawing mode in order to introduce changes or corrections. This interactive process continues until the user is satisfied that the room performs well, the actual criterion being experience based and qualitative.

The program presents a white drawing surface (a window) that can be moved around the computer screen and made larger or smaller. The scrolling facilities are automatically adjusted in all cases. All commands needed for the use of the program are presented in a menu line in the upper part of the screen, that can be made active by pointing to it with an arrow-like cursor moved by a mouse. This cursor is transformed into a cross while drawing lines. In order to write the program a high level structured language was used (Pascal) together with graphics tools available from the proprietary system of the actual machine (Apple Macintosh).

3. Conclusions.
This program is one of a series of similar applications that help in the acoustical design of rooms (reverberation time prediction, traffic noise prediction etc), forming a set of simple expert design generators, being at the same time very powerful teaching aids. Although they are stand alone applications, they are very easy to transform into background programs, to work with a foreground CAD application, instead of the pseudoCAD environment used by the described application to facilitate the graphic data input. And it is felt that this will be the sensible way to introduce acoustical consultation into the design of almost all kinds of buildings, which is an impossible task if one thinks in terms of human consultants.

260
EXPERIENCE WITH A NEW SYSTEM OF ACOUSTIC CONTROL

D. de Vries, A.J. Berkhout

Delft University of Technology, Faculty of Applied Physics,
P.O.B. 5046, 2600 GA Delft, The Netherlands

INTRODUCTION

Recently, a new system for acoustic control of auditoria (ACS) has been developed at the acoustics laboratory of the Delft University. The system aims at reconstructing the reflected sound field in an 'ideal' hall within an existing hall with acoustic shortcomings, according to the principles of acoustic holography. A description of these principles are given in another paper presented at this congress [1]. In this paper, some applications of ACS will be discussed.

CHURCHES

The first Acoustic Control Systems have been installed in several churches in The Netherlands, where lack of reverberation gave rise to complaints about the performance of the organ. Here, the major goal was to increase the reverberation times up to 3 seconds or more and to raise the reverberant level correspondingly, without causing any noticeable coloration of the sound. This criterion was easily met in all cases. For an example of reverberation characteristics obtained, see fig. 1.

![Reverberation characteristics in Petra Church, Rijnsburg (The Netherlands).](image)

Figure 1: Reverberation characteristics in Petra Church, Rijnsburg (The Netherlands).

Note that, when the microphones of ACS are selectively directed to the organ pipe front, spoken word is not significantly 'reverberated' and maintains its intelligibility.
THE DELFT UNIVERSITY SYSTEM

The auditorium of Delft University was the first in a series of acoustically 'dry' halls originally designed for speech (lectures, drama), where ACS was applied to provide variable acoustics, i.e. to optimally adapt the acoustical conditions to other hall functions as well. However, the most important role of the Delft system is that it is used for research, being a part of the acoustics laboratory. The system consists of two modules processing in real time the sound recorded by 24 microphones over the stage (see fig. 2):

1. a module to generate early reflections on the stage and in the front part of the hall through 12 loudspeakers;
2. a module to generate reverberation in the hall through 24 loudspeakers.

![Diagram of the Delft University auditorium with ACS](image)

*Figure 2: The Delft University auditorium with ACS.*

Both modules have 20 parameters to be varied, as indicated in fig. 3. This way, a wide variety of sound fields can be adjusted to be analyzed by measurements and by listening tests. We mention here some of the results of our investigations - to be discussed and documented during presentation:

1. As the input signal of ACS, the direct sound of the source (e.g., the orchestra) should be recorded, however not so closely that usual movements of a player are perceived as acoustical changes. Therefore, directive microphones should be arranged over the stage area such that each instrument (or group of instruments) is covered by at least two microphone main lobes. Depending on their directional properties, the
microphones should be positioned at a height of about 0.5 - 0.8 times the omnidirectional 'reverberation radius' of the hall; this distance is not very critical, and may well be adapted to sightlines, projector lightbeams etc.

2. The increase of reflection density with time is an important item, especially for the 'natural' perception of high frequencies. Further research should clarify which is the 'critical' density increment. In this context it is worth while to realize that in most halls the sound field is more or less two-dimensional due to the absorbing audience area, corresponding with linear increase of reflection density with time instead of the quadratic increase in the 3D shoebox-model found in all textbooks!

3. With ACS, the slope and level of a decay curve can be adjusted independently, such that the increase of RT with by factor of 2 does not necessarily yield a level increase of 3 dB, as in 'classical' solutions. Experiments with varying levels given a certain reverberation time show that many people - musicians as well as listeners - appreciate a relatively low level increase, such that also in the rear seating area good clarity of music and intelligibility of speech is maintained.

A system fully identical to the Delft set-up is in use at York University in Toronto, Canada, in the faculty of music, where research is done twinned to the Delft program. [2]

Figure 3: ACS module parameters.

MULTI-FUNCTIONAL HALLS
Meanwhile, the system has been installed in several multi-functional theatres having side-stages and flytowers. Here, an ACS stage module is applied to replace the orchestra shell with its manpower-intensive
build-and-remove procedures. Loudspeakers should have well-specified
directional properties in this case, since the musicians should be well-
addressed with early reflections, but not the ACS input microphones.
the performance of ACS is not based on acoustical feedback. Besides an
optional module generating early lateral reflections in the hall, a
reverberation module provides variable reverberation times and levels.
Since, in smaller auditoria, the direct level produced by a large orchestra
is already fairly high, the reverberant level created by ACS should be
controlled such that the total loudness perceived will be kept acceptable.
The systems have up to 8 preprogrammed settings in memory which are
optimal for different hall functions, to be recalled with a simple
pushbutton device.

STUDIOS
ACS has also been applied in a large broadcasting studio, where -
apart from a balcony for the audience during live recording sessions -
'stage' and 'hall' area coincide. Here, a double array of loudspeakers has
been mounted on the walls, a lower one to generate well-directed early
reflections to the performers, a higher one to generate diffuse
reverberation. 16 Preprogrammed settings are available to choose the
optimal acoustical conditions for a wide range of musical scores and
styles, at the benefit of the performers, and of the recording engineers as
well.

REFERENCES
speech reinforcement, to be presented at 13th Int. Congress on
reverberation preferences for acoustical instruments, paper presented
A HOLOGRAPHIC APPROACH TO SPEECH REINFORCEMENT

P. Vogel, A.J. Berkhout, D. de Vries

Delft University of Technology, Faculty of Applied Physics,
P.O. Box 5046, 2600 GA Delft, The Netherlands

ABSTRACT
In common speech reinforcement practice, preservation of true source localization is often a weak point, since this requires that the direct speech signal precedes the amplified loudspeaker signal. Finding the appropriate time delays for all listener positions simultaneously is a cumbersome task, especially in case of moving sources (e.g., cabaret, drama). However, speech enhancement with full maintenance of localization can be obtained by using the principles of acoustic holography. Here, the direct signal of the sources is recorded with an array of microphones and reradiated to the audience with an array of loudspeakers after processing such that the emitted sound field is spatially in phase with the direct field. In this paper, the design of a holographic speech reinforcement system will be discussed, as well as preliminary experience with such a system.

I. INTRODUCTION
Acoustic holography is a technique that reconstructs from measurements the full wave field at any desired point in space. Any holographic system can be described by three subsystems (fig. 1):

\[ \text{source} \rightarrow \text{incoming wave field} \rightarrow \text{processor} \rightarrow \text{reconstructed wave field} \]

*Figure 1: Principle of acoustic holography. If } x_2 = x_1 \text{ then } W \text{ equals a unity operator.}*

1. Microphone array
The individual microphone signals of the array define the source wave field at the position of the microphone array. Array length (truncation) and microphone distance (sampling) should be properly chosen.
2. Extrapolation subsystem

The extrapolation subsystem is a wave theory based signal processor that simulates propagation from the position of the microphone array towards the position of a loudspeaker array.

3. Loudspeaker array

The extrapolated microphone signals are fed into the individual loudspeakers of the loudspeaker array. After re-emission the resulting wave field resembles the true wave field as well as possible. It propagates from the loudspeaker array further in space in a "natural" way.

For direct sound control the simulation process in the extrapolation subsystem occurs in free space. For reverberation control simulation occurs in an enclosed space with desired boundary conditions. If the microphone signals are stored, then extrapolation and re-emission may occur at a later time (reproduction).

II. THEORETICAL ASPECTS

If an array of microphones measure the pressure of a propagating wave field in the plane \( x=x_1 \),

\[ p = p(x_1,y,z,t) \]

at each position, and the microphone signals are fed to related loudspeaker positions, then the resulting wave field can be made identical to the original wave field for \( x>x_1 \). Mathematically (Berkhout, 1987),

\[
P(x,y,z,\omega) = \iint_{S_1} P(x_1,\eta,\varsigma,\omega)(1+jkr)\cos\phi \frac{e^{-jkr}}{r^2} \ d\eta \ d\varsigma \quad (1a)
\]

where \( P(x,y,z,\omega) \) is the frequency domain presentation (Fourier transform) of pressure distribution \( p(x,y,z,t) \) and

\[ x > x_1 \]
\[ k = \omega/c \]
\[ c = \text{sound velocity} \]
\[ r = \sqrt{(x-x_1)^2 + (y-\eta)^2 + (z-\varsigma)^2} \]
\[ \cos\phi = (x-x_1)/r. \]

Note from Eq. (1a) that for a perfect reconstruction the loudspeakers should have a dipole characteristic. In practical situations the integral (1a) should be replaced by a summation,

\[
P(x,y,z,\omega) = \sum_{mn} P(x_1,y_m,z_n,\omega)(1+jkr_{mn})\cos\phi \frac{e^{-jkr_{mn}}}{r_{mn}^2} \Delta y\Delta z
\]

265
or

\[ P(x, y, z, \omega) = \sum_{mn} W(x-x_1, y-y_m, z-z_n, \omega) P(x_1, y_m, z_n, \omega) \]  

where

\[ W(x-x_1, y-y_m, z-z_n, \omega) = (1+jk_{mn}^r \cos \phi_{mn}) e^{-jk_{mn} \Delta y \Delta z} \]

with

\[ r_{mn} = \sqrt{(x-x_1)^2 + (y-y_m)^2 + (z-z_n)^2}. \]

Eq. (1b) represents a spatial convolution along y and z.

According to sampling theory an exact reconstruction can be made if \( \Delta x = \Delta y = \Delta z / 2 \). Hence sampling should be denser for higher frequencies. Perception experiments have taught us that the sampling intervals can be chosen coarser, particularly in reverberation control.

In the generalized version of acoustic holography, the measured microphone signals are fed into a processor and propagation ("extrapolation") to another plane, say \( x=x_2 \), is carried out numerically (Fig. 1),

\[ P(x_2, y, z, \omega) = \sum_{mn} W(x_2-x_1, y-y_m, z-z_n, \omega) P(x_1, y_m, z_n, \omega) \]  

where \( x_2 > x_1 \).

Next, re-emission into space ("reconstruction") occurs by an array of dipole loudspeakers in the plane \( x=x_2 \),

\[ P(x, y, z, \omega) = \sum_{mn} W(x-x_2, y-y_m, z-z_n, \omega) P(x_2, y_m, z_n, \omega) \]

for \( x>x_2 \). Control of direct wave fields as well as reflected wave fields should be based on holographic expressions (2a) and (2b).

In summary, by measuring the pressure of an incoming wave field with an array of microphones and applying to the microphone signals an extrapolation process, the wave field can be fully reconstructed everywhere in space by an array of loudspeakers. The listener would not notice the difference between the real field and the reconstructed field since recording, extrapolation, and reconstruction can be made as perfect as desired.

III. DIRECT SOUND CONTROL

Today most electroacoustic systems aim at direct sound enhancement. Basically, one or more directional microphones pick up the direct sound and, after some frequency-dependent amplification, re-emission occurs by a cluster of loudspeakers which direct the amplified sound to the audience. Optionally a delay may be involved to compensate for the difference in
position between microphones and loudspeakers. In many practical situations the source is perceived at the wrong location: intelligibility is improved at the cost of localization.

By using holographic arrays, intelligibility as well as localization can be handled properly: the wave front of the direct wave field is fully reconstructed at any moment of time with the correct phase spectrum and with a desirable frequency-dependent enhancement.

To explain the holographic principle for direct sound control consider the two-dimensional situation in Fig. 2. In front of the source area a distribution of microphones is mounted to measure the direct wave field at \( x = x_1 \). Farther away, at \( x = x_2 \), we want to reconstruct the direct wave field by a distribution of loudspeakers. Using the two-dimensional version of Eq. (2a), extrapolation from \( x = x_1 \) to \( x = x_2 \) is given by

\[
P(x_2, y, \omega) = \sum_{m} W(x_2 - x_1, y - y_m, \omega)P(x_1, y_m, \omega)
\]  \hspace{1cm} (3a)

where

\[
W(x_2 - x_1, y - y_m, \omega) = \sqrt{jkr_m} \cos \phi_m e^{\frac{-jkr_m}{\sqrt{r_m}}}
\]  \hspace{1cm} (3b)

and reconstruction by the dipole loudspeakers at \( x = x_2 \) is given by

\[
P(x, y, \omega) = \sum_{m} W(x - x_2, y - y_m, \omega)P(x_2, y_m, \omega)
\]  \hspace{1cm} (3c)

for \( x > x_2 \).

![Figure 2: Holographic direct sound control: acquisition by microphone array near the stage, extrapolation by signal processor, and reconstruction by loudspeaker array near the audience.](image)

In Eq. (3b) the factor \( \sqrt{jkr_m} \) gives a phase shift of \( \pi/4 \) and an enhancement of the high frequencies of 3 dB per octave, the term \( \cos \phi_m \) causes attenuation, and the term \( e^{\frac{-jkr_m}{\sqrt{r_m}}} \) represents travel time.
Fig. 3 shows the reconstructed direct sound field for different positions of the sound source. Note the important property that the reconstructed wave fronts are consistent with the position of the source.

Figure 3: Direct sound field after acquisition, extrapolation, and reconstruction for different positions of impulsive sound source.

In practical situations the array of microphones may be mounted in front of the stage and the array of loudspeakers may be mounted everywhere in the hall, at the ceiling, at the edge of a balcony, and so on. In addition, similar to conventional techniques, it is advantageous to choose the loudspeakers directional perpendicular to the line array and directed to the audience.

Experiments in our an-echoic room confirm the theory.

IV. CONCLUSIONS
A new approach to direct sound control has been represented in terms of acoustic holography. The holographic approach allows enhancement of the direct sound without loss of localization. In addition, the holographic approach opens new ways in sound reproduction.
V. REFERENCES


"THE DSP 610 - A COMPACT PROCESSOR TO UTILIZE A NEW DIRECTIONAL SOUND REINFORCEMENT SYSTEM, THE DELTA STEREOPHONY SYSTEM"

Dipl.-Ing. Wilhelm NADLER
AKG

For the "Delta Stereophony System", a multi-channel controllable delay unit has been developed. The paper describes a compact device with integrated digital signal processing.

The "Delta Stereophony System":

This special sound reinforcement system has been introduced for the first time at the 6th Acoustical Conference in Budapest in 1976 (Huey, Steffen, Steinke, Reichard, Ahnert: Ein Schallübertragungssystem zur richtungsgtreuen Beschallung größer Auditorien). The described method solved a number of problems found in conventional sound reinforcement systems. Sound reinforcement systems in large theatres or multi-purpose halls are able to suppress disturbing echoes using distributed loudspeaker systems in connection with time delayed signals. But the localization and the sense of depth are usually lost. Only when using a specially applied delay technique - the "Delta Stereophony System" - the localization and the sense of depth of the original sound on stage can be reinstated.

Based on the fundamental law of the first wavefront, the delay times for the individual loudspeakers or loudspeaker groups are adjusted in such a way, that the original sound or its simulation is arriving first at a particular place in the auditorium. If all remaining signals from all the other loudspeakers or loudspeaker groups stay within certain level and time limits, the localization cues are kept in focus to the optical cues from the original sound source (Fig. 1).

For more than one sound source and various distributed reception areas throughout the auditorium, different delay times and signal levels are required for each input signal, as well as a matrix connection between inputs and the loudspeaker groups has to be established. The calculation of the required delay times and signal levels is realized in a calculation program optimized for

![Diagram of sound system](image-url)
various significant reception areas. This procedure guarantees
perfect localization and image of depth over a wide area in the
auditorium. Additional advantages of this system are high sound
quality, even distribution of signal level and the possibility to
raise the sound level for the reception areas without losing
localization cues.

The Delta Stereo Compact Processor DSP 610

The realization of the "Delta Stereophony System" with
discrete delay units and connection matrix requires a substantial
investment in hardware and installation work. It will also miss
important features and possibilities of a perfect "Delta Stereophony System".

AKG has developed a compact Delta Stereo Processor to fulfill
all requirements asked from such a system and match it with fairly
simple operational features. Due to its compact size, but last not
least because of the operational aspect, the Delta Stereo Compact
Processor DSP 610 is well suited for small and medium sized
theatres and halls. The DSP 610 may be also expanded for larger
venues providing a system with similar operational features.

The Delta Stereo Compact Processor contains a digital matrix
with six inputs, each with its own delay network and possible
connection to one or more of the ten output channels. The delayed
signal channels may also be attenuated, if required (fig. 2).

The features:
- matrix with six delayed input to 10 outputs
- additional four non-delayed (direct) inputs
- adjustable delay times and (optional) attenuation for each
  input/output matrix connection
- each matrix connection and each input or output channel may be
  individually switched off
- computer control of delay times for moving sound sources, e. g.
  wireless microphones
- level indication for each input and output channel

The control of the DSP 610 is established via a standard RS 232
interface with an IBM AT-compatible computer system.

The Block Diagram of the DSP 610 (fig. 3):

The analog input signals are pre-amplified in the input
stages having two transformer balanced inputs each.
The conversion to the digital format takes place in the following input modules. They contain next to the sample/hold circuits and the 16 bit A/D converters a floating decimal point encoder to achieve the data format of 16 + 3 bit and consequently obtain high quality signal processing. The digital samples in the format of 16 + 3 bit are stored in EPROMs and read out with different address words. These different address words lead to the required delay times (max. 520 ms in 20 us steps). The delayed samples in the format described above are further converted into a 24 bit, fixed decimal point format for digital computation. The 24 bit digital words are led to a centralized computing stage via the common databus 1 which connects all the input modules.

The delayed signals from the input channels are then digitally added to the output channel, but may also be altered in level before this procedure by a computer controlled 24 bit multiplier circuit. The ten output modules are connected to the data bus 2 which carries the data for the D/A conversion. Each output module has its own summing point to which undelayed analog input signals may be mixed. The digitally controlled level attenuator allows the storage of the output levels. A digital level indication circuitry monitors and displays the signal level in 6 dB steps of inputs and outputs.

The Software:

All parameters within the DSP 610 are software controlled by an MS-DOS computer (IBM-AT or compatible) via a serial interface. The most important functions of the software are:
- calculation of the individual delay times and attenuation levels
- using a pre-programmed algorithm
- control of the delay times values according to the moving sound source location followed by the "computer mouse" on the monitor
- storage and recall of complete data sets to adapt to changed or alternate system/stage layouts.

273
The actual stage situation may be schematically displayed on the monitor screen for live operation of the system (Fig. 4).

Moving sound sources are displayed as specially marked cursors. These cursors will follow directly the movements of the operated "mouse".

The Application of the DSP 610 in Sound Reinforcement Systems:

The base for the correct function of the "Delta Stereophony System" is the calculation of all room and system specific delay times and attenuation levels to achieve the truthful sound image transfer typical of this kind of sound system. The calculation is based on appropriate and optimized program algorithms, considering besides room data also specific system data, like frequency response, radiation response and other loudspeaker characteristics. Also important are the positions of the sound sources or their simulation. Fixed sources will lead to a fixed set of delay times, while moving sound sources require continuous updating of all delay times within the computer and will depend on the appropriate cursor position simulating the actual source position.

All necessary data to operate the system are available within the computer once a particular setup has been initialized and set. The DSP 610 will be usually wired between the group outputs of a mixing desk and the main amplifier inputs for the loudspeakers or loudspeaker groups.

The software of the DSP 610 contains useful features for the operator: controlling the moving sound sources during a live performance, e.g. mouse control or control of pre-programmed sequences.

Please Note:

The application of the "Delta Stereophony System" is tied to an "Operation Licence" authorized by the holder of the relevant patents which is included when using a DSP 610.
QUALITÉ ACOUSTIQUE, ISOLATION ET INTIMITÉ DES IMmeUBLES D'HABITATION EN COPROPRIÉTÉS

Christian Martel et Jean-Gabriel Migneron

Laboratoire d'acoustique, Centre de recherches en aménagement et en développement (CRAD), 1624 pavillon Félix-Antoine-Savard, Université Laval, Ste-Foy, Québec, G1K 7P4

La part des condominiums devient de plus en plus importante dans le marché résidentiel de la région de Québec. L'étude présentée ici porte sur la qualité acoustique de ce type d'habitation, elle est extraite d'un rapport préparé pour le compte de la Société canadienne d'hypothèques et de logement[1].

Échantillon

Nous avons tenté de constituer un échantillon proportionnel au pourcentage d'unités en condominium en fonction du type de construction, de façon à obtenir une représentation la plus exacte possible de la situation dans la région de Québec. Ainsi, 45% des immeubles (9 projets) que nous avons étudiés sont en béton, selon le système dalle et colonnes; la dalle étant d'une épaisseur de 200 à 250mm. Les autres 35% (7 projets) représentent les constructions à ossature de bois, toutes avec une dalle de béton de 38mm sur chacun des plans. Les derniers 20% (4 projets) expriment la proportion des immeubles à structure d'acier, une moitié étant construite en "steeldeck" (dalles de béton habituellement de 100mm coulées sur un profilé d'acier et appuyées sur des poutrelles) et l'autre en système "Hambro" (dalles de béton de 75 à 100mm formant un système composé avec les poutrelles).

Ce sont les immeubles en voie d'achèvement ou encore vides qui ont été testés. Ceci nous a permis un accès beaucoup plus facile aux logements, une obtention rapide des plans et surtout des précisions sur les modifications apportées en cours de construction. Grâce à cette étude sur des immeubles récents, nous avons acquis une vision réelle de la situation présente, avec les techniques de construction et les matériaux d'aujourd'hui.

Nous avons mesuré l'isolation acoustique selon les normes ASTM tout en prenant en considération la perception des usagers. Nous avons ainsi cherché à déterminer l'ampleur du problème causé à un copropriétaire par les bruits aériens et d'impact provenant d'un autre résident ou d'un système mécanique quelconque à l'intérieur même de la construction.

Comportement des murs et des planchers aux bruits aériens[2,3]

Des mesures d'isolement aux bruits aériens ont été effectuées dans les deux sens pour les murs mioyens et les planchers. Les murs ont obtenu une moyenne assez satisfaisante de FSTC 53[4] quoique légèrement en dessous du niveau STC 55 recommandé par la SCHL[5]. Indépendamment du type de construction, 38% des murs se sont classés au-dessus de STC 55, 30% ont été inférieurs à STC 50 tandis que, probablement à cause de mauvaises conditions de mises en œuvre des techniques et des matériaux, 9% des murs mioyens n'ont pu atteindre une performance de STC 45 exigée par le Code national du bâtiment[6,7,8]. Le maximum de dispersion de 5 (STC) des mesures prises dans les deux sens s'explique du fait que le mur n'est pas nécessairement symétrique et les locaux récepteurs ont une constante acoustique (absorption) différente causée par des appartements complètement meublés ou absolument vides et sans aucun revêtement de plancher.

On constate que dans les constructions en bois, les isollements aux bruits aériens semblent excellents pour plusieurs projets quoique les meilleures performances proviennent surtout de murs couplés. Les murs mioyens des constructions à ossature d'acier sont toujours un élément de remplissage dont l'efficacité n'a pas de rapport avec aucune fonction structurale. Pour les immeubles en béton, les murs porteurs coulés sur place ne sont pas une garantie d'une bonne isolation; normalement l'isolement devrait être proportionnel à la masse mais il faut compter une épaisseur suffisante, de plus, on note des pertes significatives dans les basses fréquences autour de la fréquence critique.
Figure n°1 : corrélation entre les différents murs mitoyens en fonction de l'épaisseur

En étudiant la régression que l'on peut établir entre l'indice d'isolation observé et l'épaisseur totale des murs mitoyens mesurés, on peut conclure que l'épaisseur du mur joue un rôle appréciable quant au résultat escompté en matière d'isolation, surtout que dans la plupart des cas on a affaire à des parois multiples qui comportent une ou plusieurs lames d'air d'épaisseurs variables. En utilisant la masse surfacique, la corrélation devient moins bonne, on peut toutefois mentionner que la dispersion des résultats est plus grande et que la multiplication des couches plus ou moins indépendantes dans une paroi n'est pas la garantie d'une élévation de l'isolation par rapport à celui d'une paroi homogène.

Les difficultés de mise en œuvre sur le chantier demeurent un facteur crucial dans l'insonorisation d'une paroi. Ainsi, d'après ces résultats pratiques, pour obtenir un STC 55, il faudrait un mur de plus de 63kg/m² et de plus de 280mm d'épaisseur. De même, pour atteindre un niveau idéal de STC 60, il faudrait un mur mitoyen de plus de 370mm d'épaisseur et plus de 170kg/m².

Figure n°2 : distribution des indices d'isolation au bruit aérien des murs (propagation directe seulement) suivant le type de construction

On constate avec satisfaction que les isolements des planchers sont en général supérieurs à ceux des murs mitoyens. Avec une moyenne de près de STC 57, pratiquement aucun condominium de la région de Québec n'a un niveau inférieur à STC 50. La légère différence favorable au sens inverse de propagation est due à la dissipation du bruit dans la partie légère du plafond avant que l'énergie ne frappe la structures même. Ce phénomène prend d'autant plus d'ampleur que plusieurs planchers au moment des tests n'étaient pourvus d'aucune finition pour atténuer le bruit incident.
Les planchers dans la construction à ossature de bois ont un isolement aux bruits aériens variant très peu autour de STC 55. Même si celui de la structure d'acier est comparable, il est plus dispersé. L'isolement des planchers des structures de béton par contre est très facilement affecté par différents problèmes: faute, coincidence des ondes incidentes et vibratoires, etc. Pour les mêmes raisons que pour les murs mitoyens, l'épaisseur du plancher est encore ici plus significative que sa masse superficielle.

Comportement des planchers aux bruits d'impact[9]

Le bruit d'impact ne fait l'objet d'aucune exigence du Code national du bâtiment mais d'une recommandation pour un isolement atteignant 1IC 65 de la SCII[5]. 40% des planchers mesurés se conformaient à cette dernière recommandation; cette tranche incluant tous les tests faits sur tapis et sous-tapis. Beaucoup plus faible, le tapis employé seul offre approximativement le même rendement que la céramique tandis que le linoléum collé directement sur la dalle est encore inférieur, et même très près de cette dernière. Finalement, on ne constate pratiquement aucune relation entre la performance d'un plancher aux bruits d'impact et le type de construction utilisé.

![Diagramme de la perte de transmission en fonction de la fréquence](image)

**Figure n°3:** rendements des différents revêtements de sol aux bruits d'impact sur plancher à ossature de bois et dalle de béton de 38mm

Nous avons aussi effectué des mesures de transmission latérale aux bruits d'impact en procédant de la même façon que la méthode verticale, mais en évaluant le bruit réfléchi dans le logement voisin, sur le même plancher. Bien que cette évaluation ne soit pas normalisée, elle fait toutefois ressortir des points intéressants sur la propagation du son. Très peu d'immeubles ont réussi à atteindre un 1IC latéral de 65. Ceci nous montre le problème caractéristique des dalles continues: le son se propage dans la dalle de béton même loin du point d'excitation. Les mesures vibratoires complémentaires montrent une rémission acoustique de toutes les surfaces, même les cloisons légères et non portées car elles demeurent appuyées sur la structure et surtout sur les dalles de béton. Ici encore, le revêtement de sol joue un rôle déterminant dans l'isolation de la source.

Perception des résidents[10,11]

De façon générale, on note une bonne corrélation entre la satisfaction obtenue et les indices d'isolements mesurés sur place. Lorsque le niveau STC ne fait que respecter le Code national du bâtiment, généralement les propriétaires occupants sont plus insatisfaits et songent sérieusement à se plaindre.
Très clairement, la phase critique au plan perceptif se situe au moment de l'entrée des nouveaux occupants, et pour plusieurs raisons. Très souvent, les nouveaux occupants viennent d'une maison unifamiliale et la proximité immédiate des voisins est difficile à accepter. Aussi, ils semblent dans un immeuble récent et en grande partie inoccupé; le bruit de fond étant très bas, le moindre bruit peut être entendu, même d'un appartement éloigné.

La gêne n'est pas la même suivant que les pertes acoustiques se situent dans les basses ou les hautes fréquences. Dans le premier cas, une fréquence critique mal placée ou une résonance du type masse/air/masse laissera percevoir les basses fréquences alors que dans le second, une simple fuite causée par des ouvertures dans la paroi, une mauvaise étanchéité ou une transmission par les parois latérales (le mur de façade étant souvent en cause) laissera percevoir un mince fillet de conversation.

**Conclusion**

Malgré l'étendue forcément limitée de notre échantillon, la recherche a été très indicatrice de la situation générale en matière d'isolation acoustique pour les condominiums. Bien que les murs mitoyens soient les plus problématiques, une composition de différents matériaux et de lames d'air d'une certaine épaisseur donne d'excellents résultats. La mise en œuvre sur le chantier demeure un problème important. Suivant la surveillance, le rendement sera inférieur[12]. Les planchers se comportent mieux aux bruits aériens mais pour les bruits d'impact ils sont souvent négligés, surtout pour les planchers légers, continus et recouvert de céramique ou de linoleum; à défaut d'y installer un plancher flottant, l'isolation du plafond inférieur ou la pose d'un tapis avec sous-tapis demeurent les meilleures solutions. Malgré l'emploi des murs et planchers sur ossature de bois, d'acier ou de béton, on constate que la construction canadienne a atteint une technicité performante en matière d'isolation acoustique.

**Références**

SOME EFFECTS ON THE LABORATORY MEASURED SOUND TRANSMISSION LOSS

U. Donner

Institut für Technische Akustik,
Technische Universität Berlin,
Einsteinufer 27, 1000 Berlin 10

1. Introduction

Many papers about the effects on the laboratory measured sound transmission loss (TL) have been published in the last years. Most of the papers deal with measurements and only a few authors calculate the influence of physical effects on the TL.

This publication reports of computer simulation for two-dimensional sound fields. The simulation has been done in two dimensions to save time. The effects are the same as in three dimensions.

2. Numerical Approach

The numerical method which is used here is called 'collocation' /1/. The sound field is expanded in a finite sum of eigenfunctions and the coefficients of the eigenfunctions are determined by complying with the boundary conditions at special 'collocation points'. For rigid walls the sound fields in rooms (or ducts) I and II and niches I and II (as shown in figure 1 for rooms) are expanded in sums like

\[ p(x,z) = \sum_{n} \cos \left( \frac{n \cdot \pi}{d} \cdot x \right) \cdot \left[ A_{n} \exp(jk_{z} \cdot z) + B_{n} \exp(-jk_{z} \cdot z) \right] \] (1)

The boundary conditions at the collocation points deal with the pressure \( p \) and the particle velocity in the \( z \)-direction \( v_{z} \) as shown in figure 1. The excitation of the sound field in the room is simulated by a pressure \( p_{a} \) which is coupled to the sound field by the mass \( m_{a} \). The excitation of the sound field in the duct is accomplished by taking all modes, that exceed their cut-off frequency, into account. These modes have the same amplitude. Damping is taken into account in the form of a homogeneous porous absorber with a small flow resistance \( \zeta \).

3. Results

Figure 2 shows the narrow band and the third octave band spectrum of the TL. The geometrical parameters are explained in figure 1. The narrow band spectra are
different, but the third octave band spectra seem to be almost alike. With every new mode the duct's TL decreases due to the high impedances of each mode just above the cut-off frequency. This can't be seen in the room's TL, because the resonances of the rooms dominate this spectrum.

Figure 3 shows the TL with the same geometry as in figure 2 except for the depth of the niches. The narrow band spectrum of the duct's TL shows reductions below the cut-off frequencies of the modes. This can be explained by describing the arrangement as a system of mass (nearfield/excited duct) - spring (niche I) - mass - spring (niche II) - mass (nearfield/receiving duct), which gives reductions of the TL because of its resonances. This effect also can't be seen in the narrow band spectrum of the room's TL. But in comparison to figure 2 both third octave band spectra show a reduction of TL. Measurements and more details of this resonance effect are published in /2/.

Figure 4 shows the TL for the 125Hz - third octave of a geometrical arrangement, which is nearly the same as in figure 1. In this figure the length of the recieving room 2 is varied from half to twice of the room 1. The results are shown for three different reverberation times, which in each case are equal for both rooms. It can be seen, that at \( l_2/l_1 \)-values of 0.5, 1.0 and 2.0 the TL has minima. The reason for this is the coincidence of resonance frequencies in rooms 1 and 2. To avoid these minima, the values of \( l_2/l_1 \) must be changed at least 0.2m at the critical \( l_2/l_1 \)-values. But also at other values of \( l_2/l_1 \) minima of the TL occur.

The variation of the reverberation time shows, that the less reverberation time is the lesser is the variance. But the average value is practically the same in all three cases.

4. Resumee

The coincidence of resonance frequencies in the two rooms of a laboratory decreases the TL as well as the presence of a deep niche does.

To avoid too much variation of the TL when measured in different laboratories, the reverberation times of the rooms ought not to be too long.

Literature


/2/ U. Donner Einfluss der Nischentiefe auf das Luftschallabklingen, Fortschrifte der Akustik, DAGA 1989

280
fig. 1: sketch of the modell with collocation-points

<table>
<thead>
<tr>
<th>excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{an} - p_i = j \omega m_{an} \cdot y_{i2}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>boundary conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{v}<em>{12} = \bar{v}</em>{22} = 0$</td>
</tr>
<tr>
<td>$\bar{v}<em>{12} = \bar{v}</em>{22}$; $p_1 = p_i$</td>
</tr>
<tr>
<td>$\bar{v}<em>{12} = \bar{v}</em>{22}$; $p_2 = p_{ii}$</td>
</tr>
<tr>
<td>$p_1 - p_{ii} = \bar{z}<em>t \cdot \bar{v}</em>{12}$</td>
</tr>
</tbody>
</table>

fig. 2: $b_1 = b_2 = 6m$, $a = 1m$, $x_1 = x_2 = 2.5m$, $l_i = l_{ii} = 0m$, $\Xi = 10Ns/m^4$, $Z_T = j\omega m''$ (m'' = 25kg/m$^2$)

--- room: $l_1 = 8.7m$, $l_2 = 7.3m$ ($T = 1.6s$)
--- duct
--- mass law/45° angle of incident
Fig. 3: Same as in fig. 2 except for $l_1 = l_II = 0.2m$.

Fig. 4: TL between two rooms for the 125 Hz - third octave. 3 different reverberation times and varying $l_2$:

- $b_1 = b_2 = 6.33m$, $a = 1m$, $x_1 = x_2 = 2.66m$, $l_1 = l_II = 0.2m$,
- $Z_n = 25kg/m^2$, $l_1 = 8m$.

Mass law/45° angle of incident: 25 dB.
3.F
STC RATING VERSUS dB(A) - LABORATORY AND FIELD STUDIES

Tibor Vass

Department of Architecture, Curtin University of Technology, GPO Box U 1987
PERTH, 6001, Western Australia, Australia

Introduction

Australian Standard, AS 2107-1987, recommends design sound levels in dB(A) or in special cases e.g. Studios and Theatres in Noise Rating (NR) for building interiors in Health, Educational, Office, Public, Industrial and Residential Buildings.

Sound Insulation (SI) values of partitions separating adjacent spaces in the above buildings are measured according to AS 1191-1985, Method of Laboratory Measurement of Airborne Sound Transmission Loss of Building Partitions, and expressed in a single 'STC' rating, based on AS 1276-1979 Methods of Determination of Sound Transmission Class and Noise Insulation Class of Building Partitions.

At the sketch design stage, when designing a 'complex' building, the Architect's or his/her Consultant's task in the field of acoustic control is to:

a) Identify the various sound sources in each room and calculate or estimate the resultant sound level in dB(A) or NR

b) Identify the various noise sources in 'noisy' rooms, referred to as 'source rooms', then calculate or estimate the total noise levels in those rooms expressed either in 1/3 octave band Sound Pressure Level (SPL) or in dB(A)

c) For rooms adjacent to b) above, select the most appropriate construction/partition with a suitable Sound Insulation value which would reduce the 'source room's level in the 'receiver room' to the recommended values indicated in AS 2107.

When SI of the 'source room' is expressed in dB(A) and if the SI value of a partition could be expressed also in dB(A), (appropriately calculated/corrected), then the Architect's or Consultant's task of selecting a suitable partition at the initial design stage would be made easier.

Experimental Procedures

To examine the practicability of the above proposal, a series of laboratory and field measurements were performed. The selection of partitions for these measurements were based on the followings:

a) Partitions/constructions with no 'coincidence' dip, e.g. masonry constructions, Experiment 1

b) Partitions/constructions with pronounced 'coincidence' dip, e.g. standard framed/sheeted and some panel constructions, Experiment 2

c) Partitions/constructions with no significant 'coincidence' dip, e.g. b) above with absorbent facings (1), Experiment 3.

This paper reports test results for Experiment 1 only.

All laboratory measurements were performed according to an appropriate AS and expressed in the following units:

1. 'STC' rating, based on AS 1191 and AS 1276
2. dB(A), following procedures described in AS 1191.

285
3. $\bar{R}$, average Sound Reduction Index, which is the arithmetic mean of 1/3 octave band SPL from 100 Hz to 3150 Hz.

All field measurements were performed according to AS 2252-1979, Methods for Field Measurement of the Reduction of Airborne Sound Transmission in Buildings and expressed here in $dB(A)$ only.

Experiment 1 - Laboratory Study I

Whilst the recommended precision of measurement quoted in AS 1191 was achieved for a given test using a particular signal spectrum and sound source location in the reverberation chamber's 'source room', a variance of Sound Transmission Loss spectrum for a given wall was observed when the signal spectrum was varied.

To test the significance of signal spectrum variance introduced into the 'source room', the following signal types and the combination of these were used:

a) Pink noise No.1a) and 1b)

b) Pink noise No.2

c) White noise No.1

d) White noise No.2

A single brick wall of 150 mm thick with a surface density of 796 kg/m$^2$, built into the opening between two reverberant rooms, was used for the tests. For results see Table I below.

Note: The source(s) of sound in the 'source room' was located in one of the trihedral corners of the room.

<table>
<thead>
<tr>
<th>Group No.</th>
<th>Test No.</th>
<th>Signal(s) Type</th>
<th>STC rating</th>
<th>$\bar{R}$</th>
<th>$dB(A)$</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>1</td>
<td>Pink noise No.1a)</td>
<td>46</td>
<td>43.9</td>
<td>44.2</td>
<td>No variance of STC, $\bar{R}$ and $dB(A)$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Pink noise No.1b)</td>
<td>46</td>
<td>43.6</td>
<td>44.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Pink noise No.2</td>
<td>46</td>
<td>43.8</td>
<td>44.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Pink noise No.1a)</td>
<td>46</td>
<td>43.7</td>
<td>44.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pink noise No.2</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Group No.1</td>
<td></td>
<td>46</td>
<td>44*</td>
<td>44*</td>
<td>* Figures rounded off</td>
</tr>
<tr>
<td>II.</td>
<td>5</td>
<td>Pink noise No.1a)</td>
<td>46</td>
<td>43.8</td>
<td>46</td>
<td>No variance of STC, $\bar{R}$ but variance of $dB(A)$</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>White noise No.2</td>
<td>46</td>
<td>43.6</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Group No.1</td>
<td></td>
<td>46</td>
<td>44*</td>
<td>45*</td>
<td>* Figures rounded off</td>
</tr>
<tr>
<td>III.</td>
<td>7</td>
<td>White noise No.2</td>
<td>45</td>
<td>43.2</td>
<td>46</td>
<td>No variance of STC, $\bar{R}$ but variance of $dB(A)$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>White noise No.1</td>
<td>45</td>
<td>43.4</td>
<td>46.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>White noise No.1</td>
<td>45</td>
<td>43.2</td>
<td>48</td>
<td>There is also a variance of STC of Group No. III from Groups No.I and II.</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>White noise No.2</td>
<td>45</td>
<td>43.1</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Group No.1</td>
<td></td>
<td>45</td>
<td>43*</td>
<td>47*</td>
<td>* Figures rounded off</td>
</tr>
</tbody>
</table>

The results shown in Table I above indicate that the use of one or two pink noise test signal(s) in the 'source room' will produce the least variance of all recorded units.

See Fig.1 for graphic illustrations of the frequency spectra of these modified by the 'source room'.

286
Experiment 1 - Laboratory Study 2

To establish SI data for various types of brick walls, testing procedures described in Study 1 above was employed including the use of two incoherent pink noise signals in the 'source room'.

Test results are tabulated in Table II for single brick walls and Table III for double cavity wall construction. See Fig. 2 for graphic illustration of 'STC' and dB(A) variance. All dB figures are rounded off.

Table II - Single brick walls

<table>
<thead>
<tr>
<th>Test No.</th>
<th>Thickness mm</th>
<th>Surface density kg/m²</th>
<th>STC</th>
<th>R</th>
<th>dB(A)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>90</td>
<td>144</td>
<td>42</td>
<td>41</td>
<td>44</td>
<td>Rendered both sides</td>
</tr>
<tr>
<td>2</td>
<td>110</td>
<td>144</td>
<td>43</td>
<td>41</td>
<td>46</td>
<td>Rendered both sides</td>
</tr>
<tr>
<td>3</td>
<td>110</td>
<td>173</td>
<td>43</td>
<td>41</td>
<td>46</td>
<td>Rendered both sides</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>194</td>
<td>44</td>
<td>42</td>
<td>46</td>
<td>Rendered one side</td>
</tr>
<tr>
<td>5</td>
<td>130</td>
<td>215</td>
<td>45</td>
<td>43</td>
<td>46</td>
<td>Rendered both sides</td>
</tr>
<tr>
<td>6</td>
<td>150</td>
<td>257</td>
<td>46</td>
<td>42</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>150</td>
<td>296</td>
<td>46</td>
<td>44</td>
<td>44</td>
<td></td>
</tr>
</tbody>
</table>

Table III - Double cavity brick walls

<table>
<thead>
<tr>
<th>Test No.</th>
<th>Thickness mm</th>
<th>Surface density kg/m²</th>
<th>STC</th>
<th>R</th>
<th>dB(A)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>230</td>
<td>204</td>
<td>44</td>
<td>43</td>
<td>44</td>
<td>2x90 mm + 50 mm cavity</td>
</tr>
<tr>
<td>9</td>
<td>230</td>
<td>240</td>
<td>46</td>
<td>45</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>250</td>
<td>246</td>
<td>46</td>
<td>45</td>
<td>44</td>
<td>as above, render both sides</td>
</tr>
<tr>
<td>11</td>
<td>250</td>
<td>275</td>
<td>46</td>
<td>45</td>
<td>48</td>
<td>90 + 110 + 50 mm cavity</td>
</tr>
<tr>
<td>12</td>
<td>250</td>
<td>356</td>
<td>46</td>
<td>45</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>

Experiment 1 - Field Study 1 (2)

To test whether the SI value of brick walls measured in dB(A) under laboratory conditions, Study 2, could be used to estimate noise levels transmitted into 'receiver rooms', a number of field experiments were conducted. For results see Table IV below. In all cases there were no openings between the two test rooms, though in tests marked *, flanking paths presented weak links.

Table IV

<table>
<thead>
<tr>
<th>Test No.</th>
<th>Thickness mm</th>
<th>Surface density kg/m²</th>
<th>Lab. value dB(A)</th>
<th>Field value dB(A)</th>
<th>Calculated dB(A)</th>
<th>Measured dB(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150</td>
<td>296</td>
<td>44</td>
<td>44</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>296</td>
<td>44</td>
<td>44</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>230</td>
<td>204</td>
<td>43</td>
<td>49</td>
<td>41</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>230</td>
<td>204</td>
<td>43</td>
<td>46</td>
<td>41</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td>230</td>
<td>204</td>
<td>43</td>
<td>42</td>
<td>52</td>
<td>54*</td>
</tr>
<tr>
<td>6</td>
<td>130</td>
<td>215</td>
<td>46</td>
<td>44</td>
<td>42</td>
<td>45*</td>
</tr>
<tr>
<td>7</td>
<td>230</td>
<td>280</td>
<td>44</td>
<td>43</td>
<td>47</td>
<td>50*</td>
</tr>
</tbody>
</table>

Summary

Laboratory Study 1 confirmed AS recommendation that the noise signal to be used for sound insulation measurement, in this study brick partitions, should be pink noise.

287
Laboratory Study 2 produced Sound Insulation data for different brick walls and the results were recorded in various units including dB(A).

Field studies performed have shown good agreement, with acceptable tolerances, between predicted and measured SL in 'receiver rooms' when using dB(A) SI values for brick partitions, keeping in mind that the recommended design sound levels for building interiors in AS 2107 are quoted with a 5 dB(A) tolerance/range.

Examination of typical noise sources in offices (slides 1 and 2), hotels (slides 3 and 4) and in many industrial spaces (slides 5 and 6) show that most of those noises have pink noise spectra, thus permitting the use for noise control SI data in dB(A) for brick walls at the early stage of design of various buildings.

References

T. Vass - Lightweight wall design for specific STC/absorption - 12th ICA Toronto, E2-6, 1986
T. Vass - Curtin Consultancy Reports, 1978 to 1988, University Reports.
MEASUREMENTS ON THE SOUND REDUCTION INDEX OF WINDOWS AND GLAZED EXTERIOR DOORS

Panayotopoulos N.*, Tzekakis E.*

*Laboratory of Architectural Technology, School of Architecture, Aristotle University of Thessaloniki, 54006, Thessaloniki, Greece

The Laboratory of Architectural Technology at the Aristotle University of Thessaloniki is adequately equipped to carry out several kinds of measurements on the fields of sound, noise and building acoustics.

Since 1985 the Laboratory has a new pair of rooms for measuring and testifying the sound reduction index of wall partitions without flanking transmission (figure 1).

The source room has a volume of 50 m³, a floor area of 16 m² and a total inner surface of 95 m². Its walls, with a thickness of 24 cm are built of concrete slabs with cavities filled with sand.

The receiving room has a volume of 43 m³, a floor area of 17 m² and an total inner area of 76 m². The floor is a 15 cm thick floating screed on 10 cm thick glasswool slabs with a density of 100 kg/m³. The wall construction is the same as that of the source room. The ceiling, built of reinforced concrete, is 15 cm thick. The room’s resonance frequency lies below 10 Hz.

Among other measurements being carried out regularly, these of glazed elements, namely windows and glazed exterior doors, appear to be of great interest, because of the wide variety in the construction characteristics of these elements and their effect on their acoustical performance. Fig. 2, 3 and 4 show construction details and sound reduction curves from three measurements on glazed units of different characteristics and frame materials.

A sufficient number of measurements on glazed doors and windows made of modern frame materials, has eventually resulted in a classification of these building elements according to their acoustical performance. This classification indicates their suitability to be used as parts of sound insulating external walls. In order to achieve maximum insulation, a special 50 cm brick construction has been used as the separating wall, in which the glazed samples have been

<table>
<thead>
<tr>
<th>No</th>
<th>Test Item</th>
<th>Material</th>
<th>Number of panes</th>
<th>Thickness D (mm)</th>
<th>R_w (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>2</td>
<td>4+16+4</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>single window - sealed</td>
<td>aluminium</td>
<td>1</td>
<td>4</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>single window - sealed</td>
<td>aluminium</td>
<td>2</td>
<td>4+12+4</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>single window - sealed</td>
<td>aluminium</td>
<td>2</td>
<td>4+24+6</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>double door - sealed</td>
<td>plastic</td>
<td>3</td>
<td>4+12+6+200+10</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>door single sealed</td>
<td>plastic</td>
<td>2</td>
<td>4+11+5</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>2</td>
<td>4+14+5</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>3</td>
<td>4+6+3+6+4</td>
<td>35</td>
</tr>
<tr>
<td>9</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>2</td>
<td>4+12+6</td>
<td>37</td>
</tr>
<tr>
<td>10</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>1</td>
<td>10</td>
<td>41</td>
</tr>
<tr>
<td>11</td>
<td>double door - sealed</td>
<td>plastic</td>
<td>3</td>
<td>4+12+6+100+10</td>
<td>46</td>
</tr>
<tr>
<td>12</td>
<td>single door - sealed</td>
<td>plastic</td>
<td>1</td>
<td>5</td>
<td>31</td>
</tr>
</tbody>
</table>

Table 1: List of tested windows and doors, with glazing details and R_w values
Fig. 1: Plan of source- and receiving room

placed. A number of tested glazed units, which appear to be of interest, are enlisted in table 1. The results of the measurements described aid to the classification of the glazed elements, as shown in table 2.

<table>
<thead>
<tr>
<th>Class No</th>
<th>Sound Reduction Index [dB]</th>
<th>Unit No</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>≥60</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>45 - 49</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>40 - 44</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>35 - 39</td>
<td>4, 8, 9</td>
</tr>
<tr>
<td>2</td>
<td>30 - 34</td>
<td>1, 2, 3, 6, 7, 12</td>
</tr>
</tbody>
</table>

Table 2: Classification of the sound insulation of measured glazed units
Figure 2 (Item No 4): Single aluminium framed window, double sealed, with a 4-24-8 mm glazing. Sound Reduction Index $R_w = 38$ dB.

Figure 3 (Item No 5): Double plastic framed door, each doubly sealed. Glazing 4.12.6 mm first door, 10 mm second door. Air space 200 mm. Sound Reduction Index $R_w = 50$ dB.
Fig. 4 (Item No 12): Single plastic framed door, double sealed, with 5 mm glazing. Sound Reduction Index $R_w = 31$ dB.

Fig. 5: Comparison of the 12 sound reduction curves.
AN INVESTIGATION OF ROOM-TO-ROOM SUSPENDED CEILING INSULATION MEASUREMENT BY ONE-TENTH SCALE MODELLING.

G.Dodd and X.Meynial.

Acoustics Research Centre, University of Auckland, New Zealand.

Introduction.

The measurement of the performance of lay-in tile systems for suspended ceilings is covered in ISO 140 Pt 9 (1) where a standard facility is specified.

No facility like this has been built in New Zealand and the Acoustics Testing Service (ATS) of Auckland University has been making insulation measurements in rooms significantly smaller than 50 m³ (see Fig. 1). Although this has provided a means for comparing materials we cannot directly compare our results with those from materials measured in accordance with the ISO standard and we have used modelling, to assess the likely differences in results.

Modelling Procedure.

Models of the ATS and the ISO 140 Pt 9 (ISO) measurement facilities, both at 1/10th scale, were constructed of timber and lead-lined gypsum board. The isolation achieved is shown in Fig. 2(a). The models were designed to meet the requirements of the ISO standard (e.g., with respect to surface absorption and plenum lining) A 50 mm diameter tweeter L/S was used as source for the steady state SPL's, and a spark generator for the RT measurements. We made a selection of ceiling materials for testing in the models to cover low and high insulations and a variety of absorption coefficients (as seen from the plenum side as well). The acquisitions were made using a developed version of the MIDAS auditorium modelling system [2] and the measurement procedure followed the recommended method for obtaining the normalised level differences Dn,c:

\[ Dn,c = Ls - Lr - 10 \log(Ar/Ao) \]  

where \( Ls \) is the average "centre-of-volume" SPL in the source room.

\( Lr \) is the average "centre-of-volume" SPL in the receiver room.

\( Ar \) is the receiving room absorption obtained from RT's.

All measurements of SPL and RT were the averages from 12 microphone positions and a repeatability test using different samples showed the results to be within the requirements specified in ISO 140 Pt 2[1].

Results (a) Performance of the Models.

As an indicator of the validity of the modelling procedure, we compared the model and full-size ATS facility for

(1) Level differences with no ceiling sample,

(2) Reverberation Times with highly absorbent samples and
(3) Standard deviations of SPL's and RT's.

The standard deviation values for the respective quantities were similar in all cases, and the R.T. and level difference values were acceptably close except in the region of 500 Hz and at one or two 1/3rd octaves below (fig. 2b). Thus we concluded that with samples in place the behaviour observed in the model would be similar to the behaviour expected at full-scale but with perhaps some lowering of accuracy at 500 Hz and below.

(b) Influence of Source Room Size

Both of the models and the full size facility show a trend for Dn,c to depend on the measurement direction. The differences in Dn,c for the 2 directions appear to depend on the TI of the ceiling and also, perhaps, the ceiling absorption on the plenum side. There is also a striking similarity of behaviour in all 3 facilities for the Dn,c differences to change sign at the 315 or 400 Hz 1/3rd octave. The source in the smaller room creates a lower level difference below this transition frequency (c.f. the source in the larger room) and a higher level difference above it, (fig 3). There is a clear trend for these differences between the two directions to be larger in the ATS model compared with the ISO model and to become larger as more absorption is added by the ceiling samples to the plenum space.

(c) Effect of Facility Size

The normalised level differences for 3 different materials are shown in figs 4 & 5. At all frequencies the smaller ATS test chambers produce higher values, but three regions can be identified where the behaviour of the Dn,c plots relative to each other is similar for the various materials. (1) At 200 Hz the values produced by the two facilities for a particular sample are almost identical. Towards lower frequencies the values separate but show quite similar slopes. (2) For the octave centred on 315 Hz the results are the furthest apart and show the insulation in this region to be rather uneven. (3) Above 500 Hz the plots for each material are virtually parallel suggesting that values measured in one facility may be used to accurately predict the values for the other.

Theoretical Approach

ISO 140 regards measurement spaces as containing diffuse sound fields and to have at low frequencies a sufficiency of room modes whose resonances are evenly distributed. It is not likely that these conditions will be readily met in our small ceiling transmission rooms. However, applying the usual insulation equation (1.4.3) and taking the symbols as defined in fig 6 we can show

\[ D_{n,c} = 20 \log \left( \frac{1}{t} \right) - 10 \log \left( \frac{S_{Sr}}{A_0 A_p} \right) \]

Thus \( D_{n,c}^{ATS} = D_{n,c}^{ISO} + 10 \log \left( \frac{|S_{Sr}^{ISO}| A_p^{ATS}}{|S_{Sr}^{ATS}| A_p^{ISO}} \right) \) (3)

For the particular dimensions involved we obtain

\[ D_{n,c}^{ATS} = D_{n,c}^{ISO} + 4 \text{ dB} \] (4)
Discussion

Corrections are shown in Table 1, for obtaining ISO results from measurements for the 3 distinct frequency regions mentioned above. Whilst the mean values for these corrections at low and high frequencies are close to the correction derived in (4), there are indications that for specific materials there are dependencies not yet accounted for, and the discrepancy in the 315Hz region is large. The uneven behaviour in this 315Hz region has been noted by Veres and Meichel [3], who argue that it is connected with the half-wavelength resonance of the plenum height. In our case, the divergence of the results can be explained by the plenum resonances occurring at different frequencies.

The simple theory presented above is inadequate in taking the plenum field as diffuse, and in not taking into account the additional attenuation that can be expected in what is a partially lined duct. We might expect that as the plenum absorption is increased, these additional losses would be higher for the larger facility (e.g. see Osipov et al.[4]) and hence the corrections required for the ATS results would decrease. However, this trend is only observed at high frequencies (see Table 1).

Conclusions

1. Our measurements indicate that smaller measuring rooms result in higher values for $D_n,c$.
2. The dependence of $D_n,c$ on frequency shows three fairly distinct regions, low and high frequency regions where the insulation increases monotonically with frequency, and a mid-frequency "plateau" of uneven behaviour.
3. The direction of measurement has an influence on the measured insulation, and the effect is reversed between low and high frequency regions.
4. A diffuse-field theory is partly successful in predicting the observed behaviour, but more work is necessary to establish the form of the plenum transmission.
5. The work has shown the feasibility of predicting ISO values from measurements in non-standard facilities and has provided a set of corrections which can be applied to ATS results when estimates of $D_n,c$ and $D_n,cw$ for an ISO standard facility are required.

References

<table>
<thead>
<tr>
<th>CEILING TYPE</th>
<th>$&lt;200,\text{Hz}$</th>
<th>$315,\text{Hz}$</th>
<th>$&gt;500,\text{Hz}$</th>
<th>$D_n,cw$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High TL</td>
<td>$-6.0,\text{db}$</td>
<td>$-8.2,\text{db}$</td>
<td>$-6.7,\text{db}$</td>
<td>$-6.7,\text{db}$</td>
</tr>
<tr>
<td>Low TL</td>
<td>$-2.3,\text{db}$</td>
<td>$-7.7,\text{db}$</td>
<td>$-6.2,\text{db}$</td>
<td>$-6.7,\text{db}$</td>
</tr>
<tr>
<td>High @ Low TL</td>
<td>$-4.4,\text{db}$</td>
<td>$-9.0,\text{db}$</td>
<td>$-5.6,\text{db}$</td>
<td>$-5,\text{db}$</td>
</tr>
<tr>
<td>Av. @ all T.</td>
<td>$-4.2,\text{db}$</td>
<td>$-8.3,\text{db}$</td>
<td>$-5.5,\text{db}$</td>
<td>$-6,\text{db}$</td>
</tr>
</tbody>
</table>

**Table 1**

Corrections to be applied for estimates of ISO results.
Fig. 1 Dimensions of facilities

Fig. 2 (a) Isolation in ATS model
(b) Model and full-size comparison

Fig. 3 Effect of measurement direction

Fig. 4 Dn,c for High α ceiling

Fig. 5 Dn,c for Low and High TL's

Fig. 6 Definitions of symbols
STRUCTURAL TRANSMISSION - SOUND INSULATION OF TWO LAYER WALLS

Frigyes Reis

Dept. of Building Constructions, Budapest Technical University
H-1111 Budapest, Muegyetem rkp.1-3. HUNGARY

Introduction
The sound insulation of two layer separating walls, made of brick, gypsum, or light-weighted concrete elements, is determined by the structural transmission of vibration between the layers, through the coupling constructions (pillars, crossing walls, ceilings). A calculation method of the sound reduction index of double walls, built in laboratory or in situ is summarized here. The presented method fits to the general concept of energy-propagation on the flanking transmission paths.

The conception of the method
The two layer separating walls, built in laboratories or in situ, can be summarized by the schemes (or their combination) of fig.1. The direct path of propagation is signed by 1: this leads through the plate-airgap-plate system. All the other paths, including path 2, are considered to be flanking paths. Based on experimental results all the paths, crossing the airgap are neglected, and only the paths of structural transmission are taken into account.

The paths are characterized by their own sound reduction index, the validity of the well known formula (1) [1] is extended to the double constructions:

\[ R_k = R_1 + 10 \cdot \log \left( \frac{S_i}{S_j} \right) + 10 \cdot \log \left( \frac{\delta_i}{\delta_j} \right) + D_v \]  (1)

where i and j mark the constructions in the source room (i) and the receiving room (j), \( R_k \) is the direct sound reduction index of wall 1, \( S_i \) and \( S_j \) are the corresponding surfaces, \( \delta_i \) and \( \delta_j \) are the radiation efficiencies, \( D_v \) is the average velocity level difference between the elements i and j.

Calculating for eq. the sound reduction index of the transmission path No. 2, \( R_2 \) is the direct sound reduction of the first layer of the separating wall, \( S_1 = S_2 \), and \( D_v \), \( \delta_1 \), \( \delta_2 \) are related to the layers 1 and 2 of the double wall.

The sound reduction indexes can be summarized. Form.(1) shows, that the task is to calculate the radiation efficiencies and the average velocity level differences.

The radiation efficiency
The change of the radiation efficiency in form. (1) has two reasons: the possible change of thickness or material
... and the change of the way of excitation. The constructions looking to the source room are excited by airborne sound, while the constructions on the side of the receiving room get structure borne excitation. The radiation efficiency of the plate excited by airborne sound is always higher then the radiation efficiency of structure borne sound excitation. This phenomenon occurs under the critical frequency, and the well-known formulas don't give the correct values. In [2] the reasons were determined, and a computer-based calculation method was presented. The measured and calculated values of a bio-concrete plate are shown in fig.2.

The average velocity level difference

The level differences between the elements were calculated by means of the schemes of fig. 1, but other combinations are also possible. The details can be found in [3], [4].

The main elements of the physical model are the following:
- The walls are considered to be simply supported, thin plates of finite extent and of a two dimensional bending wave field.
- The pillar is a torsional bar, the ends are clamped.
- Only the bending wave transmission is considered through the structural junctions, the boundary conditions contain the balance of the moments, and the equality of the angles of rotation at the perimeter of the plates.

The calculation of the average velocities is made by determining the coupling between the modes, using computer programs.

To illustrate the way of working of a structural junction fig. 3 shows a corner of the scheme 1/b in equilibrium position and in motion.

Calculations and experiments

The measured and calculated results of a laboratory experiment are summarized on fig. 4. \( \Delta R \) is the improvement of sound reduction index, caused by the second layer of the wall. The effect of the floor and the ceiling can be neglected, the result is determined by the structural propagation through the brick pillars. The corresponding scheme is on fig. 1/a.

Comparing either the measured or the calculated \( \Delta R \) and \( D \), curves the importance of the radiation efficiency factor is clear in this case, since the critical frequency is about 1300 Hz. The effect of the brick walls, connected to the pillars are taken into consideration by a higher value of the loss factor.

The experiment on fig. 5 was published in [5] without any calculations. The light-weighted gypsum wall is so "weak", that the effect of the ceilings can be neglected. There were no measured results of the single layer wall, so they were replaced by a calculation, based on [6], the result of which is the curve⑧. The graphs show the calculated sound reduction indexes of the different paths of propagation, according to fig. 1/b, and the calculated and measured
resultant sound reduction index. The results show the importance of the transmission through the path (2).

Formula (1) points to two possible ways of increasing the sound reduction indexes: increasing the differences between the coupled elements (to increase $D_v$) and increasing the critical frequency with the density as high as possible.

References
1. DIN 52217 Flankenübertragung. Begriffe
4. Reis F.: Calculation of the Flanking Transmission of Two Layer Separating Walls. Noise Control 88, Cracow
5. Kotschy A.: Increasing the sound insulation of the dividing walls ... Proc. of the 9th Conference on Acoustics, Budapest
13th INTERNATIONAL CONGRESS ON ACOUSTICS
YUGOSLAVIA 1989

Fig. 4. MEASURED AND CALCULATED RESULTS OF A LABORATORY EXPERIMENT

Fig. 5. MEASUREMENT AND CALCULATION IN SITU
Sound Insulating Properties of Porous Materials.

Shun-ichi Nakamura
Tokyo National University of Fine Arts and Music
12-8 Ueno Koen, Taito-Ku, Tokyo-To, 110 Japan

Porous materials such as glass wool and rock wool are very popular and powerful sound absorbents of great utility. They are, as is well known, the key materials not only for sound absorbing structures but also for sound insulating lightweight structures. And yet, they have been thought however to give only a supporting role for sound insulation.

Nevertheless the results of the fundamental study suggest that it is possible to give them a leading role based on the idea of utilizing their advantageous region compared to usual sound insulating partitions.

Analysis of Sound Insulating Properties of Fibrous Porous Materials

Sound insulating properties of porous materials were already outlined by Nakamura [1]. The frequency characteristics of transmission loss and the relationship between thickness and TL values were examined; the range of those measurements, however, was very limited since the previous measurements were not made from a practical point of view.

Therefore, the recent study's aim was to reconfirm and expand the old report for a much wider range of thickness, density and materials of samples.

![Fig.1 Specifications of samples examined.](image1.png)
![Fig.2 Examples of the results of TL measurement.](image2.png)

Samples and Measurements

Fig.1 shows the specifications of the fifteen samples examined. Those varieties in thicknesses of samples were made by piling layers upon layers. Among those samples there were such cases as two rock wool mats of different density were layered together and three cases of [GW48K], whose results from the previous TL measurements were used again for the present analysis.

Sound transmission loss measurements were carried out using the standardized method, though the sample area was temporarily reduced to 1.62m². Flow resistance measurements were also carried out for each of four materials using test pieces of [diameter:100mm, thickness:100mm].

Frequency Characteristics of Sound Transmission Loss

Fig.2 shows the examples of the measured results arbitrarily chosen only for ease of comparison.
In Fig.3 those results have been rewritten so that the logarithmic scale is adopted to the TL axis. Then the curves in Fig.2 changed to parallel straight lines in the greatest part of the observed area. Results in Fig.4 also show the same characteristics.

The facts shown here empirically support the idea reported previously that the frequency characteristics of sound transmission loss can be approximated by \((1/2 \text{ power of frequency})\).

**Contribution of Thickness**

As shown in Fig.4 for examples, it can be assumed that the sound transmission loss value can be considered to be (proportional to the thickness of the sample) except for lower region of TL values. That means that the total value of sound insulation in dB can be cumulatively calculated as the sum of the values of each of layers piled together. In the case of the multi-layer type, the experimental value approximately agreed with the estimated value.

Those might come from the fact that the sound insulating properties of this kind of material is primarily affected by the sound waves in samples being transmitted in the normal direction.

**Contribution of Density, Surface Density**

The original data of fifteen samples were simplified into five lines each in Fig.5, 6 and 7 by applying the \{thickness law\} cited above. Therefore, the difference appearing in Fig.5 shows the contribution of some structural factors. Density must be the most important among them, but those results suggest that the contribution of some other factors such as diameter of fiber cannot be neglected. So it is quite natural that the value of surface density of a sample should, in general, not be related to sound insulating properties in a simple way as shown in Fig.6.
Relationship to Flow Resistance

Finally, the measured value of flow resistance was found to be reliable when related to sound insulating properties. According to the results of the investigation the (2/3 power of flow resistance) must be taken into account as shown in Fig.7, where the normalized flow resistance is so defined as the value estimated for a thickness of 1000mm.

On the other hand, it is known that the normalized flow resistance is, as a whole, dependent on the power of density [2].

Combining those two relationships, it can be recognized, as a whole, that the sound insulating properties are approximately proportional to density, as long as the varieties of the corrections based on the other structural factors are small enough.

Generalized TL Characteristics of Fibrous Acoustical Materials.

Synthesizing the results of those investigations cited above, the generalized TL characteristics have been obtained as shown in Fig.8.

The measured data of eleven samples were rearranged, excluding the old ones. That way they gave a unified characteristic of TL value. Four cases of [GW80K] are graphically demonstrated there for reference.
Comparison to Practical Lightweight Partitions.

As shown in Fig.9, the IL values obtained by usual thin porous materials are insignificant and unsatisfying from the practical point of view of sound insulation design, even though they are recognized as powerful absorbents. Those values are clearly less than the value expected from the mass law for single leaf partitions. However, results obtained from the heaviest sample in this study approaches almost the same values of normal partitions. In the region above this condition, this fact suggests porous material would be advantageous with respect to the design of lightweight partitions.

Monograph for Sound Insulating Partition Design

Fig.10 is an example of the monograph for sound insulating partition design based on this investigation. Using this chart, the critical thickness from where the sound insulation of [RW 150K] exceeds the mass law and or some certain measured results can be discussed.

It must be emphasized that the advantage of applying this kind of material to sound insulating partitions, would be expected to appear in the field of noise control engineering in low frequencies, since in this region, as is well known, reliable lightweight partitions have been difficult to design.

Acknowledgements

I am grateful to Mr. M. Ogawa of the Kobayasi Institute of Physical Research for his support in the practice of measurements.

References

W. Rückward

Bundesanstalt für Materialforschung und -prüfung (BAM)
Unter den Eichen 87, D-1000 Berlin 45

INFLUENCE ON THE AIRBORNE SOUND INSULATION OF BUILDING
ELEMENTS BY SLITS; DOWELS AND PARTIAL COVERINGS

The airborne sound insulation between rooms and between
the free field and rooms is first of all dependent upon the
mass of a wall or upon a good constructed double wall, second-
ly from the by-way in walls. These by-ways can be doors,
together with open joints and walls with different bricks; or
you have the influence of curtain wall constructions. It also
can be an opening in ceilings for tubes. In all these cases
the sound insulation will be reduced.

In this performance some results of measurements from
slits and joints and from walls with missing coverings in
parts will be presented. In Figure 1 there is shown a plate
with a joint in the upper part. The thickness of the plate
is 40 mm, the breideness 1000 mm, the parameter is the height.
The loss in the insulation from 0 to 10 mm height in \( \Delta R_W = 11 \text{ dB} \); the same result you receive when the slit is situated
between plate and floor (Figure 2).

The insulation underneath the slit resonance is dependent
from the length and the width and follows the equation
\( R \sim 20 \lg \frac{l}{b} \).

Another form of joints is found between tubes and walls
(or floors) e.g. in water or heating systems. When there is
a bore-hole in a ceiling and there is used a tube of 35 mm
diameter, you can find a difference from 5 dB between
padded and unpadded joint as seen in Figure 3. The free
sectional area is 10 cm². With a medium length of about
1 = 13 cm, the height is \( h = 0.75 \text{ cm} \). The same beginning
of the falling reduction is seen in Figure 4, with the
same bore-hole but the diameter of the tube is about
d = 25 mm (f = 500 Hz).

The influence of a slit with a length of 420 cm between
a wall and the ceiling is seen in Figure 5; the height of
the slit is given with about 2 mm. When the slit is made
tight, the reduction can win \( \Delta R_W = 12 \text{ dB} \), and covering
the wall with a plaster you win 3 dB again. The greatest
difference is seen in the frequency range \( f > 800 \text{ Hz} \). The
upper curve is only a function of the mass; that means:
there is no slit or joint. The joint is caused in the
behaviour of the wall after being built.
The covering of a wall for a good thermal insulation consists of damping material and plaster, with the resonance frequency from about $f_0 \approx 250$ Hz (Figure 6). When a small part of this insulation is removed (10 cm), the sound insulation ($R$) is reduced from $R = 82$ dB to $R = 79$ dB ($f = 3150$ Hz) and $R_W = 55$ dB to $R_W = 52$ dB. The $f_0$-phenomena is still existing; it is also existing when the non-covered wall has a height from $h = 100$ cm. The influence of the double wall construction (higher insulation $f > 500$ Hz) is lost, but the negative influence of the resonance frequency is remained.

**Figure 1**
Influence of a joint in a door-panel on the airborne sound insulation (joint in the middle of the panel)

**Figure 2**
Influence of a joint in a door-panel on the airborne sound insulation (joint between door-panel and floor)
Figure 3
Bore-hole in a ceiling with a tube of 35 mm in it and open/closed slit

Figure 4
Bore-hole in a ceiling with a tube of 25 mm in it and open/closed slit

Figure 5
Influence of a slit from about 1 till 2 mm between a wall and the ceiling (length 420 cm) on the airborne sound insulation
Testing rooms for walls with reduced flaming transmission

1. 10 mm gypsum plaster
2. 24 mm sand lime bricks
3. 60 mm polystyrene foamed plastic
4. 5 mm plaster

5. $R_w = 53$ dB
6. $R' = 52$ dB
7. $R'' = 51$ dB (wall only with one-sided plaster)

**Figure 6**
Partial covering of a wall by a thermal insulation system and the changing of the airborne sound insulation.
THE PROPAGATION OF STRUCTURE-BORNE NOISE THROUGH MODELS OF MODERN BUILDINGS

Alexander Osipov

Dept. of Building Physics, Moscow Architectural Institute
Poschdestvenka ul., 11, Moscow 103754, USSR

The problem of structure-borne noise propagation in modern large-sized acoustically homogenous construction buildings is of great practical importance - most of city residents are aware of this fact. Various engineering equipment or nearby underground lines can be the sources of structure-borne noise. Mechanical oscillations of walls or floors, appearing in one point, are large distance propagated almost without any damping and irradiated into the air, resulted in uncomfortable conditions in living accommodations and industrial premises.

The influence of building constructive parameters on structure-borne noise propagation was estimated by calculations, made with statistical energy analysis /1/ and by experimental modelling the fragments of panel and framework buildings. The models were made from plexiglass in the scale 1:7. The oscillations were induced by a stationary vibrator 4810 in the range of 125-6000 Hz. The root meansquare values of acceleration level were measured on each element of the model in 25 points with further statistical processing the results of these measurements, using on-line analyser with graphical registrator 2717. The dependence of structure-borne noise attenuation upon the distance from the source of oscillations was estimated by the following formula

$$\Delta L_a = 10 \log \left( \prod_{i=1}^{n} 10^{0.1(L_{ai})} \right) - 10 \log \left( \sum_{i=1}^{n} 10^{0.1(L_{mi})} \right), \quad /1/$$

where $\Delta L_a$ - acceleration level attenuation; $n$ - number of measurement points at each element of the model; $L_{ai}$, $L_{mi}$ - root meansquare value of the acceleration level on the source and on the element of the model respectively.

According to statistical energy analysis, frameless panel buildings can be considered as a dynamical system, consisting of plates, tightly connected with each other. The
calculated model of framework buildings become more complicated because of blocking masses/bars, columns/in junctions of building plates. The calculation assumes, that the great part of sound energy is transferred by bending waves. The longitudinal waves were not taken into account because of the complication of mathematical apparatus. The equations system of energy balance for each model, consisting of 42 elements, is the following

\[ \omega \mathbf{KE} = \mathbf{P} \]

where \( \omega \) - circular frequency; \( \mathbf{K} \) - loss matrix; \( \mathbf{E} \) - matrix-columns of unknown energies; \( \mathbf{P} \) - external energy matrix.

The necessary values of transmission coefficients of sound energy through the junctions of framework building constructions were determined on the base of /2/.

The theoretical regularities of oscillation energy distribution through floors, partitions and walls of framework and frameless buildings dependently on the distance from the source of oscillation were got and experimentally confirmed.

It was found strong attenuation of structure-borne noise in more than 800 Hz frequency in framework buildings in comparison with frameless buildings. While in middle range /15-600 Hz/ there was found the contrary effect, i.e. the attenuation in frameless buildings is 1-2 dB more, than in framework buildings, but in the range of 125-250 Hz the damping values were just similar /fig. 1/.

The final aim of structure-borne noise isolation is decreasing the sound pressure level in accommodations/air noise/, appearing because of irradiation of oscillating bounding constructions. There were considered the problems of predicting the efficiency of different measures in structure-borne noise isolation. Since the structure-borne noise is the spreading of elastic waves through hard bodies/plates, columns/, its decreasing is possible by influencing on physico-mechanical properties of concrete construction, on size relationship of floors and wall intersections, on using elastic paddings and blocking masses in junctions and on enlarging construction masses. The equations system of energy balance includes all the above mentioned parameters and that is why their variations could determine by numerical calculation the ways and measures of decreasing the oscillations of bounding constructions and air noise in accommodations.
The decreasing of sound pressure level $\Delta L_p$ in accommodations was determined as

$$\Delta L_p = 10 \log \frac{\sum S_i <V_i'^2>}{\sum S_i <V_i^2>}$$

where $<V_i^2>, <V_i'^2>$ - root mean square of oscillation velocity before and after measures of structure-borne noise isolation respectively.

It was shown, that the decreasing of air noise in accommodations is largely influenced by construction damping. When the loss factor is increasing twice, than in accommodation $\mathcal{L}$ the sound pressure level is decreased by 2-3 dB and in accommodation $\mathcal{L}^*$ by 5-7 dB.

References


Fig. 1. The attenuation of structure-borne noise in models of panel /a/ and framework buildings /b/

- - experiment;
- - SEA calculation.
3. G
AN ORIGINAL METHOD FOR IN SITU IMPEDANCE MEASUREMENT

E. Luzzato, A. Adobes, D. Bosselut

Département Acoustique et Mécanique Vibratoire, EDF - D.E.R.
1, Av. du Général de Gaulle - 92141 - CLAMART CEDEX - FRANCE

Most of the impedance measurement methods [1-8] require the use of particular devices such as a wave guide, an anechoic chamber or a particular type of sound source. In each of these configurations, the acoustic material to be measured may have a particular acoustic behavior because it is not exposed to the same wave propagation as the one present in the room where it is placed. In other words, the use of a wave guide will change the behavior of the material because the sample which is used is more compressed, its dimensions are smaller, or only normal plane waves are taken into consideration. Although the impedance measurement with a two microphone technique in an anechoic chamber can lead to better results, it often does not provide the acoustic absorption characteristics of the material when placed in the actual room. The only way of obtaining these characteristics is to measure them in the room with the surrounding noise. In fact, the method must be designed not only for acoustic impedance measurement, but for determining the actual boundary condition at the considered plane surface.

The method presented in this paper consists in the analysis of the acoustic field in the vicinity of the considered plane surface. Its principle is to find a local approximation of the acoustic pressure field by means of measured data at three points inside a numerically bounded domain that does not contain any source. By means of these three acoustic pressure data, it is possible, according to a given functional, to extrapolate an approximation of the complex acoustic pressure and normal velocity at the plane surface. The parameters defining this functional are unique in the domain of interest and they can be identified by means of the three measured data.

THE METHOD PRINCIPLES

Inside a small bounded domain $\Omega$ containing no sound source, one can take a local approximation of the acoustic pressure field $P$ such as :

$$P(x) = \alpha \ e^{iKx} + \beta \ e^{-iKx}$$  (1)

where $x$ denotes the position vector of the observation point in $R^3$, $K$ is a vector in $R^3$ whose components are complex numbers and $\alpha$ and $\beta$ are two functional parameters with complex values. This approximation is issued from the form of the plane wave solution of the Helmholtz equation. For instance, if all the imaginary parts of the components of vector $K$ are zero and if its modulus is equal to the actual wave number $k$, then expression (1) represents a plane wave field with respect to the direction of wave vector $K$.

![Figure 1: The measurement principle](image)

In general, the acoustic fields observed in the vicinity of plane walls cannot be accurately approximated by plane wave fields. However, in the case of large halls and in the vicinity of large plane surfaces, the acoustic fields can get close to the plane wave field as long as the sound sources are far away from the observation domain.
In order to account for the difference between the ideal plane wave field and the real one, vector \( \mathbf{K} \) is defined with complex components. In fact, in this impedance measurement problem, functional (1) is considered as an interpolation functional that proved to work better than other functionals such as polynomial expressions. Therefore, the interpolation is only valid in a small region, compared to the wavelength, next to the considered plane surface. Furthermore, since one is generally interested by a boundary condition expressed in terms of the gradient component orthogonal to the plane boundary, expression (1) can be simplified as follows:

\[
P(x) = A \exp(i k_x x) + B \exp(-i k_x x)
\]

(2)

where \( k_x \) denotes the orthogonal projection of \( \mathbf{K} \) on direction \( x \), the outward direction orthogonal to the plane boundary (figure 1), \( k_x \), A and B are three complex numbers whose values depend upon the acoustic field in the local approximation domain. The approximation of the acoustic particle velocity in direction \( x \) is derived from (2):

\[
V_p(x) = k_x A e^{i k_x x} - B e^{-i k_x x}
\]

(3)

where \( \rho \) denotes the fluid density. Therefore, the estimated specific acoustic impedance at a given observation point of the boundary \( x = 0 \), is expressed by the ratio of the estimated pressure to the estimated normal velocity at this point:

\[
Z = \frac{\omega}{k_x c} \frac{A + B}{A - B}
\]

(4)

Let \( P(x_1) \), \( P(x_2) \) and \( P(x_3) \) be the complex acoustic pressures measured at point \( M_1 \), \( M_2 \) and \( M_3 \), next to the boundary surface (figure 1) and let \( H_{21} \) and \( H_{31} \) be the ratio of the acoustic pressures defined as follows:

\[
H_{21} = \frac{P(x_2)}{P(x_1)} \quad \text{and} \quad H_{31} = \frac{P(x_3)}{P(x_1)}
\]

(5)

The identification of functional (2) with the values of the acoustic pressure measured at points \( x_1 \), \( x_2 \) and \( x_3 \), leads to the following transcendental equation in \( k_x \):

\[
e^{i k_x (x_3 - x_2)} e^{-i k_x (x_3 - x_1)} H_{21} \left[ e^{i k_x (x_3 - x_1)} - e^{-i k_x (x_3 - x_1)} \right] + ... + H_{31} \left[ e^{i k_x (x_2 - x_3)} - e^{-i k_x (x_2 - x_3)} \right] = 0
\]

(6)

Since equation (6) has an infinite number of quasi-periodic solutions, the boundaries of its solution domain must be established in the complex plane:

- The upper limit of the real part of \( k_x \) is \( \frac{\omega}{c} (1 + e_1) \) and its lower limit is \( \frac{\omega}{c} e_2 \)
- \( e_1 \) and \( e_2 \) are both positive real numbers, smaller than 1.

- The upper limit of the absolute value of the imaginary part of \( k_x \) is \( \frac{\omega}{c} e_3 \) where \( e_3 \) is a positive real number, smaller than 1.

When a solution \( k_x \) of equation (6) is found in its solution domain, one can compute the estimation of the normal specific impedance:

\[
Z = \frac{\omega}{k_x c} \frac{H_{21} \left[ e^{i k_x x_1} - e^{-i k_x x_1} \right] - e^{i k_x x_2} e^{-i k_x x_2}}{H_{21} \left[ e^{i k_x x_1} + e^{-i k_x x_1} \right] - e^{i k_x x_2} e^{-i k_x x_2}}
\]

(7)
For a given measurement configuration and for a given pulsation \( \omega \), there may not be any solution of equation (6) in the solution domain and therefore no estimation found for the impedance. However, the measurement is generally performed for a large number of frequencies (1024, 2048 or 4096 points) and the impedance curve in terms of frequency can still be accurately estimated.

THE NUMERICAL AND EXPERIMENTAL IMPLEMENTATION

The Newton method is used to solve equation (6). In the Newton iterative process, various initial values of \( k_z \) are chosen inside the solution domain for each considered frequency. The maximum number of iterations is 20; it allows one to find an accurate value for the solution.

The method has been implemented with a 3 channel FFT analyzer with a time record size of 4096 points, that is 2048 frequency points for the frequency transfer functions. Three uncalibrated microphones of 5 mm diameter are used for the experimental implementation and an overall calibration of the three microphone device is made before each measurement. For the calibration, the three microphones are set at the position of microphone \( M_1 \) so that the acoustic pressure signal is almost the same for the three microphones. The two reference transfer functions \( H^{21}_{0} \) and \( H^{31}_{0} \) are then measured. For the impedance measurement, the microphones are placed at their respective positions \( M_1, M_2 \) and \( M_3 \) so that the transfer functions \( H^{21} \) and \( H^{31} \) can be determined. The transfer functions used in the numerical process are therefore:

\[
H^{21} = \frac{H^{21}_{0}}{H^{21}_{0}}, \quad \text{and} \quad H^{31} = \frac{H^{31}_{0}}{H^{31}_{0}}.
\]

In the experimental implementation of the method, the values chosen for the parameters that define the solution domain of equation (6) are: \( \varepsilon_1 = 0.05 \), \( \varepsilon_2 = 0.1 \) and \( \varepsilon_3 = 0.2 \).

If the considered plane surface is large compared to the wavelength, if it has a local reaction behavior and if the sources are far away from this surface with respect to the wavelength, the acoustic field next to the surface can be correctly represented by a superposition of plane waves. In this case, one can sever the various incidence angles for the measurement by respectively setting the two limits of the admissible real part of \( k_z \) to \( \cos \varphi_1 \) and \( \cos \varphi_2 \) where \( \varphi_1 \) and \( \varphi_2 \) are two given incidence angles between 0 and \( \pi/2 \).

CONCLUSION

In large rooms, this method allows one to determine the actual acoustic boundary conditions of walls with the ambient noise. It helps the user to find some particular absorption behavior such as acoustical or mechanical resonances, that are generally difficult to observe. When used on local reaction materials such as fiberglass and foam surfaces in a large reverberating room, the results are similar to the measurements made in an anechoic chamber or in a plane wave tube with a two microphone technique. This method has also been used to determine the local complex impedance at surface points of plate resonators in the domain of their resonance frequency and the results were in very good agreement with the numerical simulation results.

REFERENCES

Figure 2: Measured values of complex specific impedance and absorption factor of a glasswool material. The curves are drawn in terms of frequency. The smooth curves represent the values measured with a two microphone technique in a plane wave guide. The other curves presenting large variations show the values measured with the three microphone method in the machinery hall of a nuclear power station; the three microphone method was implemented using the ambient noise generated by the machines in the hall.
* 13th INTERNATIONAL CONGRESS ON ACOUSTICS *
* YUGOSLAVIA * 1989 *

RESTORATION AND RECONSTRUCTION OF THE LIBRARY BUILDING
OF THE CITY OF BELGRADE

Ilios Gavrilovic
Arhitektura i urbanizam, Djure Djasovic b,
Prtumn-Beograd, Yugoslavia

1. SHORT HISTORY

The building was erected as a hotel in the second half of the 19th
century. The name was "Serbian Crown". Sometime during the last deca-
des the purpose of the building was changed, and lately the building was
dilapidated that a thorough reconstruction had to be undertaken with
the idea to convert it into the City Library. The library depot for
books required more space than originally encircled by the basement; 3
years ago, while digging to enlarge the basement, remnants of the entry
tower of the Roman Castrum Singidunum showed up. Stone blocks, bricks,
and a layer of thin stone bricks and even a leaden water-pipe of 20 cm in
diameter had been found. They all belonged to the complex of the castrum
built by Romans for the Fourth Flavian Legion in the 1st c. A.D.,
now covering mainly the area of the Park Kalemegdan. Our reconstruction
may be considered as a time bridge over 20 centuries.

The restoration and reconstruction were designed by the highly spe-
cialized office for protection of monuments in Belgrade III, which guar-
anteed appropriate preservation of all relevant elements and styles.

2. SOME DATA ABOUT THE BUILDING

A four-storeyed edifice has massive brick walls. The ceilings above
the first two floors were made of bricks forming arches; and above the
3rd and 4th floor the ceiling structure was made of wood, and although
well preserved had to be changed with a reinforced concrete structure.
The complete area is 3700 m². The brick walls have a surface mass of
about 1500 kg/m². As a contrast, newly designed party walls are light
prefabricated partitions of 10 cm.

Noise inside the building is made by mechanical rooms for water sup-
ply, ventilation and air-conditioning. The boiler room with electric
heaters is situated in the ground floor. There are no more noisy rooms.
Quiet rooms are grouped in the other side of the building.

The ground floor comprises of: Entrance, Issue Department with Card
Catalogue and Open Shelves, an atrium with an Art Gallery; a small re-
freshment bar, offices for the Administration and Accounting, a large
reading room for arts, and a computer terminal room.

Each floor is supplied with its own lavatory.

3. THE SOUND INSULATION DESIGN

Although the Yugoslav Regulations for Design and Construction of
Buildings contain criteria for noise control in libraries, no details
are given for elements of a library building. Therefore, criteria had
to be chosen and defined for all walls, ceilings, doors and windows,
based on good practice and experience.

Calculations of sound transmission loss had to allow for two things:
the old wall structure varied to a great extent, and its layer materials

319
sometimes were not very well known. When in great doubt, measurement on the spot had been made. Analogous and similar structures have been studied in order to obtain the necessary results. It was all done in a very conservative way so that the designed sound insulation has a safety margin of about 1 to 3 dB.

Masonry walls, having a high surface mass, did not require particular attention, but the light partitions made of gypsum cardboards had to be further tested for their acoustic qualities in an acoustic laboratory.

All provided types of prefabricated wooden partitions also had to be tested according to Yugoslav Standards. The lowest sound insulation required was $R_W = 42$ dB. In some cases the demand was as high as $R_W = 52$ dB, which called for careful elimination of sound leaks and flanking paths both in the structure and during the mounting of the walls.

All floors are so designed as to diminish the foot-step noise. Carpets on resilient underlayments are used as the most efficient ones in mid- and high-frequency range. To achieve higher values of impact noise insulation over the entire frequency range, an elaborate construction consisting of concrete slabs with suspended ceilings and floating floors has been applied. Rigid floor finishes, such as ceramic and marble floors, are also designed as floating structures.

Some special precautions have been carried out: the resilient layers under the floating screed have been carefully protected against moisture. The structural floor was made as even as possible, and all around the screed, along the walls, a non-hardening caulking was applied, to separate the screed from walls. No "sound bridges" are allowed, and final measurements "in situ" should prove the correct performance. - The floating screeds also comply with Yugoslav Standards as to the permissible load.

All floor coverings were acoustically tested according to YU-standards and they had to satisfy the requirements stated in the project.

Doors. - To act as an efficient sound barrier, door should be heavy. In the Library building there are solid-core wooden doors and hollow-core metal doors filled with fibrous and lim materials. We preferred doors with raised sills, since they are more efficient than those with flush sills. The metal doors are fitted in a metal frame with a neoprene seal.

4. PROTECTION AGAINST COMMUNITY NOISE

The library building has three fronts towards three streets, two of which carry traffic of medium density including street-car lines (Paris Street) and a busy trolley-bus terminal (Najićevo Street). An improved sound insulation of windows facing the streets was therefore necessary. Double glazed windows with wide air-space between the panes were requested. Application of non-hardening sealants secured good sealing of cracks. No unforeseen gaps were left during the mounting.

5. THE RULES FOR CONSTRUCTION WORKS

Some rules had to be followed during construction works. When mounting ducts and pipelines care had to be taken to seal all passages through concrete slabs, partitions and masonry walls. Windows and doors have been precisely checked while mounting in order to prevent sound leakage. Supervisor and the acoustic designer had to check every door
regarding the data provided by the manufacturer. The data should contain the value of the sound insulation as tested by an official institution, and comply with the project requirements. For instance, the metal doors should have at least $R_w = 40$ dB.

6. MECHANICAL NOISES AND VIBRATIONS

To insulate both low and high frequency vibrations, steel springs with ribbed neoprene blankets are used. Also neoprene-in-shear mounts of various stiffnesses are applied to absorb vibration effectively.

Cooling tower vibrations involve low frequency vibrations of the slowly rotating propeller-type cooler and high frequency vibrations resulting from the impact of falling water. The library building required appropriate noise control technique for the roof-mounted cooling tower.

Electrical energy is supplied by a dry-type transformer, insulated by several waffle-shaped layers. All electrical connections are made of flexible braided cable.

7. AIR DUCT LAYOUTS

The insulation of noise transmitted through air ducts is applied in both supply and return-air systems. The library building has many premises sensitive to noise, and large mufflers and linings, sized to satisfy requirements for less sensitive premises, are installed in main ducts, while highly sensitive rooms called for additional mufflers in branch ducts.

The mechanical room is far away from the noise sensitive areas. Its enclosing building structure consists of heavy walls masonry and concrete. The ceiling is covered with thick sound absorbing material. All penetrations are sealed with caulk and pipes are sleeved with fiber packing. Also, cantilevered steel springs are used.

Cross talk is usually transmitted through short common air ducts running from one room to another. It is dealt with the additional mufflers in branch ducts.

To avoid transmission of the system vibrations to partitions, the ductwork was nowhere rigidly mounted to walls. Care was also taken of ducts passing through noisy areas in order to avoid penetration of noise into the duct.

REFERENCES

[1] Main architect Mr. S. Vučenović; Sound insulation project by Mr. Z. Perolo.


INTRODUCTION

Owing to their low initial and operational costs, fabric structures are currently used in architectural applications. Their use is particularly favoured for temporary installations, especially when traditional permanent structures are not allowed to be built up. Rather large spaces are utilized for indoor sport activities, musical and drama shows, and other recreational activities which involve more or less critical acoustical qualities. Usually, a fabric structure is supported by a network of cables and pylons (tent) or by pressurized air within either an enclosed fabric envelope or the main volume of the structure itself (airhouse). Peculiar geometrical and constructional features are often responsible of poor acoustics in these large spaces which result in an uncomfortable acoustical environment.

Generally, the shape of these structures, although simple, presents large concave surfaces to the impinging sound; therefore focusing effects often result. Classical equations based on the ideal diffuse sound field theory are of questionable use when predicting room acoustical parameters of spaces having an odd sound distribution. Ray acoustics seems to be a more appropriate approach to cope with the above mentioned problem.

This note reports an application of a computerized ray-tracing technique to the prediction of some room acoustical features at four locations in an airhouse containing two tennis courts. Computer results are compared with measured data.

GENERAL FEATURES OF THE AIRHOUSE

Fig.1 shows the general appearance of the airhouse under consideration. It has a dome shape with a quadrangular base whose side lengths average 36.5 m.

Fig.1 External (left) and internal (right) views of the airhouse
its central height is about 11.2 m. The flexible membrane structure is made up of a single PVC impervious sheet fixed to the ground all along its lower edge. It is kept standing by a moderate overpressure of the air in the main volume which is generated by a continuously running fan. In normal operation the total surface of 2800 m² encloses a volume of about 9500 m³.

SOUND ABSORBING PROPERTIES OF THE INTERNAL SURFACES

By neglecting some comparatively small surfaces (e.g. doors, fan outlet), those most relevant are the membrane surface $S_m$ and the tennis court ground surface $S_g$. They amount to 1500 m² and 1300 m² respectively. In order to calculate an energy echogram, and the room acoustical parameters deriving from it, for given sound source and receiver locations, the computerized ray-tracing procedure needs suitable sound absorption coefficients of the above mentioned surfaces. Unfortunately, only a few measured data about the acoustical properties of fabric structures are reported in the open literature. Croome [1] suggested to evaluate the sound absorption coefficient $\alpha$ of a flexible membrane by using a classical formula for calculating the normal incidence sound transmission coefficient of a limp wall, namely:

$$\alpha = 1 + \left( \frac{\pi f M}{p c} \right)^2$$  \hspace{1cm} (1)

where $f$ is the frequency (Hz), $M$ is the mass per unit area (kg/m²), $p$ is the air density (kg/m³) and $c$ is the sound speed (m/s). However predicted values by eq.1 were found by Croome in fairly good agreement with measured data for two membranes only at low frequency (125 Hz). Experimental values were higher than predicted ones by an order of magnitude at 500 Hz and 1 kHz. Byrne [2] published a procedure for calculating the diffuse sound field absorption coefficient and the diffuse field sound reduction index of fabric constructions. Fig.6 of Ref.[2] reports a set of curves from which the 1/3-octave diffuse field sound absorption coefficient for a uniformly tensioned impervious sheet as a function of the frequency can be derived. Like eq.1, also these curves show that the sound absorption coefficient for a low-tensioned impervious single fabric sheet depends solely on its mass per unit area. The 1/3-octave band sound absorption coefficients predicted by Byrne for an impervious sheet with $M=1.27$ kg/m² were in good agreement with those calculated from 1-octave band reverberation times measured in an actual fabric roofed structure of the same mass per unit area.

By weighing some samples of known area a value of about 0.61 kg/m² was obtained for the flexible membrane of the airhouse under consideration.

Table 1 reports the sound absorption coefficients calculated through eq.1 ($\alpha_C$), by using Fig.6 of Ref.[2] ($\alpha_B$), and those actually used in the computerized

<table>
<thead>
<tr>
<th>$f$(Hz)</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_C$</td>
<td>0.420</td>
<td>0.150</td>
<td>0.040</td>
<td>0.011</td>
<td>0.003</td>
</tr>
<tr>
<td>$\alpha_B$</td>
<td>0.630</td>
<td>0.340</td>
<td>0.140</td>
<td>0.050</td>
<td>0.017</td>
</tr>
<tr>
<td>$\alpha_m$</td>
<td>0.380</td>
<td>0.150</td>
<td>0.040</td>
<td>0.011</td>
<td>0.003</td>
</tr>
<tr>
<td>$\alpha_g$</td>
<td>0.030</td>
<td>0.040</td>
<td>0.090</td>
<td>0.110</td>
<td>0.120</td>
</tr>
</tbody>
</table>

324
ray-tracing calculation ($\alpha_m$). Last row of Table 1 shows the sound absorption coefficients used for the tennis court ground ($\alpha_g$). The values of $\alpha_m$ and $\alpha_g$ were selected by a trial and error searching. This was performed bearing in mind that i) the sound absorption coefficient of the membrane decreases with increasing frequency, as predicted both by eq.1 and Fig.6 of Ref.2, ii) the tennis court ground is a highly resistive porous surface, therefore its sound absorption coefficient is a mildly increasing function of the frequency. As a matter of fact, the tennis court ground is made of a bed of finely milled red bricks, very hard-packed and mixed with a binder. Over the resulting compact surface a thin layer (0-1 mm) of loose red brick dust is shed. It is often moistened with water. Embleton [3] reports values of the flow resistivity for "Quarry dust, fine, very hard-packed by vehicles" in the range 5000-20000 (CGS units) and values higher than 20000 (CGS units) for "Asphalt, sealed by dust and use". The values of $\alpha_g$ in Table 1 lay in the range bounded by the normal incidence and uniform incidence (locally reacting surface) sound absorption coefficient predicted by the normal impedance of a very thick layer of porous material having a flow resistivity of about 25 $10^6$ kg.m$^{-3}$s$^{-1}$ (1 CGS unit = 1000 kg.m$^{-3}$s$^{-1}$). Its intrinsic acoustical properties were evaluated by Mechs's formulas [4].

MEASUREMENT RESULTS AND RAY-TRACING CALCULATIONS

One location of the sound source (a small fire-cracker) and four locations of the microphone were used in the experimental work. Eight responses to sound pulses generated by fire-crackers were tape recorded at each receiving point. Subsequent analysis of the recording with 1-octave band filtering yielded the average reverberation times reported in rows TM of Table 2. Rows TC report the corresponding results obtained by the ray-tracing calculation. Insertion of the sound absorption, air absorption included, used in the computer calculation into Sabine equation produces reverberation times consistently higher than those in Table 2. Results calculated by the ray-tracing technique disclose that the sound absorption coefficients of the membrane appear to be closer to the normal incidence ones ($\alpha_e$) rather than to the diffuse field ones ($\alpha_B$). Furthermore, most calculated echograms show an uneven time distribution of the sound energy.

Table 2. Reverberation time in seconds.

<table>
<thead>
<tr>
<th>Receiving location</th>
<th>$f$(Hz)</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TM1</td>
<td>1.60</td>
<td>4.24</td>
<td>6.15</td>
<td>6.46</td>
<td>5.07</td>
</tr>
<tr>
<td></td>
<td>TC1</td>
<td>1.57</td>
<td>3.89</td>
<td>6.39</td>
<td>6.25</td>
<td>5.13</td>
</tr>
<tr>
<td>2</td>
<td>TM2</td>
<td>1.41</td>
<td>4.06</td>
<td>6.63</td>
<td>5.70</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>TC2</td>
<td>1.49</td>
<td>3.65</td>
<td>6.04</td>
<td>5.92</td>
<td>4.83</td>
</tr>
<tr>
<td>3</td>
<td>TM3</td>
<td>1.30</td>
<td>4.02</td>
<td>6.24</td>
<td>6.17</td>
<td>4.85</td>
</tr>
<tr>
<td></td>
<td>TC3</td>
<td>1.59</td>
<td>3.99</td>
<td>6.32</td>
<td>6.22</td>
<td>5.20</td>
</tr>
<tr>
<td>4</td>
<td>TM4</td>
<td>1.51</td>
<td>3.93</td>
<td>5.91</td>
<td>6.15</td>
<td>5.10</td>
</tr>
<tr>
<td></td>
<td>TC4</td>
<td>1.53</td>
<td>3.83</td>
<td>6.18</td>
<td>6.09</td>
<td>5.02</td>
</tr>
</tbody>
</table>
This corresponds to audible echoes in the airhouse. For example, Fig. 2 shows the echogram calculated at the receiving location N.2 for the frequency 1 kHz (left).

![Echogram and recorded decay](image)

Fig. 2. Calculated echogram (left) and corresponding recorded decay (right)

and the corresponding experimental 1-octave band sound pressure level decay (right). A periodic sequence of sound energy concentration is visible in the echogram, as well as periodic peaks show up in the experimental sound pressure level decay. In both cases the average sequence period is about 0.25 s (an audible repetitive echo).

**CONCLUSIONS**

The observed airhouse presents acoustical defects which are frequently encountered in these structures, that is exceedingly long reverberation times, uneven sound distribution and repetitive echoes. Owing to some lack of sound diffusing surfaces, these features can be predicted fairly well by a computerized ray-tracing technique, once suitable sound absorption coefficients are found for the internal surfaces. According to the results reported in this note, they may lay in the range bounded by their normal incidence and diffuse field values.

**REFERENCES**


DESIGNING TSS HALLS

Tomislav Stanojevic*
Goran Sakota**

* SAVA Centar, Belgrade, Yugoslavia
** AVC Consulting & Engineering, Belgrade, Yugoslavia

1. Introduction

The Total Surround Sound System (TSS - patents pending) [1], as a possible future technological standard, assumes a multichannel audio system in which the audio signal is processed by a special TSS processor (Figures 1 and 2).

The fundamental properties of the TSS system are as follows:

- digital audio channel processing;
- spatial (3-D) sound reproduction;
- possibility of recreating the original sound field with exact locations of individual sound sources;
- space and speed simulation of sound effects (Travelling Sound Effect) in a large number of combinations using all angles in space;
- full correlation between the sound and visual axes; and
- compatibility with systems using the Delta stereophonic principle (DSS) [4].

The TSS system may be used in future cinema halls using HDTV projectors, theatre and multifunction halls, in various show halls, for spectacles and multimedia presentations, as well as in rock concert halls, night clubs, discotheques, and much like.

The TSS system concept also caters for the possibility of having a large number of creative performances in the fields of movies, theatre and "live" concert sound. In rock music, using a TSS system and a TSS processor operating in a real time mode, one can achieve significantly better results than using a classic system, in conjunction with impressive possibilities for creating sound effects.

In movies and theatre plays, one can produce most realistic sound effects by simulating air and sea battles, effects from nature (storms, ocean storms, earthquakes), street sounds, and so forth.

Future TSS theatre and multifunction halls must be designed in a different manner from those operating today, in full accord with the TSS
system. This article explains the fundamental conditions which must be satisfied by TSS halls, as well as certain principles of the new design philosophy.

2. Basic Characteristics of TSS Halls

According to the TSS concept, the sources of direct sound (3 TSS processor channels) are catered for by three loudspeaker groups behind the projection screen (of the movie or video film), or above the stage (of the theatre and concert hall), while the other sources (5 TSS processor channels) are to be located in the walls, the floor and the ceiling, depending on the form and dimensions of the hall.

In order to satisfy the condition of optimally covering the audience and using the most favorable locations of sound sources, an analysis yielded the required form of a TSS hall (Figure 3). Due to the required visibility of the screen or stage, the halls must have a gentle audience slope (around 5%), and because of the loudspeakers in the ceiling (Ceiling Travelling Sound), it is necessary for the ceiling to follow the slope of the floor (the constant distance of the ceiling from the floor is required for optimal calculations regarding a matrix network of the so-called "Cricket" loudspeakers in the ceiling). Galleries or balconies in the hall will require a special matrix sound network with a different step (mutual distance among loudspeakers).

For optimal distances among sound sources regarding effects and the audience, the relationship between the hall height, length and width should be 1:2:3, and side walls may gently slope outward from the stage towards the back wall of the hall (by about 5–10%).

As the TSS system is in essence an electro-acoustical system of sound reinforcement and effects reproduction, the reverberation time of halls should be as short as possible in order to decrease the influence of space on the reproduction of recorded or live amplified sound. In medium sized TSS halls (5,000 m²), it is recommended that the average reverberation time be no longer than 0.9-1.1 sec, while the frequency-dependent reverberation time should be as linear as possible, with a gentle drop in higher frequencies. The back wall of the hall should be lined with absorption materials so as to eliminate harmful reflections and doubled sound on the stage (in concert versions of the hall).

In order to cater for effects from the floor, it is necessary to provide corresponding floor openings for mounting sound sources, which must be constructed in such a way as to provide for maximum possible mechanical protection, but without inhibiting optimal source radiation capabilities.

3. Designing TSS Systems and Halls

The electro-acoustical calculations for the three front loudspeaker sound sources in TSS halls may be done as shown in the literature [1], [2], and the calculations for the matrix network of "Cricket" loudspeakers may be performed using the calculations for a distributed network as defined
in the literature [3]. The hall ceiling and walls should contain efficient sound sources with a directivity characteristic of 60°×60°, while the floor may contain less efficient sources with a directivity characteristic of 80°×80°. The front sound sources should use powerful and efficient loudspeaker groups (systems) with a constant directivity at high and medium frequencies, as well as corresponding subwoofer sections for reproducing extremely low frequencies (explosion effects, earthquakes, etc).

For calculations of locating the sound sources using 3-D graphic displays for the audience, the TSS hall may use a program by the NEXO company (NEXOCAAD - Computer Aided Acoustic Designer), which was presented on the 86th AES Convention in Hamburg. Using this program developed for IBM PC/AT and compatible computers, it is possible to effect the analysis and calculations of spacial acoustics of TSS halls (reverberation time, early reflections, impulse response, etc). Using the NEXOCAAD program has been simplified by implementing high quality menus, which in conjunction with a mouse and zooming enable quick access to individual program windows and their optimal use for displaying calculation results.

4. Conclusion

Because of the specific character of the TSS system, it is necessary to simultaneously design TSS halls for both the spacial acoustics and for the location of sound sources and their optimal covering of the audience area. It is thus necessary to use computers with such programs which interactively make it possible to input and change all data related to the form, dimensions and finish of the hall walls, as well as data related to sound sources (directivity, efficiency, etc). A graphic display of acoustics parameters and analysis results regarding audience coverage enable fast producing of acoustic processing designs of halls and locations of sound sources for all eight TSS channels. It is the authors' opinion that such an approach provides for optimal implementation of complex TSS parameters in future TSS halls.

5. References


Fig. 1 Functional block diagram of TSS system

Fig. 2 Symbolic presentation of loudspeakers in TSS hall

Fig. 3 TSS hall - plan and longitudinal section
CALCULATION OF THE TRANSMISSION LOSS OF PANELS WITH NON-HOMOGENEOUS CROSS-SECTIONS

Miašina Kurtović

Faculty of Electrical Engineering, Beograd-YU, Bul. Revolucije 73

The intention of this paper is to propose a method for calculating the transmission loss of non-homogeneous partitions, such as multilayer, hollow, ribbed or corrugated ones, shown in Fig.1.

For calculating the transmission loss the necessary data are:

- the surface mass (\( m_s \))
- the coincidence frequency (\( f_c \))
- the plateau height (\( R_{PE} \))
- the internal damping factor (\( \eta \)).

Knowing these parameters, one can choose one of the already accepted approximate methods to plot the diagram of the transmission loss versus frequency.

For all non-homogeneous partitions in Fig.1 it is suggested to use in all equations the real surface mass. After having found the coincidence frequency, this mass should be used to establish the product \( m_s f_c \) by means of which one can find \( R_{PE} \) (for example by shifting \( R_{PE} \) for the homogeneous plate of the same mass by \( 20 \log \left[ (m_s f_c)_{non-hom}/(m_s f_c)_{hom} \right] \)). The damping factor, needed to calculate the transmission loss above \( f_c \), changes for such partitions only in the case presented in Fig. 1(a), but methods for calculating the new combined damping factor are well known.

What remains as a problem, still not solved adequately, is to find the coincidence frequency of such partitions.

The impedance of a homogeneous, relatively thin panel, excited by a sound wave, coming under the angle of incidence \( \theta \), neglecting losses, is given by [1]:

\[
Z_c = j \omega m_s + \frac{B' \omega^2}{\rho c^2} \sin^2 \theta
\]

where \( B' \) is the bending stiffness. This impedance becomes ze-
ro for $\theta = \pi/2$ at the frequency:

$$f_c = \frac{c^2}{2\pi} \sqrt{\frac{m_s}{B'}}$$

known as the coincidence frequency.

If the panel is a combination of two (or more) different panels, stiffly joined, denoted by 1 and 2 in Fig.1, then as both panels must have a unique bending wave velocity - their impedances are connected in series, we have then:

$$Z_c = j\omega \sum m_i + \frac{\omega^3 \sum B'_i}{j\rho c} \sin^4 \theta$$

The coincidence frequency is now, for $\theta = \pi/2$:

$$f_c = \frac{c^2}{2\pi} \sqrt{\frac{\sum m_i}{\sum B'_i}}$$

As $m_s = \rho d$ ($\rho$ is the density and $d$ the thickness of the material), and $B'$ is proportional to $Ed^3$ ($E$ being the Young modulus), the last equation can be rewritten as:

$$f_c [Hz] = 6.4 \times 10^4 \sqrt{\frac{\sum (\rho_i d_i)}{\sum (E_i d_i^3)}} = 6.4 \times 10^4 \sqrt{\sum \left(\frac{m_i}{E_i d_i^3}\right)}$$

(1)

For the partition in Fig.1(a) use Eq.1 to calculate the new coincidence frequency.

For the partition in Fig.1(b) the stiffness is practically the same in all directions and therefore in Eq.1 only one thickness $d$ , the real one, should be used.

The other two examples are similar. For these partitions the stiffness of the vertical part 2 does not affect the stiffness of the whole panel in the X-axis direction. One must take into account only the greater surface mass and calculate the coincidence frequency as:

$$f_{cx} = 6.4 \times 10^4 \sqrt{\frac{m_s}{E d_i^3}}$$

In the Y-axis direction we use Eq.1, but adequately modified. The impedances of parts 1 and 2 are mechanically connected in series because they have the same velocity. Their specific acoustic impedances $Z_{11}$ and $Z_{22}$ should be multiplied by $\ell_1$ and $\ell_2$ respectively, the obtained mechanical impedances per length summed and then divided by the total length $\ell_1 + \ell_2$ to obtain the new common specific impedance. The ge-
neral solution is:
\[
f_{cy} = 6.4 \cdot 10^{-4} \sqrt{\frac{Z_{r}(p_{r}, d_{r})}{Z_{l}(p_{l}, d_{l})}} = 6.4 \cdot 10^{-4} \sqrt{\frac{m_{s}}{Z_{r}(p_{r}, d_{r})}}
\]
and in this particular case:
\[
f_{cy} = 6.4 \cdot 10^{-4} \sqrt{\frac{m_{s}(p_{r}+d_{r})}{E(p_{l}, d_{l})}}
\]

The bending wave velocities are \(c_{fx}\) and \(c_{fy}\) in the two axes directions. For other directions \(c_{f}\) must be found. In an arbitrary direction under the angle \(\varphi\) the bending wave covers a path \(\sqrt{x^2 + y^2}\) (see Fig.2). One part of it, belonging to the \(x\)-axis, is \(x \cos \varphi\), and the time needed to cover it is \(\frac{x \cos \varphi}{c_{fx}}\). For the \(y\)-axis the corresponding propagation time is \(\frac{y \sin \varphi}{c_{fy}}\). Thus the velocity under the angle \(\varphi\) is:
\[
c_{f} = \frac{\sqrt{x^2 + y^2}}{x \cos \varphi + y \sin \varphi} = \frac{c_{fx} c_{fy}}{c_{fx} \cos \varphi + c_{fy} \sin \varphi}
\]
\[
= \frac{c_{fx}}{\cos \varphi + \sqrt{f_{cy}/f_{cx}} \cdot \sin \varphi}
\]

because \(c_{f}\) is proportional to \(\sqrt{f}\), hence \(c_{fx}/c_{fy} = \sqrt{f_{cy}/f_{cx}}\).

This angle distribution of velocities, given by Eq.2, provides correctly for \(f_{cx} = f_{cy}\) the same velocity in all directions. An example when \(f_{cx} \neq f_{cy}\) is given in Fig.3.

The mean value of \(c_{f}\) in the region between the angles \(\varphi_1\) and \(\varphi_2\) is:
\[
c_{f} = \frac{1}{\varphi_2 - \varphi_1} \int_{\varphi_1}^{\varphi_2} c_{fx} \, d\varphi = c_{fx} \left( \frac{1}{\varphi_2 - \varphi_1} \int_{\varphi_1}^{\varphi_2} \frac{d\varphi}{\sqrt{f_{cy}/f_{cx} \cdot \sin^2 \varphi}} \right)
\]
\[
= \frac{c_{fx}}{\varphi_2 - \varphi_1} \int_{f_{cy}}^{f_{cx}} \frac{\sqrt{f_{cx} \cdot \tan \varphi}}{\sqrt{f_{cy}}} \left( \frac{1}{\sqrt{f_{cx}}} \right) d \varphi = c_{fx} \cdot D
\]

When there is a small difference between the values of \(f_{cx}\) and \(f_{cy}\) then we can take \(\varphi_1 = 0\) and \(\varphi_2 = \pi/2\), and obtain \(f_{c} = \sqrt{f_{cy} \cdot f_{cy}}\). When there is a large difference between these values, it is necessary to divide the region between \(\varphi = 0\) and \(\varphi = \pi/2\) in several (\(\pi\)) equal parts, to find out for every partition the velocity \(c_{f}\), using Eq.3, then to calculate the corresponding \(f_{cy} = c_{f}/D\), and \(P_{ef}\) by means of \(m_{s}/c_{f}\), and finally to draw for the part \(i\) of the partition the whole diagram of \(R_{ie}\).

Having at his disposal \(\pi\) such diagrams, one can calculate the resulting transmission loss of the partition as:
\[ R = 10 \log \left( \frac{n}{\sum_{i=1}^{n} 10^{-c_i/R_i}} \right) \]

In Fig. 3, where \( f_x/f_y = 25 \), the plane is divided into \( n = 5 \) regions. Good results have been achieved using this method.

Fig. 1

Fig. 2

Fig. 4 - Calculated curve and measured values (circles)

REFERENCES
MODEL EVALUATION IN BELGRADE NATIONAL THEATRE RECONSTRUCTION

M. Mijic, H. Kurtovic, Z. Perolo, V. Onicin
Faculty of Electrical Eng. Belgrade

The National Theatre in Belgrade is an edifice built in the 19th century. During its history its hall changed its looks several times. The reconstruction of the theatre, carried in the course of last several years, and which is just being completed, had the aim, among other things, to redesign the interior of the hall and to bring it into the state close to the original from the 19th century.

The analyses in the field of room acoustics during the process of designing and carrying out the reconstruction were made by the physical model of the hall which was made for that purpose.

This paper presents the characteristics of the model of the Belgrade National Theatre Hall and some results obtained during the measurements in it.

THE PHYSICAL MODEL OF THE NATIONAL THEATRE HALL

The physical model of the interior of the National Theatre Hall was made according to the interior design. On the basis of hall dimensions, and on the basis of the possibilities of the measuring technology, the optimal proportion of the model of this hall was chosen, and the model is scaled 1:10.

The hall interior has the horseshoe form. From the aspect of the modelling work, complex forms of interior boundary surfaces are the main model characteristic. The hall consists of the orchestra and three galleries. On the edge of the orchestra and on the first two galleries the space is subdivided by boxes. In the front part of the orchestra, before the stage, there is the orchestra pit. It is designed so that it could be closed at the stage level (whereby the stage is enlarged), or at the orchestra floor level.

Exterior dimension of the model are approximately 190x200x175cm. Prior to the onset of the elaboration, a concept was adopted that the model be formed of two parts, with the crossection along the longitudinal axis of the hall.

Both halves of the model are placed separately into a basic construction formed by wood slabs (5x8 cm in cross section). In such a way the necessary physical stiffness of the construction is realised. Within such a construction the sheating of wood panels were fixed (18 mm thick) which limit

335
the size of the model on the exterior. Such frame, in the shape of a box, served to set the elements of the hall interior.

In order to make the elements of the model interior different wood materials were applied (panel slab, solid wood, etc). The variety of the materials applied is conditioned by the variety of interior elements, characteristic for hall from that period. They are arc-shaped gallery or box fences, a vaulted ceiling, etc. Characteristic are also the varied levels in the hall floor. The lateral edge walls of the hall are made of plywood, because this material permit the adjustment to the existing curves in the hall. The columns, which in the hall itself pass through all floors, are made of solid wood. In order to obtain good precision in model construction, they were all made out of one piece, and passed through all floors. The complexity and variety of forms conditioned some final corrections of the surfaces. Whenever it was needed, the corrections were made by the corresponding putty for wood. The smoothness of the surface, required for transposed frequencies on which the measurements were carried out, was achieved by painting the surfaces.

RESULTS OF MEASUREMENTS

Echogram measurements were the basic measurements carried out in the model. On the basis of these measurements, the parameters relevant for the evaluation of the acoustical properties of the hall were determine ("Definition", "Clarity"). The electrical sparkler was used as a source of an excitation pulses. All measurements were made by the sparks of 14 J energy. The registration of the acoustical signals was made by a 1/4" microphone.

In the course of designing, there was a large number of measurements performed in model as well as changes made in the configuration of some elements. However, the very concept of the building adaptation did not permit big physical changes in the hall, because it also represents a historical monument. Therefore, the majority of measurements was made in order to test the accepted concept.

Due to lack of space, only a small part of the results are displayed here. In figures 1, 2 and 3 some characteristic results are shown, describing the hall features in the octave at 1 kHz (10 kHz in model) when the sound source was set in the middle of the stage opening. Figure 1 shows the relative change of sound level in the hall, defined relative to the closest point in the orchestra (measuring point No. 1). The results show that the variations of sound level in the hall are within the interval of 7 dB. Fig. 2 shows the changes of "Clarity", while Fig. 4 shows the changes of "Definition" (in model is between 0,3 and 0,8 and C between 0 db and 12 db).
The measurements on the galleries were made so that the microphone was placed almost at the point where the listeners head in the position of easy sitting in the armchair should be. As the result of such measurements it was found that the worst places in the hall are the seats at the sides of higher galleries, especially at the third gallery (place No. 20). The same fact has influence in the values obtained for "Clarity" and "Definition". However, it is certain that the spectators in these seats always lean towards the gallery fence, which results in better listening conditions.

CONCLUSION

The measurements carried out in the model of the National Theatre Hall during its redesigning and reconstruction permitted undoubtedly the prediction of its acoustical features, and thereby the verification of the designer's ideas prior to the completion of works. However, the model and the work on it designing caused some investigations in the field of the technics of model constructions. Also the corelation between electrical and acoustical properties of the electrical spark was also investigated in some aspects.

LITERATURE

1. H. Els, "Ein Messystem fur die akustische modelltechnik" Wirtschaftsverlag NW, Dortmund (1986)
ACOUSTICS OF THE BELGRADE NATIONAL THEATRE

Zoran Perolo

Televizija Beograd, Projektantski biru TVR,
YU-11000 Beograd, Kneza Višeslava 88, Jugoslavija

Introduction

National Theatre in Belgrade is an eclectic XIX century building with a very tumultuous history. Built in 1869 supposedly by Architect. Aleksa Bugarski its front looked at that time very much like Teatro alla Scala of Milan (1), and the interior was based on theatre blueprints of the Austrian company "Fellner und Helmer" (2) An intimate auditorium, it had initially 15 rows (224 seats), 14 boxes and some standing space on the floor, 12 boxes and Princ's box on the first and 14 boxes on the second gallery. The third gallery had standing places only. The volume above the central floor area of 101 m² is Vc = 1855 m³ and total volume of auditorium V = 2433 m³. As far as acoustics is concerned. Item 8. Report of the Theatre Board from the year 1869 states:

... About acoustics of the building, in a hurry, we had no time to contemplate...

Since 1869 the theatre was modified and reconstructed many times. Two frontal wings with staircases for public added in 1911-1922 and many constructional changes (concrete roof girders and auditorium ceiling) changed the theatre perceptibly. The fourth gallery was added in extension to the central part of the third and the capacity increased from 714 to 944.

Major reconstruction of the auditorium took place in 1969. The boxes were removed, the seating capacity further increased. All walls received wooden panelling.

Finally, the present reconstruction (1986 1989) aimed at returning the 1922 look of the historical part of the theatre building included tearing down the haphazardly added technical areas behind the front stage and erection of a completely new 10000 m² wing.

Design concept

The building of the National Theatre in Belgrad under Historic Monument Protection Regulation was to be returned to the original state by 1986 blueprints (Architects Zdravković and Drinjački, acoustics by author). One of the problems was dating of the "original". The building from 1869 lacked space for public communications that was later added and consequently it was decided that appearance of the building after the 1911-1922 reconstruction was to be set as a designing guideline.

Acoustic concept of 1986 1989 restoration was based on reinstalling the boxes and corridors around the horseshoe shaped auditorium (III.1). Previous analysis of the XIX
* 13th INTERNATIONAL CONGRESS ON ACOUSTICS *
* YUGOSLAVIA * 1989 *

![Image](image.png)

century theatre of the same type in Pula, Yugoslavia (2) proved without doubt such theatres to be well studied "acoustic machines" whose vital parts are galleries, boxes and its elements (parapets, undersides, lateral partitions etc.). The very act of restoration of boxes is consequently a major acoustic contribution in respect to the previous state. It remained however of paramount importance to execute consequently the detailing and choice of materials. The existing reverberation time was judged as being on the low side and lacking low-frequency rise, since the theatre is used not only for drama, but as a main opera hall as well. Since the total volume was to be decreased by boxes and corridors formation it was necessary to secure as little additional absorption as possible.

Detailed analysis of the sound propagation within the auditorium resulted in a list of improvement possibilities. One of them was increase of the stage height in reference to the first row floor level and consequent increment of the floor rake. Gallery undersides were more favourably shaped where possible. Sound diffusion orchestra pit wall elements intended to improve the stage-orchestra pit communication are introduced (credit to Prof. H. Kurtović). Several other desirable interventions (shape of the gallery, slope of the gallery undersides etc.) were not feasible for various reasons, but were also not considered as crucial from the standpoint of acoustics.

The third segment of design concept consisted in provision of all interventions necessary for established acoustic defects to be removed.

Throughout the design stage the three-dimensional sound propagation control was executed using 1:10 scale model of the auditorium (3). No defects were observed. The measured acoustic properties were within acceptable limits and confirmed objectively author's conclusions derived by the graphic method.

Preliminary RT measurements are shown in Fig. 3 and compared favourably with measurements of the previous state.

General observations

Sound propagation analysis in case of National Theatre in Belgrade and National Theatre of Istria in Pula (2) showed remarkable similarities due to (probably) same origin i.e. plans made by company Fellner und Helmer of Vienna, Austria. Main reflecting surface for 1st order sound reflections is the ceiling. The only non-local lateral reflections of the 1st order cover very small portion of the auditorium and their origins are portal wall and gallery parapets. Most other low order reflections are local, i.e. from the vicinity of the listener - box partitions, rear wall of the box, box ceiling, underside of the gallery above, upper surface of the gallery parapet etc. Since only uniformly distributed reflections arrive from the ceiling, the level variations across listening plane are quite large (4). Reflections from the vicinity of the listener increase the subjective feeling of intimacy

340
primarily based on relatively small distance to the stage (max. 22 m for National Theatre Belgrade). The capacity of theaters of this type is limited to about 1300 seats. The limiting factor is (in XIX century) the empirically derived maximum height of the ceiling before the echo is perceived in the last row of the floor.

Conclusion

National Theatre of Belgrade is a fine if minute example of a European theatre from the second half of the XIX century. The auditorium type used at that time represents good "acoustic machine" even by today’s standards. In restoration work it's not so much the question of what can be done to improve the acoustics but rather what is to be avoided in order not to diminish the initially correct concept of the hall.

References

(1) Popović, Z., "Istorijska arhitektura pozorišta Jugoslavije u Europe", Gradjevinski fakultet, Beograd 1986
(3) Mijić, M., et al "Ispitivanje akustičkih osobina adaptirane sale Narodnog pozorišta u Beogradu pomoću fizičkog modela", ETAN, Novi Sad, 1989
1 - BEFORE RESTORATION
2 - AFTER

ILL. 3 RT FOR EMPTY HALL

342
IMPROVEMENT METHOD FOR FLOOR IMPACT SOUND INSULATION ON WOODEN FRAME STRUCTURE
(CONTROL OF FLOOR IMPACT SOUND ON WOODEN FRAME STRUCTURE: PART 1)

K. Inoue and S. Kimura

College of Science & Technology NITON Univ.
1-8 KANDA-SURUGADAI, Chiyoda-ku Tokyo 101, JAPAN

1. INTRODUCTION

In Japan there are many wooden frame structures for multiple dwellings. However, there has hardly been any research or development relating to improving floor impact sound insulation performance of party floors, which is an important consideration for the acoustic performance of a living space, and in many cases, the so-called conventional construction method is being used without modification. The floor impact sound insulation performance of wooden multiple dwellings in the case of a heavy impact source is between L-75 and L-80 in a room directly below the floor structure, and this is a serious problem except in the case of row houses where dwelling units are connected only by party walls.

S. Linnarsson and others have conducted a study on so-called heavy impact sources, because floor impact sounds at lower frequencies have become a problem in wooden houses, which are frequently built in Scandinavian countries. Another reason was that the results of an ISU test using a tapping machine did not coincide with the evaluation given by the residents of the multiple dwellings. They showed a section of an independent double flooring structure and an independent sound insulation ceiling which gave an improvement of about 7DB at the first peak of characteristic C against the impact given by a 2kg ball filled with sand and dropped from a height of 1.35m. On the other hand, Yamaguchi and others have used double flooring, and added dynamic vibration absorbers, which consist of a 3.2mm thick steel sheet and AD form, on the floor boards. They reported that the floor impact sound insulation performance can be improved by about 2 ranks (10DB) for the sound insulation grade of a heavy impact source by this method compared to the same performance of a conventional wooden floor. However, they are unable to improve upon the efficiency, which is considered to be necessary for party floors of multiple dwellings, and their study did not include radiated sound from the walls or the sound insulation effect of the ceiling of a room below the floor structure.

For this reason, we manufactured a full-size wooden frame structure to study a method to reduce floor impact sound against heavy impact sources of wooden floor structure. In this study, we examined the contribution of ceiling and wall vibrations against the generation of floor impact sound, and conducted various experimental research on a method to improve the floor impact sound insulation performance. As a result, we found that if total measurement are carried out on the floor, ceiling and wall structures of a room below the floor structure, wooden houses can have an insulation grade of about L-55, which is equivalent to the sound insulation performance of concrete multiple dwellings, against a heavy impact source.

In this report, we will summarize and explain the concept of the improving method in part 1, outline the development and process of improving each section of the test house in Part 2.

2. BASIC STUDY ON IMPROVEMENT METHOD
The heavy floor impact sound insulation performance of a conventional wooden frame structure is very poor, generally only 1-80. Several reasons have to be considered for this poor performance, such as, insufficient mass of floor face, insufficient flexural rigidity, reduced fastening ability of members at the joints etc. The following methods can be provided to improve the above conditions:

1. Make the floor section and wall section more rigid in order to control the vibration response. (Fig. 1)

2. Provide a dynamic vibration absorber on the floor face to control the vibration response. (Fig. 2)

3. Provide a floating floor on the floor face to reduce vibration transmission. (Fig. 3)

4. Provide a high performance independent sound insulation layer at the ceiling and walls of the room below the floor structure to insulate the radiation of sound from conventional floor and wall structure. (Fig. 4)

The method described in paragraph (1) is what we are proposing. The floor section and wall section are completely adhered by the combined use of nails and adhesive agents to make high rigidity in order to increase the geometrical moment of inertia and to reduce the vibration amplitude. This method can be an orthodox method on the basis of structural mechanics, and improved results are definite. The increased rigidity increases the strength of a structure and the stability of the floor face and also the living comfort may be improved in some way. However, it may be difficult to integrate members at the construction site. For this reason, it is better to introduce factory production to make panels for floor face. In this way the objective performance can be achieved.

For the paragraph (2), especially in the case of heavy impact sources, when considering the objective frequency of evaluations, the sound pressure response at 63Hz band surpasses other bands. For this reason, a dynamic vibration absorber to absorb the vibration of this frequency band should be used. Although the conventional floor is said to be light, it still has great mass. For this reason, great mass is required at the mass section of the dynamic vibration absorbers, which results in the system becoming too large. The paragraph (3) is to provide a basic floorboard, which has more than a certain amount of mass and rigidity, and an elastic material is spread on top to provide a floating floor layer. If this is done accurately, a great transfer loss of impact sound is expected. When the elastic material is thin, the effect of the elastic material spring is expected to greatly depend on the cushion of air inside the shock absorbing layer. For this reason, it is
very important to allow the air inside the shock absorbing layer to shift around when impact is given to the floor. When materials, such as iron grating or holed plywood are used, which gives sufficient ventilation performance to the floating floor layer, the floor impact sound insulation grade will be improved even more.

The paragraph (4) is to provide an independent high performance sound insulation layer inside a conventional wall structure of a room below the floor structure to insulate the impact sound generated at the conventional floor structure. If the sound insulation layer inside the wall can completely cut off vibration, a full performance of the sound insulation layer can be expected. However, this method still presents problems as to cost, the space of the room below the floor structure and the construction method.

We have described four countermeasures for heavy floor impact sound of wooden houses. However, all these methods may be unrealistic if sufficient improvement effect is expected by using these methods independently. It may be necessary to use these methods in combination with others in order to gain effective sound insulation results. This report will explain the high rigidity method of paragraph (1).

3. STUDY ON FLOOR SECTION USING HIGH RIGIDITY METHOD

In order to get high rigidity of floor section, it is better to increase rigidity by making a sectional structure as shown in fig.5, integrating the floor ribs and the surface boards, and increasing the geometrical moment of inertia. An increase of rigidity is estimated by using the point impedance which is proportioned to the square root of the rigidity. The sectional dimension is estimated on the supposition that the floor impact sound is conducted only to the ceiling surface of a room below and the sound insulation performance against the heavy impact source is L-55.

Supposing that the size of the room below the floor structure is about 12.5 m², and the calculation objective is 63 Hz band only, then the floor impact sound level of the 63 Hz band (L₀) can be expressed by the formula (1).

\[ L₀ = 20 \cdot \log_{10}(F_{max}/Z₀) + 10 \cdot \log_{10}(ρ₀ \cdot C₀ \cdot k) + 10 \cdot \log_{10}S + 10 \cdot \log_{10}(4/A) + 120 + ΔL - D \]  

(1)

Where, \( Z₀ \): point impedance for flexural wave of floor panel (kg/s)
\( F_{max} \): effective value of impulsive force of heavy impact source (N)
\( ρ₀ \cdot C₀ \): specific acoustic resistance of air (kg/m/s)
\( k \): sound radiation coefficient
\( S \): sound radiation area (m²)
A: sound absorbing power of the room below (N/m)

ΔL: difference between integral level per unit time and FAST peak
value of sound level meter (dB)

D: sound insulation effect by ceiling (dB)

The floor impact sound level of 63Hz band which is equivalent to L-55 sound insulation grade is set to 78dB. The sound insulation efficiency of ceiling - independent sound insulation ceiling to be provided - of the room below the floor structure is assumed to be 5dB considering safety and workability. The effective value of impact force is assumed to be F=55(N) based on our measurement. If the acoustic radiation coefficient at the lower board (16mm thick) of the floor panel is -3dB/oct when it is below the coincidence frequency (610Hz), then the acoustic radiation coefficient at 63Hz band is k=0.1. The sound radiation area is assumed to be S=12.5m². The sound absorbing power of a room below the floor structure is assumed to be A=10m³. The ΔL is set to +10dB based on our study results of actual waveforms.

\[78 > 20 \cdot \log_{10}(65/Z_s) + 10 \cdot \log_{10}(414 \times 0.1) + 10 \cdot \log_{10}(12.5) + 10 \cdot \log_{10}(4/10) + 120 + 10 - 5\]

\[Z_s = 2.1 \times 10^9 \text{ (kg/s)}\]

If the Young's modulus of wood is \(E = 7 \times 10^9 \text{ (N/m²)}\), then

\[Z_s = 8\sqrt{\frac{E}{\rho}} = 8\sqrt{\frac{7 \times 10^9}{\rho}} \quad \text{(2)}\]

\[I = 6.9 \times 10^8 / E = 0.0086 \quad \text{(3)}\]

In addition, geometrical moment (I) and mass per unit area (m) can be expressed as

\[\pi = \rho(0.072 + 0.29 \cdot h) \quad \text{(4)}\]

\[I(1.072 - 0.76 \cdot h^2) / 12 \quad \text{(5)}\]

Therefore, according to the floor panel section of Fig.5, the necessary height (h) of the floor rib to satisfy the value of formula (3) is h ≥ 0.21m, and when the total floor panel thickness is made to approximately 28cm, L-55 can be achieved.

This estimation is, of course, a guide for manufacturing floor panels. Although many factors, such as supporting points of floor panels at the periphery and beams, eigenvalue of the floor panels, vibration characteristics at the wall surface of the room below the floor structure, and acoustic characteristics of the space of the room below the floor structure should be involved in the actual calculation. However, it is clear what kind of basic performance is required for the basic floor structure against the floor impact sound generation system.

References:
1.S.Einsson:Low Frequency Impact from Timber Joist Floors, Inter-Noise 83,July 1983,pp 615-618

348
PRACTICAL APPLICATION OF METHOD FOR REDUCTION OF FLOOR IMPACT SOUND TO WOODEN HOUSE.
(CONTROL OF FLOOR IMPACT SOUND ON WOODEN FRAME STRUCTURE: PART 2)

S. Kimura & K. Inoue

College of Science & Technology, NIHON Univ.
1-8-Kanda-Surugadai, Chiyoda-ku, Tokyo 101, JAPAN

1. TEST AND ANALYSIS METHODS

We remodeled an experimental wooden model house of the real size in the following sequence and examined the improvement method of the floor impact sound insulation performance by obtaining the test results. This experiment was conducted to find countermeasures for the problems of the conventional floor construction method, and was conducted on condition that the conventional structure and the construction method should not be changed drastically.

First stage: Making a floor with highly rigid panels and adding a sound insulation ceiling.

Second stage: First stage remodeling plus making rigid wall surface in a room below the floor structure.

Third stage: Improvement of each section based on the results of the second stage remodeling.

Because the objective is a wooden structure, we considered it would be difficult to calculate the contribution of each section correctly in advance. For this reason, we decided to judge and deal with the effect of countermeasures basically according to the floor impact sound level and vibration measuring results at each stage. At each stage of improvement, the measurement of the floor impact sound level was conducted according to a heavy impact source based on JIS A1418. In order to study the degree of contribution from each section of a room below the floor structure, the vibration velocity responses on the

Fig.1: Floor Plan of Experimental Wooden Structure

Fig.2: Section of Highly Rigid Floor Panels and Vibration Proof Independent Ceiling
floor panel near the vibrating point, the ceiling surface directly under the vibrating point and the center of each wall surface of the room below the floor construction were simultaneously recorded in a data recorder when point S4 of Fig.1 was vibrated by a heavy impact source. The recorded waveform was calculated as an octave band level of 8Hz to 250Hz after being A/D converted at 4000Hz and going through a Fourier transformation.

2. STUDY OF TEST RESULTS

2.1 TEST RESULTS AT FIRST STAGE

In the first stage, eight highly rigid floor panels as shown in Fig.2 were installed on the floor surface. Each floor panel was joined together by the application of surface materials for the integrating of floor panels, onto the floor panels by using nails and adhesives to integrate the whole panel floor. The ceiling of the room below the floor structure was made into an independent sound insulation ceiling, the carrying rods of which were supported at the wall surface with the vibration isolating material, as shown in Fig.2, and two 12mm thick planter boards were applied to the ceiling. Other sections were made in the same way as conventional construction methods.

The measuring result of the floor impact sound level due to a heavy impact source is L65, and is three ranks better than the conventional floor on the sound insulation grade. However, it is not as much as the performance of L55 as calculated in a trial. According to Fig.3, which shows the measured values of vibration response at each section, the response of the wall surface sections indicated 10dB or higher than the response at the ceiling surface at over 63Hz band. The characteristics of the floor impact sound level is considered to be determined by the radiation of sound from these wall surfaces.

Based on the above results, the floor impact sound level in the first stage is assumed not to show the performance of the highly rigid floor panel as the wall surface is not rigid. This result indicates the necessity of conducting some countermeasures on the wall surface section.

2.2 TEST RESULTS AT SECOND STAGE

In the second stage, the wall surface of the room below the floor construction was improved. First of all, the fixation of studs to girth and ground-sill, which is done only by nailing under the conventional construction method, was
increased by using adhesives together. A layer of 12mm thick plaster board and two layers of 12mm thick plywood were applied on the wall surface inside the room with both adhesive and nails, which made the plaster board and plywood to integrate with the studs. The external wall was also integrated with 12mm thick plywood by using both adhesives and nails, and a brushing finish of 20mm thick mortar was applied to the surface.

Because of the improvement of the wall surface section, the floor impact sound level became L-55 (L number is L-57), and thus, the original object value was achieved. According to Fig.4, the vibration responses of each section of the room below the floor construction was more balanced than the ones in Fig.3. However, the response of the wall surface section, especially the north side wall surface where the floor panel was supported, was large. Therefore, if the balance of vibration response in these sections is adjusted more, the performance is expected to improve more.

According to the effect of the sound insulation ceiling in Fig.5, around 10dB of sound insulation was obtained in the lower frequency band. In the case of wooden houses, this type of independent sound insulation ceiling is said to be indispensable, as comparatively large air layer can be obtained behind the ceiling. In an estimation of the floor impact sound level in part 1, the effect of the sound insulation ceiling was set to 5 dB at 63Hz band in consideration of safety. According to this result, however, we assumed that around 10dB, which is near to the value of sound transmission loss (11.8dB) based on the mass law, is expected to be obtained.

2.3 TEST RESULTS AT THIRD STAGE

In the third stage, the wall surface section was remodeled further by referring to the vibration measurement results of Fig.4 in the second stage. In this stage, the rigidity of the wall surface section was increased by the addition of a stud (105 x 35mm) between each conventional stud (105 x 35mm) (Fig.6). The sound insulation performance of the ceiling was increased by the addition of a layer of 12mm thick plaster board, which has 0.5mm thick lead, on top of the two-layer plaster board by nails. Fig.7 shows the comparison of the floor impact sound measuring results at each stage. Fig.8 shows the measuring results of vibration responses at each section of the room below the floor construction.

According to Fig.8, the increase of wall surface rigidity reduced the vibration level of the wall surface at the frequency band of 63Hz and over, and the responses at each section, including the ceiling surface, were well balanced compared to Fig.4 in the second stage. Especially, the reduction at the north side wall surface was remarkable, in that the response at each section was within 3dB in the frequency band of 63 Hz and over.

As a result, according to the floor impact sound level in Fig.7, it was reduced by 4dB in the 63Hz band compared to the second stage, which indicates that the sound insulation performance was more than L-55 (L-53 in L number) in the sound insu-
3. CONCLUSION OF STUDY AND FUTURE RESEARCH POLICY

As a result of investigating into an improvement method for floor impact sound insulation against heavy impact sources by making high rigidity, which was conducted by using a full-size model wooden house, the following results were obtained:

1. It is important to consider providing wooden floor sections as thick as possible and making a complete integrated structure to increase geometrical moment of inertia in order to realize high rigidity.

2. It is considered to be convenient from the standpoint of manufacturing and accuracy to manufacture the floor panels elsewhere, rather than making them at the construction site, in order to secure the high rigidity performance of paragraph (1).

3. For the ceiling of the room below the floor structure, it is indispensable to provide an independent vibration-proof sound insulation ceiling, the ceiling carrying rods of which are supported by walls in such a manner as to prevent vibration transmission. It is possible to obtain sound insulation of 10dB in a low frequency band, even at a regular air layer of a wooden house and with two layers of 12mm thick plaster boards.

4. The effect of radiated sound from the wall surface of a room below a floor construction is very large, and thus it is impossible to obtain a sound insulation performance of L-65 or above without providing some measure to make highly rigid wall surfaces.

5. It is necessary to balance the radiation of sound from all the wall surfaces, including the ceiling of the room below the floor construction, as a countermeasure to the floor impact sound of a wooden structure, and partial measures for the floor impact sound does not bring any effective results in many cases.

6. If the idea of paragraph (5) is used effectively, the sound insulation performance of L-55, which is equivalent to the one that the 15cm thick concrete floor of 15~20m has, can be obtained against a heavy impact source without changing the conventional construction method greatly.

7. It is important to adhere the plywood completely, which are applied on the installed floor panel as the surface material for the floor panel integration. According to the test results, the sound insulation effect is improved by at least one rank (5dB) by this method.
THE PREDICTION OF NOISE FROM MACHINE VIBRATION

R. Walker, C.D. Mathers and J.A. Fletcher

British Broadcasting Corporation, Research Dept., Kingswood Warren
Tadworth, Surrey, England, KT20 6NP

1. Introduction.

In principle, the prediction of airborne noise resulting from machine vibration requires knowledge of the source of the vibrations, of the structural transmission path between source and receiver and of the efficiency of the conversion of vibration to airborne noise. In practice, the transmission path and the conversion efficiency may be combined into one overall transfer function. Practical values of these parameters have been measured experimentally in detail to produce predictions for the noise levels in rooms resulting from a nearby vibration source. A real case has also been investigated in the field. All of the measurements were carried out using an FFT analyser giving 400-point spectra over the range 0–500Hz.

2. Basis of the method

Fig. 1 shows an impedance analogue of the complete system. In general, the machine will have an output impedance, \( X_m \), the antivibration mountings (avms) and the building structure will both have a complicated equivalent arrangement of series and parallel elements. The building structure will, additionally, contain the component representing the radiation of the sound as airborne noise.

This complete equivalent circuit may be considerably simplified. Firstly, because measurements may be carried out which characterise both the transfer function for vibration velocity to sound level and the input impedance (at point A in the circuit), there is no need to consider the details of the 'structure' block at all. Secondly, if the 'anti-vibration' system can be represented by a single perfect compliance which has an impedance very much less than any other impedance in the system (particularly, less than the machine output impedance) then the avm block can be replaced by a single element and the machine impedance neglected.

If \( T \) is the transfer function from velocity to sound level, in \( \text{ms}^{-1}/20\mu\text{Pa} \)

\[ \frac{d}{d t} \text{ is the machine displacement at its foot, in m} \]

\[ J \text{ is the roof impedance, in N/m} \]

\[ S \text{ is the stiffness of the avm system, in N/m} \]

then the predicted sound pressure level (sp) in units of 20 \( \mu\text{Pa} \) is given by:

\[ \left( \frac{d \cdot S}{J} \right) / T \]

The calculation is carried out for all of the frequencies in the 400-point spectra. The resultant narrow-band prediction may be converted to \( 1/4\text{th} \) octave by summing the contributions in each idealised frequency band.

3. Experimental Investigation.

The measurements outlined above were carried out in a laboratory situation and the predicted noise levels compared with measured spl values. The machine was simulated by mounting a large electromagnetic shaker on suitable commercial vibration isolators. The 'machine' displacement was measured directly with an accelerometer. The transfer function and 'foundation' impedance were measured together using a calibrated force generator (hammer) on the mounting point and a microphone in the receiving area. The avm stiffness was deduced from the system fundamental resonant frequency. Fig. 2 shows one set of results.
4. Field measurements

The field measurements actually pre-dated the experimental confirmation of the method which was developed in response to an urgent real requirement. The practical problem was that of the installation of building services plant very close to a broadcasting studio. In this case, the building exterior formed the upper part of the studio. On an adjacent roof, alongside and structurally connected to the studio wall at a distance of less than 0.5 m was the proposed location of a pair of 300kW (thermal) chillers, each containing a 75kW electric motor and multi-cylinder reciprocating pump, together with four powerful, high-speed fans for moving the cooling air.

The three factors required for the prediction of potential noise levels were reasonably easily measurable – a nominally identical unit was already working nearby and the structural mounting points for the plant already existed. The measurement of the plant vibration displacement presented no problems. However, those of the transfer function and impedance were more difficult. At the time, no calibrated force transducer capable of generating measurable noise levels in the studio was available. The measurement was therefore carried out in two steps, using a large sledgehammer to measure the velocity-to-sound level transfer function and a small force transducer to measure the mounting-point impedance. (For later work, the sledgehammer was calibrated for use directly as a force generator.) One problem was the tailoring of the excitation spectrum to give sufficient low frequency energy. This required the selection of different types of resilient pads between the hammer and the hard concrete upstands.

Fig. 3a shows one of the predicted spectra calculated from one set of the measured data, together with the design criterion.

5. Omissions and assumptions.

One assumption built into all of this work is that all of the avms are perfectly behaved. For the measurement of machine displacement it is assumed that the avms do not affect the vibration levels. This will be true if they are very compliant compared with the machine structure at all frequencies. For the predictions, the assumption was made that the avms were perfect with the same compliance at all frequencies. For small, well-behaved mounts the dynamic-static compliance ratio may be allowed for, if it is known. For large and heavy avms with modal behaviour in the frequency range of interest, the problem will generally be too complex to allow for. For the measurement of displacement these modal effects will cause the prediction to be too low at some frequencies because of the consequent under-estimate of the source force. For the actual installation, they will be such as to increase the actual noise levels at some frequencies. For both reasons, the use of avms showing pronounced modal behaviour should be avoided.

Another implicit assumption is that of the correlation between different transmission paths. For these predictions, the entire machine was assumed to be on one avm with the aggregate characteristics of the six beneath the existing machine and positioned at one of the mounting points. It would be more accurate to analyse many paths and take the vector sum of their contributions.

6. Further work

Following the prediction of high levels of interference in the studio, the proposed installation was modified to include additional stages of vibration isolation. This greatly complicated the prediction, to an extent where the uncharacterised, additional components rendered further progress impossible. Measurements were carried out after the completion of the installation to enable a retrospective 'prediction'. Fig. 3b shows the result. The purpose of this retrospective work was to obtain information about the additional stages of vibration isolation, further to test the prediction method in the field and to allow the final result to be interpreted more meaningfully.
7. Results, discussion and conclusions

Fig. 4 shows (a) the measured studio noise levels, (b) the structure-borne noise levels predicted as described in Section 6 and (c) the calculated studio noise levels based on airborne noise and airborne sound level difference measurements.

Comparing first the 'predictions' of the noise via the two paths. At frequencies of 63, 80 and 100 Hz there is a slight dominance of the structure-borne noise of between 2 and 14 dB. At 125 and 160 Hz, the two predicted contributions are about equal. Above 160 Hz, the predicted structure-borne noise falls very sharply to become insignificant in comparison with the airborne component.

Comparing both predictions with the measured studio noise levels, the agreement is sporadic at best. At low frequencies, where the noise was predicted to be structure-borne, the difference between the prediction and the result takes a range of -2.5 to 11 dB and an average value of 4 dB. At higher frequencies, where the airborne component was predicted to be dominant, there is more correlation but a probably significant, consistent error of 4.1 dB average. This error was most probably due to the change in the acoustic fields as a result of the installation of the plant.

All of the field measurements are based on very few sample positions because of the short timescale and the awkward situations in which the measurements were made. They are therefore subject to large statistical errors. In the much better controlled environment of the laboratory, the method shows reasonable agreement, although with some errors of up to about 10 dB.

The overall conclusion which may be reached is that the measurements are practicable in the field and that the prediction method is accurate enough to form the basis of informed design decisions. Potential errors exist in the difficulty of generating measurable noise levels, especially in the presence of existing noise and in the possibility of other coupling paths such as pipework, conduits, etc. The performance of real vibration isolation systems, in particular, modal behaviour is also a potential source of serious errors.

8. References


9. Acknowledgements

This paper is published with the permission of the Director of Engineering of the British Broadcasting Corporation.

The consultant responsible for the acoustic design of the plant installation was Dr. M.S. Langley of Sound Research Laboratories.
Fig. 1. Complete system equivalent circuit.

Machine vibration

Avm system

Structure

Fig. 2. Experimental results

Fig. 3 Predictions of structure-borne noise levels

Fig. 4. Predicted and measured noise levels
EQUIVALENT FORCE AND ITS APPLICATION IN STRUCTURE-BORNE NOISE BETWEEN MACHINE AND BUILDING

Ing. Jan Stěniška, CSc

National Research Institute for Machine Design
190 11 Praha 9 - Březnice, Czechoslovakia

SUMMARY

Until recently has the equivalent force become the only feature to characterize the machine as a structure-borne noise source. It can act as complete substitute for three quantities used so far: velocity on the machine's base-supported footing, internal source impedance and reference impedance of its base structure. The computation consists in adding levels.

INTRODUCTION

Until recently has the equivalent force become the only feature to characterize the machine as a structure-borne noise source. It can act as complete substitute for three quantities used so far:

1) velocity on the machine's base supported footing $v_F$
2) internal source impedance $Z_M$
3) reference impedance of its base structure $Z_p$

This force $F_{eqv}$ or its level

$$L_{Feqv} = 20 \log \frac{F_{eqv}}{1 \text{ N}}$$  (dB)

cannot be measured straight away and does not correspond to the size of the force transported to the adjoining building structure. Its size can be experimentally assessed only by engineering methods or civil engineering acoustics (e.g. National Research Institute for Machine Design) and, exceptionally by computation from available data.
\[ L_{Feqv} = L_{VF} - 10 \log \left( 10^{\frac{-L_{ZM}}{10}} + 10^{\frac{-L_{ZF}}{10}} \right) - 180 + 20 \log N \] (dB)

where \( L_{VF} \) is the velocity level on a machine's base supported footing between the machine and the building.

\( L_{ZM} \) is the module level of the inner source impedance.

\( L_{ZF} \) is the module level of the reference building impedance at the point where it was ascertained.

\( N \) is the number of comparable places of contact between machine and building.

Valid remains the existing condition that in solid places of contact between machine and building the vibration prevail only in one predominant translating direction. While using the equivalent force, the condition \( Z_M \ll Z_F \) is negligible.

THE EMPLOYMENT OF EQUIVALENT FORCE TO CALCULATE BUILDING STRUCTURE BORNE NOISE

The main purpose of calculations of structure borne noise is to find out the magnitude of vibrations, i.e. the vibration velocity of a building surface within audible frequencies. The vibration magnitude is further used [1] to assess to radiated noise quantity in protected building room. The vibration magnitude is expressed by velocity level

\[ L_V = 20 \log \frac{V}{10^{-9} \text{m.s}^{-1}} \] (dB)

and largely attains values from \( 10^{-4} \) to \( 10^{-6} \text{ m.s}^{-1} \).

The magnitude of vibrations transferred to part of buildings \( L_{Vk} \) where the machine is attached will be calculated from

\[ L_{Vk} = L_{Feqv} - L_{Zk} + 180 - 20 \log \left( \frac{A_{24}}{N} \right) \] (dB)

where \( L_{Zk} \) is the module impedance level of the part of the building as in the computing method used so far [1].

\( A_{24} \) is the transfer parameter of the spring or another damping element on the link between the machine and building structure [1].
Fig. 1 Comparison of computed and measured vibration data on the floor during machine operation

\[
\Delta_1 = L_{v_k}(v_k, x_m, z_f) - L_{v_k}(\text{measured}) \\
\Delta_2 = \frac{L_{v_k}(f_{eqr})}{\text{computed}} - L_{v_k}(\text{measured})
\]

DETERMINATION OF EQUIVALENT FORCE

The equivalent force is now computed as a product of mean velocity in immediate vicinity of the contact between machine and building structure and what is termed the common admittance of the machine and building. In the process the measured difference of velocities on the side of the machine and building is mathematically considered. The so called impedance is usually found out by hand hammer blow. The method also takes the orientation features of the building structure mobility (without the machine) into account [4] and is the subject of further research.

CONCLUSION

The employment of equivalent force to characterize the machine as a source of vibrations substantially improves the accuracy of the structure borne noise computation especially in the first stage, i.e. in the transfer of vibrations to that part of building structure where the machine is situated.
The size of the equivalent force for different machines can be found in [2].

LITERATURE

[1] Data to be used in design and assessment of industrial buildings; Part 3 - Civil engineering acoustics, Instructions of Research Institute for High Buildings, Praha, 1985


DESIGN OF PARTITION WALLS WITH HIGH STANDING INSULATION VALUES IN THE RECONSTRUCTED ARCHITECTURAL STRUCTURE

Prof. dr. EDO ŠMIDIHEN, architect

Faculty of Zagreb University of Zagreb
41000 Zagreb, Kašićeva 26, Yugoslavia

In the reconstruction of traditionally erected buildings as well as in changing purposes and substantial modifications in recently constructed buildings arises the problem of acoustic isolation with regard to the completed state and unchangeability of the primary supporting system of the building. As a rule, only partition walls can be dislocated or constructionally redefined whereas the supporting elements of horizontal and vertical strucution, facade-walls and the arrangement of windows are preserved in their integrity.

In practice two kinds (types) of building arise as essential: traditionally erected buildings of the structural system with supporting walls of greater mass (brick, stone) and a ceiling structure made of timber joists, and buildings of modern construction design in the system of skeleton structure made of reinforced concrete with ceiling structure made of reinforced concrete slabs or a combination of joists and slabs. In the course of the reconstruction of traditionally erected buildings ceiling structures should be transformed into stiff diaphragms (reinforced concrete) because of seismic stability thus becoming partitions of satisfactory mass concerning airborne sound insulation.

**Buildings with system of massive walls**

On the assumption that supporting ceiling structures of satisfactory weight be erected the structural system makes spatially closed wholes whose peripheral structures are heavy massive partitions of 350 kg/m² surface mass. These partitions are favourable concerning the isolation of airborne sound.

The acoustic problem is related to the erection of partitions that create rooms within the space of the primary structure. These partitions should not be heavier than 70 kg/m². Partitions with sound insulation values $R = 50$ dB ($I = -10/-3$ dB) are available on the market that offers a range of systems of light double leaf partitions (Richter and the like). In some cases light stiff walls with uni- or bilaterally hung forewalls "resilient skin" can be erected.

Solutions how to achieve higher insulation values ($I = 0$ to $+5$ dB) should be looked for in the erection of the double partition with staggered studding where the limitation factor is defined by the limited dimension of interspace within the partition structure. As a
rule, partitions are erected between the supporting heavy walls so that the unfavourable flanking sound transmission does not appear.

**Skeleton construction buildings**

Primary supporting structures of a building do provide satisfactory sound insulation of airborne sound only by ceiling structures. This fact enables a relatively simple erection of acoustic isolation between rooms in vertical sequence. Rooms in a building are sound transparent in horizontal direction. Closing the space with light front construction causes the additional problem of flanking sound transmission.

In buildings with modest requirements concerning sound insulation (administrative offices etc.) satisfactory solutions may be achieved by building the elements of light partition walls that are produced on a large scale with isolation values $I = -10$ dB. But these values may be increased either by building double-layered skin walls of a double partition or by insertion of an additional 0.8-1.5 mm thick leaden sheet into a double-wall system. Isolation values $I = -4/-1$ dB on condition that partitions be erected between pillars may be expected.

However, free disposition of partition walls within the open space of skeleton system or continuing area between supporting walls of the building requires developing of partition as a concept of consequent realization of discontinuity in the quasi-box-in-box system on condition that the value of acoustic isolation of airborne sound be higher than $\bar{R} = 52$ dB.

**Fully discontinuous construction**

Taking into account the need of limited weight the optimum solution would be the erection of manifold double stud constructions with surface walls varying in material, thickness and width of interspace.

Two realized cases demonstrate approach to such a solution. In both cases acoustic isolation greater than $I = +2$ dB was required between rooms in a sequence. So partitions for sound transmission loss $R = 55$ dB were erected because of the predicted influence of flanking sound transmission.

**Case I**

Within the free space between two massive stone-walls (50 cm, 900 kg/m²) rooms in a sequence should be located and separated by partitions with sound insulation values $\bar{R} = 55$ dB. An accompanying corridor is designed as a link between them. A massive stiff wall resistant to mechanical blows was necessary between the corridor and the rooms because of the characteristics of traffic in the passage.
Partitions are erected as double staggered stud constructions with double - leaves: 12 mm gypsum board and 14 mm chipboard panel attached to the separated supporting studs 5/12 cm at a 60 cm distance. Gypsum board and chipboard are connected into a double sheet by 10 nails per m². On the one side the chipboard was covered by 0.8 mm sheet. The distance between the outside partition leaves is 28 cm. A 2x5 cm thick mineral wool blanket was introduced into the cavity between the walls. A gap of 2 cm was realized in the floor construction.

In the partition towards the corridor one of the described structures was substituted by a 12 cm thick brick wall plastered on both sides. This wall was erected on the passage side. The timber wall was erected as a part of the described structure with double gypsum board sheet towards the room. Between the massive and the timber wall freely hung a sheet of heavy bituminous felt. Transversal brick walls were erected in every second double-leaf partition in order to provide stability of the wall in the corridor. Thus an absolutely closed continuity of separate single walls was realized with a negligible flanking transmission. The obtained values of sound insulation between rooms are $R = 57$ dB (I = +6 dB).

Case II

Changing the building of an administrative office having skeleton structure into the building with special purpose, the already existing partitions could not provide sufficient sound insulation between the rooms. Partition walls were constructed as light-double-walls with a single supporting structure of tin profile studs. The front of the building was constructed as a curtain wall so that the flanking sound transmission was considerable and the sound insulation about 40 dB (I = -11 dB).

In order to achieve the necessary value I = +3 dB the following reconstruction took place: the existing walls were retained and additional partitions as the implementation of the concept of fully separated partition system erected. The additional partition was constructed as a double-leaf partition with highly damped inner space and a unilaterally erected quasi skinwall. The main supporting structure of timber studs 50/120 mm at a 55 cm distance with bilateral linings: on the side of the existing partition wall a 12 mm thick chipboard and towards the room a 5 mm asbestos-cement board. In front of it was fixed a double-leaf skinwall on the propped up small boards. The skin was made of a 12 mm thick gypsum board and a 12 mm chipboard. The already existing partitions were erected beside the pillars of reinforced concrete of supporting structure of the building. The new fitted partitions were erected between or beside the pillars so that the inside walls formed a "closed box".
A special problem was the linking of the front wall with the transversal walls. Because of the limited width of the partition leant upon window verticals there was erected another partition instead of the already described partition. It was a 2 mm thick leaden sheet coated with a mineral wool blanket and the external lining made of perforated plywood board.

The obtained values of sound insulation between rooms are
I = +1/3dB (an improvement of 12-14 dB was achieved). These values are obviously lower than the real ones because of the low sound insulation value of the front wall and continuity of a floating floor.

Conclusion
The new spatial arrangement of partition walls in the already existing primary structure will provide satisfactory sound insulation values if constructed as double-leaf partitions with multi-layered flexible sheets. Higher insulation values (I ≤ + 3 dB) could be achieved only in the concept of fully discontinuous double construction when a separate internal "independent box" of each room is formed.
SOUND INSULATION OF MULTIPLE GLAZING FILLED WITH LIQUID
T.I. Galaktionova, K.B. Rozova, S.P. Solovjov
Moscow, U.S.S.R.

The protection of rooms against external noise is an important index of building performance characteristics. As is generally known, window is the weakest point when ensuring necessary sound insulation of the external constructions. It applies equally to the buildings situated near highways with intensive traffic.

In this case the level of noise penetrating through windows can exceed the permissible level by 10-30 dBA depending on traffic intensity. In Moscow the noise levels can reach 62 dBA. The rational solution of the problem of buildings' protection against traffic noise requires the correct choice of window constructions. The optimum variant of such construction is the multiple glazing. Presently the multiple glazing, including special insulating glass units, are widely used in various buildings in different climatic regions.

The multiple glazing consists of two glasses connected by a distance frame and hermetics so that a hermetically space forms between the glasses. The frame is made of aluminium; the vulcanizing hermetics and non-hardening mastics can be used as hermetic materials.

In many cases the solution of noise control problem is connected with an attainment of the optimum internal microclimate. One of the possible solutions here may be the use of an insulating glass unit filled with liquid.

The construction tested represented a multiple glazing 100x500mm; the width of inter-glass space is 1.5mm; thickness of glass is 5 mm. The whole inter-glass space was filled with thermochromic solution on the base of isopropyl alcohol.

The filled insulating glass unit is an automatically regulated type of solar protection device, owing to the thermochromic solution's ability to change its colour (optical density) depending on temperature; this allows to maintain the required light and temperature conditions in a room all day along.

The sound insulation tests of special multiple glazing filled with liquid have been carried out in the conditions of big reverberation chambers according to the standard methodologies with the application of electroacoustic precision apparatus produced by "MKR" (GUM).

Fig.1 presents the frequency characteristics of sound insulation when filling the inter-glass space with air and liquid. One can see from comparison of the curves 1 and 2 that the filling of space with liquid provides a considerable increase of sound insulation practically in all dispaion investigated. Simultaneously the character of sound insulation frequency dependence smooths out a bit, and the critical frequency of wave coincidences for glass removes into a high frequencies range. Filling of space with liquid increases considerably the losses inside the multiple glazing causing a great increase of sound insulation. On low frequencies sound insulation grows with the increase of liquid layer density, and on high frequencies - with the increase of layer attenuation constant. It is known that the sound insulating qualities of external walls with window openings are determined by the insulation of window assembly against noise, and in lower degree they depend on the blind wall construction.

Traffic noise, as it is known, has a low-frequency character. As the result of it, the window sound insulation in the low and middle frequencies range takes on special significance, while the high frequencies range influences little on the sound level in dBA of the noise penetrating into a room.

Since the calculations of expected traffic noise levels and the measurements of city noise parameters are carried out according to the noise levels in dBA, it seems advisable to carry out the sound insulation valuation also in dBA. The method of sound insulation valuation based on the
use of some "standard" spectrum of traffic noise is proposed by the institute MNIITEP (Moscow). For that it is necessary to subtract the values of air noise insulation from the levels of the given spectrum and to determine the sound level in dBA according to the spectrum of penetrating noise obtained. The difference between sound levels corresponding the standard spectrum and the penetrating noise spectrum is the window sound insulation \( R_A \), dBA. The noise with the level of 75 dBA is assumed as "standard" spectrum. The value of sound insulation of multiple glazing filled with liquid makes 33 dBA when calculated by this methodics. Therefore the sound insulation of multiple glazing can be increased considerably with the help of liquid layer. That's why this insulating glass unit may be recommended as the protection against traffic noise, especially in the IV and V climatic regions of the USSR.

Sound insulation frequency characteristics of multiple glazing with different filling of inter-glass space

1. Filling with liquid
2. Filling with air

Fig. 1
REPRESENTATION OF THE SPEECH SIGNAL WITH ELEMENTARY WAVEFORMS: A PRELIMINARY PERCEPTIVE STUDY
S.GROVEL, J.S. LIENARD, C. D'ALESSANDRO
1.I.M.SI-CNRS  B.P. 30 - 91406 ORSAY Cedex - France

INTRODUCTION

Decomposition of the speech signal in the form of a sum of elementary functions (or elementary waveforms) provides an opportunity to apply powerful methods to the joint time/frequency representation. This type of representation, by a set of discrete elements that are well-located in spectro-temporal domain, can be found in communication theory (Gabor transform, short-term Fourier transform, wavelets) and in speech (auditory models [ref1], production models [ref2]). The general form of such an expansion is:

\[ x_n = \sum_{n} \sum_{k} c_{n,k} f_{n,k} \]

where \( c \) and \( f \) both depend on time (\( n \)) and frequency (\( k \)).

The \( f \) functions, elementary short-term waveforms with a limited bandwidth, appear as sinusoids modulated by a temporal envelope. Numerous variations of the envelope exist in the literature. The waveforms studied here are defined by six parameters (see Fig. 1): four envelope parameters -- reference instant \( \tau_r \), attack time \( att \), decay time \( dec \) (which can be an exponential or a sinusoidal decay), amplitude peak \( amp \); and two carrier parameters -- frequency \( freq \) and phase \( phi \), with respect to the envelope reference instant.

![Waveforms parameters](image)

Fig. 1 Waveforms parameters

This paper reports preliminary tests on the perception of parameters of several elementary waveforms. This study evaluates the perceptual pertinence of these parameters at a non-speech level with the perspective of using them in later automatic speech processing.

The tests concern the perception of temporal envelope parameters (attack and decay) for an isolated waveform, and the perception of the influence of attack and time delay for two simultaneous waveforms.

EXPERIMENTAL CONDITIONS

Each test stimulus consists of two tokens (X and Y) each having a duration of 10, 40, or 60 ms. The tokens were separated by a silence of 200 ms in order to avoid masking. Each stimulus was presented ten times consecutively. The amplitudes of the test tone pulses were normalized so that the total energy in X and Y was always the same.
Tokens X and Y differed only by the one parameter value being tested. Subjects were requested to decide whether X and Y were the same or different. When a positive response was obtained ("they are different"), the subjects were asked to describe in what way they differed. The tests were not forced choice, no judgement being solicited when a subject was unsure. In the results, the undecided responses are counted separately from the others.

Ten subjects from our laboratory, having no known auditory deficiencies participated in the tests. Each subject took all of the tests in a single 45-minute session. Tests were carried out using an IBM PC AT containing an UROS AU20 board with a 16 bit D/A converter at a sampling rate of 10 KHz. The volume was adjusted to a comfortable level, and the subjects listened through Beyer Dynamic DT48 headphones. The session was supervised and the sequence of stimuli was the same for each subject. At the beginning of the session, ten samples were played to the subjects to allow them to get accustomed to the tokens. Finally, since there was a large number of tokens, and they were very short, there were regularly spaced pauses to avoid auditory fatigue.

PERCEPTION OF ENVELOPE PARAMETERS

ATTACK PARAMETER

The perception of attack duration was tested at four different frequencies (166Hz, 333Hz, 500Hz, 1500Hz), for tones composed of a single waveform at durations of 10 and 40ms. For each stimulus, the X waveform had a fixed attack time (the first tone pulse of each stimulus was considered to be the reference), and the waveform of token Y varied in attack duration (see Figure 2).

![Fig. 2 Illustration of single-waveform tokens differing in attack duration.]

Figure 3-a shows the results at a fixed frequency of 500Hz and token duration of 10 ms, for differences in attack duration (between X and Y) of 1, 3, 4, 5, and 7 ms. Figure 3-b shows the results obtained under the same conditions for a token duration of 40 ms.

![Text and graphs related to the results of the experiment.]

Note that, for the 10 ms tokens, perception of the difference between the two tones happens for the majority of the subjects at a difference of 4 ms whereas for the 40 ms tokens, a difference of 7 ms was needed. Thus, there seems to be a close relationship between the total length of the tone pulse and the minimum perceptible difference in attack duration.

Figure 4 shows a similar relationship for two of the three other frequencies tested. For the 166Hz tones, this effect was not observed.
Fig. 4. Percent discrimination perceived for some differences in attack duration between tokens X and Y, for stimuli at three different frequencies (166, 333, and 1500 Hz) of various durations (10, 40, or 60 ms).

**Decay Parameter**

For tokens of 10 and 40 ms, shortening or lengthening the decay (while using the same type of envelope, the attack parameters, and constant total energy) causes a slight change in the perceived pitch, a phenomenon similar to that described by Rossing and Houtsmale [ref3], and Hartmann et al. [ref4]. For longer tokens (around 200 ms), a change in decay time does not provoke a "difference" response, except in the perception of the overall duration of the token, when the variations are long relative to the total token duration.

The exact form of the decay (sinusoidal or exponential) may, in certain situations, not have a prominent role to play in the token perception. For some reference tokens (X) of 10 or 40 ms, it is possible to adjust the exponential decay parameter of the test (Y) such that the two tokens are perceptually equivalent even though the shapes of their envelopes are different (see Figure 5).

Fig. 5. Example of two waveforms at 333 Hz frequency, which are perceptually equivalent even though the shapes of their envelopes are different.

**Perception of Grouping of Two Simultaneous Waveforms**

**Perception of Onset Disparities**

This experiment tested the perception of onset disparities in two-waveforms tokens.

The tokens were comprised of 166 and 500 Hz tones, or 333 and 500 Hz tones.

For 61 ms two-waveforms tokens, subjects perceived a difference when the 2 waveforms had a mean starting difference of 6 ms. The subjects perceived differences at offsets ranging from 2 to 7 ms depending upon the waveform frequencies. It also appears that shorter onset delays can be detected when it is the lower frequency component that is delayed. Most subjects only noticed the existence of two separate elements in the token when the delay was greater than 7 ms.

These results may be compared to those of Zwicker and Feldtkeller [ref5] which show that for a complex token of at least 50 ms in duration, temporal delay results in the perception of two successive and separate tones.
VARIATIONS OF ATTACK DURATION

In these tests, token X was made up of two waveforms having the same onset time, whereas token Y had a different attack time for one of its two waveforms (see Figure 6). When the difference in the attack times of the waveforms of Y is between 4 and 7 ms (depending on the waveforms frequencies all of the subjects perceived the tokens to be different. Some subjects perceived two separate components where only one had been perceived when the attack difference was smaller. These results show that the effects of different attack times are comparable to the effects of different onset times.

Fig. 6 Illustration of two-waveforms tokens differing in attack duration.

An analogous phenomenon was noted by Rasch [ref 6] on the perception of simultaneous notes. He showed that if two notes start simultaneously but have different attack times, the higher of the two notes may be detected more easily than when the attack times are the same, if it is the higher sound's attack time which is the shorter one.

For stimuli comprised of two simultaneous waveforms with frequencies harmonically related (125Hz, 250Hz), differences in attack are perceived more readily than in other cases. However, these results must be confirmed by other tests using waveforms at other harmonic frequencies.

CONCLUSION

Perceptual tests investigating the parameters of synthetic stimuli have been conducted. For single waveform stimuli, the shape of the temporal envelope, attack duration and type of decay, have been studied. The perception of grouping of two different waveforms which are time-delayed or which have different attack durations, has also been studied. These preliminary results raise questions about other phenomena such as overall envelope perception, the relationship between the way the envelope is perceived and the dimension of phase, and groupings of more complex objects made up of these elements.

REFERENCES

[ref1] LIENARD J.S. "Speech analysis and reconstruction using short-time, elementary waveforms", ICASSP-87
[ref2] D'ALESSANDRO C. and LIENARD J.S. "Decomposition of the speech signal into short-time waveforms using spectral segmentation", ICASSP-88
[ref6] RASCH R.A. "The perception of simultaneous notes such as in polyphonic music", Acustica 40, 21-33 (1978)
UN NOUVEAU MODÈLE ACOUSTIQUE DE PRODUCTION DE PAROLE.

M. Marayati, R. Carré, B. Guérin.

Institut de la Communication Parlée, INPG-ENSERG,
46, avenue Félix Viallet, GRENoble France
*) Centre d’Etudes et de Recherches Scientifiques,
BP 4470, DAMAS Syrie

INTRODUCTION

Dans cet article, nous rappelons les éléments d’une théorie (Marayati, Carré, Guérin ; 1988) mettant en évidence des régions spécifiques le long du conduit vocal. Ces régions ont des propriétés qui facilitent grandement la formalisation de relations simples entre la fonction d’aire du conduit vocal et les caractéristiques acoustiques de ce conduit, d’une part ; entre la fonction d’aire et les caractéristiques phonétiques des sons produits d’autre part. Un modèle du conduit vocal constitué de régions a été simulé sur ordinateur et une réalisation mécanique a été effectuée. Ces deux systèmes ont été testés pour produire des sons de parole : voyelles, transitions voyelle-voyelle, voyelle-consonne-voyelle.

THÉORIE DES RÉGIONS ET DES MODÈLE DISTINCTIFS


Dans un tube acoustique, fermé à l’une de ses extrémités et ouvert à l’autre, il existe des régions distinctives (R) ayant un comportement acoustique spécifique en ce qui concerne les transitions de formants. Ces régions sont déterminées par les passages par zéro des fonctions de sensibilité du conduit vocal uniforme pour les formants considérés. La figure 1 représente ces fonctions de sensibilité pour les trois premiers formants et pour un tube uniforme de 17 cm et d’une section uniforme de 5 cm.

Figure 1. Fonctions de sensibilité et régions pour un tube uniforme de 17 cm, en considérant les trois premiers formants.

Si les 2 premiers formants sont considérés, on définit alors 4 régions (AB, CD, DF, FA).

Le comportement acoustique d'un système à 4 régions est représenté sur la figure 2 par des homogrames calculés sur un tube sans perte fermé à l'une de ses extrémités et ouvert à l'autre. Ces homogrames représentent, pour chacune des 4 régions, les fréquences des deux premiers formants en fonction de la variation de la section de la région considérée, les 3 autres régions restant à leur valeur uniforme de 5 cm².

Figure 2. Homogrames obtenus à partir d'un tube uniforme par variation de la section de chacune des régions. La section varie de 0,005 à 14,14 cm² par bonds logarithmique. Le mode OTM est défini pour les valeurs supérieures à environ 1 cm², le mode TM pour des valeurs comprises entre 0,05 et 1 cm² et le mode TTM pour les valeurs inférieures à 0,05 cm².
Lorsque la section varie d'environ 1 à 15 cm², on voit clairement apparaître l'aspect monotone des variations de formants. Par exemple, lorsque la section de la région AB croit, les fréquences des formants F₁ et F₂ augmentent ; lorsque la section de la région CD croit, le premier formant augmente, le deuxième formant décroit. Dans ce cas, on considère que le conduit vocal forme un tout fortement couplé (mode de fonctionnement à un seul conduit : OTM - One Tract Mode). Le mode OTM correspond principalement à la production des voyelles. Lorsque la section varie d'environ 0,05 à 1 cm², le comportement change (mode de fonctionnement transitoire : TM - Transition Mode). Le mode TM correspond principalement à la production des constrictives. De 0 à 0,05 cm² on a alors deux conduits non couplés acoustiquement (TTM : Two Tract Mode). Le mode TTM correspond à la production des plosives.

On a représenté figure 3 la position des régions (8 dans le cas présent) à l'intérieur du conduit vocal. Le milieu de ces régions correspond aux points d'articulation des consonnes. Par ailleurs, les régions C, D, D' et C' sont les 4 régions de constrictions mises en évidence par Wood (1979) pour la production des voyelles.

Figure 3. Positions des 8 régions à l'intérieur du conduit vocal. Les points de constrictions lors de la production des consonnes /d/ et /k/ sont représentées (d'après Perkell, 1969).

Le caractéristique d'anti-symétrie du conduit vocal permet de mettre en évidence plusieurs conséquences exploitées lors de la commande en régions d'un modèle de conduit vocal : la compensation et la synergie. En effet, deux mouvements égaux et de même sens appliqués à des régions symétriques ne produisent pas d'effet acoustique ; deux mouvements opposés produisent un effet acoustique maximal. On peut alors en déduire une stratégie de commande générale en régions par contrôle vertical des dimensions de ces régions au lieu d'un déplacement horizontal de la constriction.
La théorie des régions et des modes distinctifs a été testée sur un modèle simulé par ordinateur. Dans ce cas, deux programmes différents permettent de calculer la fréquence de formant du système d'une part et le signal dans le domaine temporel d'autre part. Dans ce dernier cas, le conduit vocal constitue de 8 régions est associé à un modèle de source à deux masses. Par ailleurs, une réalisation mécanique a été effectuée. Elle est aujourd'hui commandée par des actionneurs contrôlés au moyen d'un ordinateur de type PC. La source est soit un haut parleur, soit une anche vibrante.

Les deux systèmes ont été utilisés pour produire différents types de sons : des voyelles, des plosives et des ensembles voyelle-voyelle, voyelle-consonne-voyelle. On a représenté figure 4, des trajectoires V-V obtenues par transition à bons logarithmiques entre deux configurations de voyelles.

Figure 4. Trajectoires dans le plan F1-F2 de transitions voyelle-voyelle.

CONCLUSIONS

Les nombreux résultats obtenus en synthèse de parole à partir du concept de régions et de modes distinctifs sont très encourageants. Une étude systématique des transitions V-V a été effectuée. De bons résultats en synthèse d'ensembles VCV ont été obtenus. La mise en place automatique d'une source de bruit en fonction de la constrictive et de la pression est à l'étude pour la production des constrictives. Ces études permettent la mise en place d'une nouvelle démarche de synthèse par règles.

BIBLIOGRAPHIE


376
ON THE EXTENT OF BITE-BLOCK COMPENSATION IN VOWEL ARTICULATION

Lawrence J Raphael* and Alice Faber*

*Herbert H. Lehman College, City University of New York, Bronx, NY 10468, USA
*Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, USA

Introduction
Several experiments have provided acoustic evidence that speakers are capable of compensating for the articulatory constraints imposed by a bite-block in their production of vowels (Lindblom & Sundberg, 1971; Lindblom, Lubker & Gay, 1979; Gay & Turvey, 1979). Although speakers do not compensate completely for the presence of a bite-block, they appear to do so immediately and in such a way that the vowels they produce are highly identifiable as the vowels they intend to produce (Fowler & Turvey, 1980). The acoustic effects of bite-block compensation have been well documented, but little is known of the changes in muscular activity which give rise to the acoustic and perceptual equivalence of normal and bite-block vowels. Of particular interest are the questions of the manner and extent of articulatory reorganization when speech production is constrained. Since compensation appears to occur for all vowels, and since different sets of muscles need to be recruited for the articulation of different vowels, it seems reasonable to hypothesize that speakers must be able to employ, instantly, a variety of muscular strategies to compensate for the presence of a bite-block. In the electromyographic (EMG) and acoustic experiment described below, we have begun to test this hypothesis.

Methods and Procedures
The single subject of the experiment, a speaker of New York City English, produced fifteen vowels and diphthongs (i, i, ey, e, s, s, ow, u, u, a, a, ay, aw, oy) in a /ap/ frame. Fifteen tokens of each utterance type (separated by tasks unrelated to this experiment) were spoken consecutively during the first half of the run in two conditions: with and without a 10 mm. bite-block. The utterances were repeated, in the same manner and conditions, later in the experiment (after another set of unrelated tasks), yielding a total of 30 tokens for each utterance type in each of the two conditions. We shall refer to the various conditions in the experiment as NOE (no bite-block, early), BBE (bite-block, early), NOE (no bite-block, late), and BBE (bit-block, late). About one hour separated the earliest and latest sets of utterances.

EMG signals were recorded from six insertion sites. These included one surface electrode (Orbicularis Oris: O01) and five bipolar hooked-wire electrodes, three of which monitored activity in the anterior portion of the Genioglossus muscle (GGA1, GGA2, GGA3), one which recorded from posterior genioglossus fibres (GGP), and one which recorded from the superior longitudinal muscle (SL). The EMG and acoustic signals were tape recorded for subsequent computerized analysis, employing the Haskins Laboratories Physiological Signal Processing (PSP) System for EMG data and the SGM program of the Interactive Laboratory System (ILS) for the acoustic data.

Results
The English vowels produced by the subject fell into two categories which are typical of speakers of this dialect: (1) those which are basically monophthongal in nature, marked by relatively minor shifts in formant frequencies and (2) those which are diphthongal in nature, marked by relatively large changes in formant frequencies. The latter category comprised what are often referred to as simple vocalic nuclei with offglides (ey, ow, s) and phone-

---

1The GGA3 placement was originally intended to access the styloglossus muscle. On-line monitoring of the signal during the run, however, as well as subsequent data analysis, indicated that genioglossus activity was recorded from this insertion.
logically distinct diphthongs (lay, aw, oy). Figures 1-3 show the formant frequencies for these vowel groups in the non-bite-block and bite-block conditions from the early and late parts of the experiment. The essential congruence of the vowel spaces and formant trajectories is somewhat mitigated by the results of the statistical analyses. Three-way analyses of variance indicated that the second formant frequencies of the vowels in the two conditions were significantly different (F[1,1]=79.9, p<.0001). Differences between first formant frequencies approached but did not reach significance (F[1,1]=3.56, p=.06). Nonetheless, we interpret these results, as have others (Fowler & Turvey, 1980; Lindblom et al., 1979), simply as an indication that bite-block compensation is not total, although it is sufficient for a speaker to produce vowels which are perceptually equivalent to those produced without a bite-block.

Figures 1-3: F1/F2 values for simple vowels (1A, 1B), vowels with offglides (2A, 2B), and diphthongs (3A, 3B). 1A, 2A, and 3A cover NOE and BBE; 1B, 2B, and 3B cover NOL and BBL. Solid lines connect bite-block values, and dashed lines non-bite-block values.

For the vowels in the NOE and BBE conditions, the EMG data provide clear evidence for the basis for the articulatory compensation inferable from the acoustic information. The vowels in question are those which are described as having a high front articulation

2 The EMG data for the BBL condition were not available at the time this paper was written.
(including diphthongs with a high front component) and the vowel /a/, generally described as a mid or high central vowel produced with the tongue tip in a retroflexed position. Let us consider the high front vowels first.

Figure 4. Average peak muscle activity for three GG sites (GGA1, GGA2, GGA3) and one SL site (U) in NOE and BBE. Values for each insertion are normalized in terms of percent of maximum activity.

Compensation for the vowels /i, i, ɛ/ and for the diphthongs /ay, aw, ay/ is evident in the activity of the genioglossus muscle which is primarily responsible for raising and protruding the body of the tongue (Smith, 1970; Raphael & Bell-Berti, 1975). Figures 4A-C show averaged peaks of genioglossus activity from all three electrode placements normalized in terms of the maximum activity for that muscle across all conditions. The data displayed are from the early bite-block and non-bite-block conditions. We note that there are greater average peak values for the bite-block vowels and diphthongs from at least two of the three electrode placements (and from all three for /ay/).

Compensation in the case of the vowel /ə/ (Fig. 4D) is manifest primarily in the activity of the superior longitudinal muscle, which is responsible for tongue retroflexion. In the bite-block condition this muscle reaches its maximum average activity, whereas in the non-bite-block condition average activity is less than forty per cent of maximum. There is also a small difference in GGA3 activity in the expected direction.

Discussion

The data of this experiment provide support for the hypothesis that articulatory compensation in bite-block speech is muscle-sensitive. We hope to provide further evidence in this regard in future EMG experiments in which we obtain data for compensation in other than high-front and retroflex vowels by recording activity from the styloglossus, palatoglossus and mylohyoid muscles. In addition, when the EMG data for the BBL condition are available, we should better understand if, and in what way, compensation is time-linked for the muscles we have reported on here.

Acknowledgements

This research was supported by NINCDS grants NS-13617 and NS-07217 to Haskins Laboratories.

References


ISOLATION OF MICROEVENTS IN SPEECH

*13th INTERNATIONAL CONGRESS ON ACOUSTICS*
*YUGOSLAVIA* • 1989 *

William J. Strong and David A. Berry, Department of Physics and Astronomy
Brigham Young University, Provo, UT 84602, U.S.A.

Introduction

In the past few years, increased interest has been shown in isolating acoustical events in the speech waveform as an initial step in the speech recognition process (e.g., Silverman, Zue). It is believed that by so doing, the time smearing of some important phonetic features can be minimized. However, the isolation of acoustic events in speech is a formidable task because of the extreme variability of events—some are very abrupt, while others are very slow.

We define a microevent (ME) as a portion of a speech waveform initiated by an input of energy such as that from a pitch pulse, a plosive burst, etc. An ME is terminated at the initiation of the next ME or at the end of a 20 ms interval. The duration of an ME is restricted to a range of 2 to 20 ms. This is based on the range of pitch periods expected and on the conjecture that other significant acoustical events will have durations of at least 2 ms.

An algorithm was developed to isolate MEs in the speech waveform. Preliminary tests of this algorithm have been performed on utterances from a male and a female. Test data included consonant-vowel (CV) pairs as well as running speech.

Method and Procedure

The speech signal was digitized at a sampling rate of 20 kHz. Originally, seventeen one-third octave, linear-phase, bandpass filters were used spanning the frequency range of 70 to 9000 Hz. However, slightly better results were obtained when sixteen additional filters were used with center frequencies corresponding to the crossover points of the original filters. Linear-phase filters were selected partly as a visual convenience, allowing the filter outputs to be precisely time aligned with the original wave. However, it was also found that a reliable way to mark the beginning of an ME was to land on a zero-crossing of one of the filters.

First, the algorithm marked relative maxima in the speech signal, but did not accept them if more closely spaced than 2 ms. Each of these marks was treated as a preliminary microevent (PME). A histogram was generated for each PME based on spacings and amplitudes of nearby relative maxima in the filters. Ideally, maxima in the histogram would correspond to the period in voiced speech and to intervals between bursts of energy in noise.

In voiced speech, the histogram could easily yield double or triple the true period. For this reason, the five largest histogram peaks were chosen as period estimates. Then the value $S(n, K)$, previously used by Silverman, was calculated, where $K$ is one of the period estimates and $x(i)$ is the speech signal.

$$S(n, K) = \sum_{i=n+1}^{n+K} x^2(i) - \sum_{i=n-K}^{n-1} x^2(i)$$

$$\sum_{i=n+1}^{n+K} x^2(i) + \sum_{i=n-1}^{n-K} x^2(i)$$

$S(n, K)$ should be close to zero for $K$ equal to any multiple of the period. The true period was chosen to be the lowest period estimate with a histogram entry within 75% of the maximum histogram entry and with an $S(n, K)$ magnitude that was comparatively low. This method yielded the true period for nearly all of the cases tested in voiced speech. Furthermore, errors were generally isolated, with correct period estimates leading and trailing them.

Once the ME-length was determined, $S(n)$ was determined for each point in the original signal, with $K$ set to half the ME-length. Then, relative maxima were located in $S(n)$, but
were not accepted when more closely spaced than 2/3 of the ME-length. Each maximum was associated with one of the original PMEs. Each PME that could be linked to a maximum of \( S(n) \) was called an ME. The exact starting (point of the ME was determined by looking through the filters to find the largest two relative maxima that were close to the PME. The zero-crossing immediately preceding one of these maxima marked the beginning of an ME. Proximity to a PME, amplitude of the relative maximum, and size of the corresponding \( S(n) \) value determined which zero-crossing to choose.

**Results**

In this summary paper only preliminary results of the ME markings are presented. Figures 1-5 illustrate marking of MEs for sections of male utterances. Figure 1 illustrates a 40 ms section of a speech waveform for the /I//z/ sounds in "shells." Five MEs are marked with vertical lines in the figure. The outputs of all 33 filters are shown (on the same time scale) in the lower part of the figure, again with vertical lines indicating five MEs. Based on the outputs of filters 29-33 it appears that MEs 4 and 5 should have been marked somewhat earlier in time. Figure 2 illustrates a transition from unvoiced to voiced speech in the /sh//e/ of "shells," where six MEs have been marked. The four ME markers (MEMs) in the unvoiced section of the wave are placed somewhat arbitrarily as is to be expected because there is no periodicity to cue them. The fourth ME is rather long in duration because there is insufficient energy in the excitation to initiate a new ME until a voice pulse initiates the fifth ME.

The /u/-waveform in Figure 3 is nearly periodic and presumably associated with stable pitch pulses occurring at regular intervals. (The primary deviation from periodicity is caused by slight period-to-period changes in wave-shape.) Five MEs are marked in the figure at regular 8 ms intervals. The /i/-waveform in Figure 4 is also nearly periodic with a period of approximately 8 ms. However, in this case the change in wave-shape (and its related effect in the filters) has resulted in the misplacement of MEMs. One might judge, for example, that MEMs one and three should be shifted to somewhat earlier times to be consistent with MEMs two, four, and five. The /I/-waveform in Figure 5 illustrates a difficult, but not atypical, problem in event marking. Does the portion of the waveform between MEMs two and three represent a long voiced event or a normal length voiced event followed by a shorter voiced event? Apparently something erratic was happening in the voice production at this point. At least the event marking scheme has marked MEs three and four correctly even after the difficulties.

The waveform from a female talker in Figure 6 illustrates the transition from a plosive aspiration into the vowel /e/. The first ME (which is in the aspiration) and the second ME at the apparent beginning of voicing appear reasonably marked. However, one would think that there are two additional MEs between MEMs two and three which have not been marked. The female /u/-waveform in Figure 7 illustrates well marked MEs. This has occurred even though the positive pulse of the fifth ME is small and distorted relative to those of the other MEs.

Even though the algorithm performed well over the speech samples tested, most of the preceding examples have illustrated difficulties which must be dealt with if the algorithm is to be useful.

**References**


FATIGUE EFFECTS ON VOWEL ARTICULATION

Alice Faber and Lawrence J. Raphael

Haskins Laboratories, 270 Crown St., New Haven, CT 06511, USA
*also Herbert H. Lehman College, CUNY

Introduction
As part of a series of experiments (Alfonso & Baer, 1982; Baer & Alfonso, 1984) designed to correlate various methods of determining tongue position and shape (ultrasound, electropalatography, x-ray) with lingual muscle activity levels, we recorded a lengthy session (c. 1.5 hours) in which one speaker of English (LJR) produced several repetitions of an elaborate set of non-linguistic tongue gestures, in addition to four blocks of /opVp/ pseudo-words. One pseudo-word block occurred at the beginning, one in the middle, and two at the end of the experiment. An additional variable, for both the gestures and the vowels, was the presence of a 10-mm biteblock during one early and one late vowel set. Our purpose was to compare EMG activation patterns for the vowels and the gestures, in order to extrapolate from the known tongue position for the gestures to its position for particular vowels; we were also interested in patterns of compensation for the presence of the biteblock (see Raphael & Faber, elsewhere in this volume). However, in order to compare the EMG activation levels for the 'same' vowel at different points in the experiment, we first had to compare the vowels themselves. To eliminate, we thought, the possibility that differing EMG patterns for a given vowel target might be correlated with differences of vowel color. It turned out that we could not completely eliminate this possibility (Faber & Raphael, 1988). Our purpose here is to discuss the differences among vowels and to offer some suggestions about the reasons for these differences.

Data and analysis
The vowel portion of the experiment consisted of four blocks containing 15 repetitions each of 15 pseudo-words of the form /opVp/, in which V ranged over /i, y, e, e, a, u, o, ow, ay, ey, ay, aw, /; the tokens of each pseudo-word were produced in sequence, with no carrier. For purposes of discussion, these blocks will be referred to as BBE (biteblock, early), NOE (no biteblock, early), NOL (no biteblock, late), and BBL (biteblock, late). There were a total of 900 tokens, 60 of each vowel. The non-speech tongue gestures were performed after the BBE and NOE conditions. EMG signals were recorded (using hooked-wire electrodes) from five lingual insertions: Superior Longitudinal (SL) and three anterior and one posterior Genioglossus (GG) sites. An additional signal, from the Orbicularis Oris inferior (OOI), was recorded with a surface electrode. The acoustic and EMG signals were recorded on parallel tracks of a 14-track FM tape recorder, along with a synchronizing signal. The acoustic signal was digitized at a sampling rate of 10kHz, with 6 dB/octave preemphasis. The EMG signals were digitized, with an effective sampling rate of 200 Hz. They were then calibrated, on the basis of a 300 μV reference tone recorded on all channels, and smoothed with a 35 msec triangular window.

The acoustic and EMG signals for each token were examined using the WENDY (Waveform Editor and Display) program at Haskins Laboratories. The following points were marked for each token: closure for the first /p/, release of the closure, onset of periodic voicing for the vowel, closure for the second /p/, and release of the closure; where possible, onset of the

1 The phonetic quality of /p/ for this speaker, [λj], is such that it must be analyzed with the other diphthongs /ey, ow, ay, ey, ay/.
2 The locus of these insertions was established by comparison with other published EMG studies using this subject (e.g., Raphael & Bell-Berti, 1974). The targets for the three anterior insertions were Styloglossus left side and anterior and posterior Genioglossus, and, for the posterior insertion, transverse and vertical intrinsic fibers. Additional target insertions, Styloglossus-right side and Mylohyoid, gave erratic signals and, hence, were not analyzed.
Results

As expected, there were durational differences among vowels, ranging from /a/ (180 msec) to /a/ (280 msec). Somewhat surprising, however, was a consistent, and strong, cross-block durational difference: vowels in the later three sets were shorter than their counterparts in the BBE block (see Fig. 1). A three-way ANOVA showed significant main effects of vowel (F[4, 828] = 451.041; p < 0.001), time (early or late) in experiment (F[1, 828] = 122.592; p < 0.001), and presence or absence of the biteblock (F[1, 828] = 97.723; p < 0.001), with all possible interaction effects equally significant. Further analysis of the NOE, NOL, and BBL conditions showed no significant difference among them (F[2, 618] = 1.215; p = 0.2975), accounting for the time-biteblock interaction in the earlier analysis.

While the acoustic analysis showed small but mostly significant cross-condition effects for F1 (F[1, 835] <time > = 51.581, p < 0.001; F[1, 835] <biteblock > = 3.558, p = 0.06) and F2 (F[1, 833] <time > = 9.225, p < 0.003; F[1, 832] <biteblock > = 79.866, p < 0.001), interpretation of these differences is hampered by their small size relative to the resolution of the ILS software. The F0 results are, in contrast, unambiguous. In addition to the expected intrinsic pitch differences among vowels (Peterson & Barney, 1952; Lehisite & Peterson, 1961), F0 had a strong tendency to rise across conditions. Except for /a, a, w, y/ average F0 for tokens of any vowel in any block was higher than for tokens of that vowel in any previous block. When presence or absence of the biteblock is discounted, the difference is absolute, as shown in Fig. 2; late tokens of a vowel have higher F0 than early tokens of the same vowel (F[1, 864] = 1273.789, p < 0.001), and the difference ranges from 9 Hz for /iy/ to 29 Hz for /a/.

The EMG signals were analyzed in two ways. Using the Physiological Signal Processing System (PSP) at Haskins Laboratories, a set of six ensemble-averaged tokens was created for each of the 15 pseudo-words in each of the four conditions, a total of 360 tokens. For each of these average files, an integral was computed for a time span of 110 msec, 100 msec prior to the line-up point and 10 msec following it. The resulting figure, in µV/Secs, provided a measure of muscular activation in the period during which the appropriate vowel posture was being achieved. While this muscular activation might have been expected to change in the course of the experiment, perhaps as a result of fatigue, no significant temporal effect was observed. The

---

3 The formant values returned by SGM tended to cluster in multiples of 78 Hz.
4 The F0 values, while substantially higher than those reported for male speakers by Peterson & Barney (1952), are compatible with those reported by Lehisite & Peterson (1961).
second analysis, of the individual, non-averaged tokens, produced clearer results. For each non-biteblock token, the EMG signal peak for each insertion was marked, using WENDY, and its amplitude in μV was computed. Level of muscular activity was highly correlated across the two conditions, with correlation coefficients ranging from .873 for the SL to .985 for the posterior GG. One-way ANOVAs showed for each insertion a significant vowel effect (p < .001). Subsequent analysis for each insertion was limited on the basis of post hoc paired t-tests to those vowels whose activity peaks were significantly greater than baseline, with p < .01. In both the NOE and the NOL conditions, there were strong correlations among the amplitude peaks for the three anterior GG insertions, ranging from .61 to .801 in NOE and from .583 to .677 in NOL. But two-way ANOVAs showed different patterns of variation for these insertions across the two conditions: activity peaks were higher in NOL than in NOE for the first (GGA2, Fig. 3B, F [1, 139] = 27.232, p < .0001), significantly lower for the second (GGA3, Fig. 3C, F [1, 219] = 38.629, p < .0001), and not significantly different for the third (GGA4, F [1, 138] = 2.302, p = .13). The posterior GG and OOI peaks were also strongly correlated in both conditions (NOE: r = .841; NOL: r = .857), showing significantly less activity in the NOL condition than in NOE (post. GG: Fig. 3A, F [1, 116] = 10.212, p = .0006; OOI: F [1, 275] = 20.225, p < .0001). The SL insertion, which showed significant levels of activity only for /a/ and /u/, also decreased in activity, but less significantly (F [1, 56] = 6.272, p = .01).

Fig. 3: Average EMG signal amplitudes (in μV) for 3 insertions (A: posterior Genoglossus; B-C: anterior Genoglossus). Solid lines connect NOE tokens, and dashed lines, NOL tokens. Only vowels with activity levels significantly above baseline in one or both sets are plotted.

Discussion
Given our design, it is not possible to provide a complete explanation for the results presented above; we did not (and could not) monitor all muscles involved in speech production. Nor did we monitor air flow (nasal or oral) or sub-glottal pressure, variations in either of which might have provided clues to which subsystems play a role in producing the F0 differences we observed. Even so, some tentative suggestions can be made. The duration differences, while they might reflect fatigue, might as easily reflect a different setting. The magnitude of the F0 increase is at the top end of the range observed from morning to late afternoon for males by Garrett & Healey (1987). This similarity in the magnitude of F0 changes suggests that our increase resulted from comparable factors, but that something about our experimental procedures compressed the effect of these factors into a shorter time span. That is, if Garrett & Healey's results are to be interpreted as reflecting the effects of cumulative fatigue, vocal or otherwise, so are ours.

5 Depending on the number of tokens per set that were adventitiously missing, df was 14, 217–220, and F ranged from 23.68 to 85.212.
6 Cf. Nittouer, McGowan & Beehler (1988), who found no significant time of day effects in F0 of male speakers.
The question remains, though, what in our experimental paradigm exacerbated these effects. There are two possibilities. One, of course, is physiological stress. The subject endured eight EMG insertions (one insertion, not discussed previously, never had an active signal); his mobility was restricted during the entire experiment; and he could not eat or drink while the electrodes were in place. The second possible explanation lies in the special demands that we placed on the speech production system. The subject was asked to produce, to the extent possible, uniform tokens of each pseudo-word. (The need for uniform tokens was a consequence of the intention to analyze ensemble-averaged EMG signals.) The duration results reported above suggest that he was successful, within experimental blocks. A result of this requirement, however, was that the speaker was focussing on the physical form of his utterances more than in ordinary speech. This non-normal focussing on form may have been sufficient to compress a day's worth of vocal fatigue into a morning. It is our intuition, pending unambiguous experimental results, that the nature of the speech materials themselves played a much greater role in producing the observed F0 increase than did physiological stress. (While our physiological measurement techniques were, indeed, invasive, they should have had no direct impact on those laryngeal muscles commonly implicated in F0 control, the Cricothyroid and the lateral Cricothyroidenoid, e.g., Honda, 1988).

Interpretation of the EMG results is less straightforward. The clear decrease in activity for four out of the six insertions analyzed can, indeed, plausibly be attributed to fatigue, but it is not immediately obvious that the same is true for the equally clear increase in one of the anterior GG insertions. What our data reflect, we believe, is changes in muscle fiber recruitment over time; fatigue in some fibers is compensated for by increased activity in other, less fatigued fibers. This interpretation is supported by the patterns of correlation among the three anterior insertions. The two insertions which showed significant change from NOE to NOL were highly correlated in both conditions, but the correlation decreased, from .801 (F [1, 14] = 23.213; p = .0003) to .677 (F [1, 14] = 10.995; p = .0056). (By way of comparison, the correlation between the OOI and posterior GG insertions was essentially unchanged, despite the significant decrease in posterior GG activity.)

Acknowledgements
We would like to acknowledge the technical assistance of Dr. Hiroshi Muta, Mr. Richard Sharkany, and Mr. Vincent Califano during the data acquisition and signal processing phases of this experiment. We would also like to thank Drs. Katherine S. Harris and Fredericksa Bell-Berti for helpful comments on an earlier draft of this paper. This research was supported by NINCDS grants NS-13617 and NS-07237 to Haskins Laboratories.

References
DIE ERKENNUNGSZEIT DER SPRACHLAUTE

Miroslav Gregurić - Momir Vujnović
Elektrotehnički fakultet Zagreb, Jugoslavija
Zavod za elektroakustiku


Abb.1. Darstellung einiger Vokale und ihrer Spektrogramme.


Die Prüfungen der Konsonanten auf die Verständlichkeit ihrer Segmente ergaben folgende Resultate: die Verständlichkeit der Konsonantensegmente ist von ihrer Position und Zeitdauer abhängig. Zuerst wurden die Segmente mit guter Verständlichkeit erforscht. Um ihre Position zu bestimmen wurden die Anfangsteile und Schlussteile der Konsonanten beseitigt bis die Verständlichkeit der Segmente noch bewahrt war. Damit wurde auch ihre Zeitdauer bestimmt. Der Anfangsteil der beseitigt wurde, hatte eine Zeitdauer bei Plosivlauten (p,t,k,b,d,g).

391
Abb. 2. Konsonantensegmente: AB = Anfangsteil, BC = gut verständliches Segment, CD = Schlussteil, AE = verlängerter Anfangsteil, ED = Restteil, BE = kennzeichnender Teil, AD = Zeitdauer des Konsonanten

von 5 ms und bei allen anderen, 10 ms ausser "r" und "l". Bei Konsonant "r" war sie 30 ms und bei "l" war sie 50 ms. Die Zeitdauer der gut verständlichen Konsonantensegmente betrug zwischen 35 und 70 ms. Die kürzeste Zeitdauer von 35 bis 45 ms hatten die Segmente von Konsonanten p,t,k. Eine Zeitdauer zwischen 45 und 55 ms hatten Segmenten von Konsonanten d,d,g,s,z, und alle anderen hatten eine längere Zeitdauer bis 70 ms. Die Zeitdauer der vereinzelt ausgesprochenen Konsonanten war zwischen 140 und 250 ms. Wenn man den Anfangsteil der Konsonanten "p" und "t" auf 15 bis 20 ms, und bei allen anderen auf 30 bis 40 ms, verlängert und dann beseitigt, wird der ganze Restteil von Konsonanten unverständlich, obwohl seine Zeitdauer über 120 ms ist (Bild 2). Das heisst das am Anfang des gut verständlichen Segmentes ein kennzeichnender Teil vorhanden ist. Er ist allein wegen seiner kurzen Zeitdauer nicht verständlich. Alle diese Prüfungen wurden auf den vereinzelt ausgesprochenen Konsonanten durchgeführt. Diese Prüfungen wurden mit den Konsonanten, die gekürzt waren durch die regelmässigen Ausscheidungen von kurzen Abschnitten, wiederholt. Das Resultat war fast gleich mit den Prüfungen von Konsonanten die nach der ersten Methode gekürzt waren. Die Zeitdauern von den verständlichen Konsonantensegmente und der gekürzten Konsonanten unterschieden sich nicht mehr als um 5 ms. Daraus kann man schliessen das für die Verständlichkeit oder Erkennung eines Konsonanten eine bestimmte minimale Zeitdauer des gekürzten Konsonanten nötig ist und das er den kennzeichnenden Teil, der auch gekürzt sein kann, einschliesst.
APPLICATION OF DIGITAL SIGNAL PROCESSING
ON THE ARABIC LANGUAGE

O. Abdel Alim, N. El Bognadly
Faculty of Engineering, Alexandria University, Alexandria, Egypt

Introduction
Vector quantizer (VQ) is a system for mapping a sequence of continuous or discrete vectors into a digital channel. For LPC, VQ involves coding the spectral coefficients for each analysis frame as a block and transmitting a single code rather than a sequence of scalar coded parameters. This achieve a saving of 20-27 bits/frame over scalar quantization. Using VQ in vocoders and speech coding make the application of speech signal processing in this area language dependent. For VQ vocoders in Arabic language, the construction of an Arabic codebook is important. In this paper, the methods of digital signal processing, specially the VQ technique of speech coding are used to construct an Arabic codebook. The codebook is tested using objective and subjective tests. Finally, speech signal processing techniques are applied to make an attempt to extract the glottal waveform for the voiced sounds contained in the codebook.

Analysis
There are several ways to perform an LPC analysis of a given frame. The most frequent methods used are the covariance method, the autocorrelation method and the lattice method [1]. In each method a set of \((F+1)\) LPC parameters and a gain parameter \(\sigma\) are extracted. In this paper, the autocorrelation method is used in which the autocorrelation coefficients of a given speech frame \(x = (x_0, \ldots, x_{N-1})\) are computed as

\[
\Gamma_{xx}(n) = \sum_{l=1}^{N-1} x_l x_{l+n} \quad 0 \leq n \leq N-1
\]

The quantization step is made using the VQ scheme which involves an encoder a decoder and a codebook. VQ codebooks are designed to minimize the average distortion that results from encoding a long training sequence of a particular frames. If \(x_j, j=1, \ldots, L\) is such a training sequence, the codebook \(C\) is designed such that [2]

\[
D = \frac{1}{L} \sum_{j=1}^{L} \min_{l} \|x_j - y_l\|
\]

where \(y_l\) are the codewords of the codebook and \(d(x_j, y_l)\) is the distortion measure which is the Itakura-Saito distortion measure and is given by [3]:

\[
d_{IS}(x, y/A) = \frac{\alpha}{\sigma^2} + \ln(\sigma^2)
\]
σ/A is the model filter

\[ A(z) = 1 + \sum_{k=1}^{P} a_k z^{-k} \]

σ is the residual energy term given by

\[ \sigma = r_a(0) r_x(0) + 2 \sum_{k=1}^{P} r_a(k) r_x(k) \]

Where

\[ r_x(k) = N^{-1} \sum_{i=0}^{N-K-1} x_i x_{i+k} \]

\[ r_a(k) = P^{-1} \sum_{i=0}^{P-K-1} a_i a_{i+k} \]

The method used for the construction of the codebook is the Shape Gain Vector Quantization (SGVQ) which is recently developed by Sabine et al [4]. The (SGVQ) is a quantizer whose codebook is the cartesian product of a codebook of vectors (Shape) and codebook of scalars (Gains). The final construction of the codebook is consistent with the idea of Durton [5] using multiseciton vector quantization.

The methods used for the test of the codebook are:

1. The distortion measure test which is the main objective test. The distortion measure used in both open test and closed test is the Itakura-Saito distortion measure.
2. The formant trajectories comparison test, in which the formant trajectories of the original frames are compared to those extracted from the codebook. The formant trajectories are obtained by using the LPC analysis [1].
3. The waveform comparison test. The constructed speech waveform using the LPC parameters from the codebook and the residual signal as excitation source are compared to the original speech signal.
4. Subjective test

The glottal waveform could be used as an excitation source for the codebook. The proposed method for the extraction of the glottal waveform is based on the method developed in [6] with some modifications to reduce the calculation time and the memory requirements and it can be summarized in the following procedures:

1) Each frame of speech is pre-filtered, sampled and pre-emphasized with a pre-emphasis factor \( \mu \). 2) A \( P \) order linear predictive filter is calculated using the autocorrelation method instead of the covariance methods to reduce the required calculation. 3) the residual signal is obtained by inverse filtering then the source signal is obtained by de-emphasizing and then it is integrated. 4) The glottal system is linearly modeled using the ITIF algorithm [7] and the normal equations are solved using the sequential regression method [8] instead of the least square method to reduce the memory requirements. 5) An impulse train separated by the pitch period is used as input to the FIR filter so obtained.
Another approach is to obtain the impulse train from the residual signal obtained by inverse filtering the integration of the source signal by the glottal filter. Step 4 is repeated using the modified least square autocorrelation compensation method [9] instead of the ITIF method and this yield to the same results but with reduction in the time needed for calculations.

**Experimental simulation**

The speech material selected for the construction of the codebook consists of 48 arabic sounds spoken by five different speakers, three males and two females. The speech material was composed by the following sets:
1. CV monosyllables where C : /d/, /t/, /s/, /, /, /x/, / and V : /a/, /i/, /u/.
2. CVC monosyllables the first C is the glottal stop /\. V and the second C as in the CV monosyllables .

The speech signals are prefiltered at cutoff frequency 5 KHz and sampled at a sampling frequency of 10 KHz. LPC parameters of order P=12 as well as the gain are extracted using the autocorrelation method with a pre-emphasis factor \( \mu = 0.9 \) and a hamming window for each frame of \( N = 200 \) samples i.e. 20 msec duration. Two codebooks are constructed and tested; a male codebook and a female codebook. The number of frames in the male codebook is 1500 frames and in the female codebook is 900 frames. The number of codewords in each codebook is 64 shape codewords and 10 gain codewords.

**Results and discussion**

1. From the results of the formant trajectories comparisons we can observe the following:
   a) The differences in these trajectories in the case of CV syllables appear in the ONSET of the of the vowel /u/ with the emphatics /s/, /d/ and of the vowel /i/ with /, /, /x/, / and the vowel /a/ with the pharyngeal /\. b) In the case of the CVC syllables the differences appear in the OFFSET of the vowel /a/ with /, /d/ and of the of the vowel /u/ with /s/, / and of the vowel /i/ with /, /x/.
2. The lowest distortion measure is 0.002 and the highest one is 0.3 These values are in the same range as those of [4]. The variations of the distortion have the same tendency as the variations of the formant trajectories.
3. The glottal waveforms obtained are shown in Fig 1.
4. The results of the waveform comparisons are shown in Fig 2.

**Conclusion**

1. The constructed codebooks are the first codebook for these arabic sounds.
2. The codebooks could be used in digital speech storage with remarkable reduction of the memory requirements since the stored data has a memory size equal 4.5% of the size required for the overall data.
3. The codebook could be used in speech coding with a reduction in the bit rate of about 33%.
4. The method used in the construction of the codebook dealt to the reduction of the search time.
5. The test of the codebook shows that it is a good matching between the original frames and those extracted from the codebook.
6. The proposed method for the extraction of the glottal waveform reduce the computational time and the memory requirements.

References
ASSP Vol. ASSP-31 N 6 pp 1556-1565 Dec. 83

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{glottal_waveforms.png}
\caption{The glottal waveforms}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{original_waveform.png}
\caption{a) Original waveform of /ka/; b) Waveform of /ka/ from the codebook.}
\end{figure}
4. B
MACHINE RECOGNITION OF SPEECH
USING TIED MIXTURE CONTINUOUS PARAMETER MODELS

Jerome R. Bellegarda and David Nahamoo
IBM Research, T.J. Watson Research Center,
Yorktown Heights, New York 10598, U.S.A.

ABSTRACT
Discrete and continuous parameter approaches to the acoustic-modelling problem in the
machine recognition of speech are unified through a class of very general hidden Markov models
which can accommodate sequences of information-bearing acoustic feature vectors lying either
in a discrete or in a continuous space. More generally, the new class allows one to represent the
prototypes in an assumption limited, yet convenient way, as tied mixtures of simple multivariate
densities. Speech recognition experiments, reported for a large (20,000-word) vocabulary office
 correspondence task, demonstrate some of the resulting benefits.

INTRODUCTION
Acoustic channel modelling is a crucial problem in an automatic speech recognition system
such as described in [1]. If $\tilde{A}$ denotes the acoustic sequence corresponding to the sequence of
words uttered, $W$, the problem is to find an appropriate hidden Markov model for the quantity
$\Pr(\tilde{A}|W)$, so as to represent the speech waveform in a parsimonious and meaningful fashion. We
have recently introduced a new class of such models [2], with enough flexibility to preserve the
information necessary for good recognition, and yet sufficiently simple to be computationally
 tractable. The purpose of this paper is to report on further experimental results obtained with
the new models on a 20,000-word vocabulary office correspondence task.

The new framework evolves from the unification of the discrete and continuous parameter
approaches to the acoustic-modelling problem in speech recognition. Suppose that the front-
end processor extracts from the speech waveform one vector of acoustic parameters per frame,
such as the energy in each of $D$ spectral bands. In the first approach, the resulting sequence
of vectors is vector quantized into a string of labels which is then assigned a (non-parametric)
multinomial probability distribution; as a result, a severe loss of information about the original
speech waveform may occur. In contrast, in the second approach, the sequence of acoustic pa-
rameter vectors is assigned a multivariate probability distribution directly, most often Gaussian
for cost effectiveness; since, however, a single Gaussian distribution can only model unimodal
behavior, this may lead to gross inaccuracies in the resulting model [3], [4]. In its most general
form, the new class of hidden Markov models allows continuous parameter modelling based on
tied mixtures of simple (unimodal) probability distributions. This compromise tends to reduce
the modelling inaccuracies arising from a single unimodal distribution while at the same time
retaining some of the flexibility of the discrete model.

GENERAL HIDDEN MARKOV MODELS
This section reviews the unified treatment of discrete and continuous hidden Markov mod-
delling presented in [2]. Assume that the acoustic evidence $\tilde{A}$ is the observed output of a hidden
Markov model, $M$. If we denote by $x_n$ the state of the underlying Markov chain at time $n$, $M$
admits the following set, $S$, of parameters: initial probabilities $\Pr(h_0) = \Pr(x_0 = i)$, transition
probabilities $\Pr(a_{ij}) = \Pr(x_i = j|x_{i-1} = i)$, and output probabilities $\Pr(y_n | x_n =
Note the double meaning of \( \tilde{y}_n \): in the discrete case, \( \tilde{y}_n \) refers to the label at time \( n \) (a scalar); otherwise, \( \tilde{y}_n \) is the acoustic parameter vector at time \( n \). In the latter case, the function \( \Pr\left( \cdot | a_n \right) \) is defined on a continuous space, and represents the output probability density function attached to the particular output-producing transition \( a_n \); typical examples include Gaussians, Laplacians, and \( K_0 \)-distributions [5].

Let \( \{ C_k \} \) be a generic set of different classes, regrouping various types of feature vectors and/or arcs. It is always possible to write:

\[
\Pr\left( \tilde{y}_n | a_n \right) = \sum_k \Pr\left( \tilde{y}_n | C_k, a_n \right) \Pr\left( C_k | a_n \right).
\] (1)

To simplify the term \( \Pr\left( \tilde{y}_n | C_k, a_n \right) \) and achieve a manageable set of parameters, we invoke a tying between these parameters, chosen in such a way that the term \( \Pr\left( \tilde{y}_n | C_k, a_n \right) \) for a given \( k \) is independent of \( a_n \). In other words, we assume that the output probabilities/distributions conditioned on the set of classes are independent of the arcs. Then, (1) can be reduced to:

\[
\Pr\left( \tilde{y}_n | a_n \right) = \sum_k \Pr\left( \tilde{y}_n | C_k \right) \Pr\left( C_k | a_n \right).
\] (2)

This equation is the key to our general approach to acoustic channel modelling. As shown below, usual types of modelling can be retrieved from (2) by appropriate choice of the quantities \( \Pr\left( \tilde{y}_n | C_k \right) \) and \( \Pr\left( C_k | a_n \right) \).

To recover the discrete model from (2), it suffices to choose for the set \( \{ C_k \} \) the space of all prototypes, i.e., the space where all the feature vectors are regrouped in a finite number of clusters, and to enforce the constraint:

\[
\Pr\left( \tilde{y}_n | C_k \right) = \begin{cases} U_k(\tilde{y}_n) & \text{if } \tilde{y}_n \in C_k; \\ 0 & \text{otherwise}. \end{cases}
\] (3)

In this expression, \( U_k(\cdot) \) is a uniform distribution over all acoustic feature vectors in the \( k \)-th cluster \( C_k \). The constraint (3) amounts to a tying which identifies with the vector quantization task, since each feature vector is replaced by its cluster identity. Of course, \( \Pr\left( C_k | a_n \right) \) then correspond to the (discrete) output probabilities.

On the other hand, to retrieve the continuous, single distribution model, it suffices to choose the set \( \{ C_k \} \) to span the space of all arcs, and to enforce the constraint:

\[
\Pr\left( C_k | a_n \right) = \begin{cases} 1 & \text{if } C_k \equiv a_n; \\ 0 & \text{otherwise}. \end{cases}
\] (4)

Now, the only terms contributing to the summation in (2) will be those for which the class is tied to the arc \( a_n \). In the Gaussian case, it remains to associate to this arc, through \( \Pr\left( \tilde{y}_n | C_k \right) \), a distribution of mean \( \mu_k \) and covariance matrix \( \Sigma_k \).

It is clear how to obtain a more general, continuous model from a mixture of such distributions. Let us choose for the set \( \{ C_k \} \) the space of all prototypes as in the discrete model, but assign to each cluster a Gaussian distribution as in the single continuous Gaussian case:

\[
\Pr\left( \tilde{y}_n | C_k \right) = \mathcal{N}(\tilde{y}_n | \mu_k, \Sigma_k).
\] (5)

Of course, the parameters \( \Pr\left( C_k | a_n \right) \) satisfy \( \sum_k \Pr\left( C_k | a_n \right) = 1 \), and will be referred to as mixture coefficients. The expression (5), together with (7), characterizes our general mixture model for the output distribution associated with a given output-producing transition. (Obviously, any other class of distributions could be used in (5) in lieu of the Gaussian family.) The parameter set \( \mathcal{S} \) for this model comprises all of \( \Pr\left( \tilde{y}_n \right) \), \( \Pr\left( a_n \right) \), \( \Pr\left( C_k | a_n \right) \), \( \mu_k \), and \( \Sigma_k \), which can be obtained via maximum likelihood estimation through the forward-backward algorithm.

The reader is referred to [6] for a complete derivation of the associated re-estimation formulas.

As shown in Fig. 1, this model may be rendered pictorially by replacing each arc \( a_n \) by a number of null (non-output-producing) transitions, depicted in dashed lines, followed by a
number of regular (output-producing) transitions, depicted in solid lines. At time \( n-1 \), the null transitions go from state \( x_{n-1} = i \) to some intermediate states, \( \{x_{n-1}^{(i-j)}\} \), the \( k \)th null transition being assigned a transition probability \( \Pr(C_k | a_i) \). In turn, the regular transitions go from each state in \( \{x_{n-1}^{(i-j)}\} \) to state \( x_n = j \), the \( k \)th one having the output distribution \( \Pr(y_n | C_k) \) attached to it. Because of the tying introduced earlier, the set of classes is the same for all the output distributions, regardless of the current arc; thus, the values that can be taken on by the intermediate states \( \{x_{n-1}^{(i-j)}\} \) depend only on the classes \( C_k \).

LARGE VOCABULARY EXPERIMENTS

The speech of ten speakers was recorded with a Crown PZM 8S desktop microphone and digitized at 20,000 Hz with a 12-bit A/D converter. Training and test data, each consisting of 100 sentences drawn from typical office correspondence, were read as isolated word speech. Acoustic feature vectors of dimension \( D = 20 \) were extracted from the signal every 10 milliseconds using a noise resistant front-end incorporating the auditory model proposed by Cohen [7]. The total number of test words was approximately 1700.

Three types of experiments were performed. (A) Discrete parameter experiments as benchmarks, with each feature vector quantized into one of 200 codebook entries; conventional K-means clustering and prototype labeling (e.g., see [1]) was performed on both training and test corpora. (B) Continuous parameter experiments with diagonal Gaussian distributions after quasi-rotation of the acoustic vectors in the direction of principal discriminants [3]; this is to make the off-diagonal terms in the covariances matrices as small as possible. (C) Continuous parameter experiments with pseudo-full covariance Gaussian distributions: to allow for a better directivity of each processed feature vector, quasi-rotation is performed on a prototype-dependent basis, which differs from full covariance processing only inasmuch as the rotation parameters are not adjusted during training; we refer to (C) as "multi-rotation" processing. In (B) the principal discriminants, and in (C) the prototype-dependent rotation parameters, were extracted from all the training corpus; the number of distributions considered was 200.

The decoding results are shown in Table I, which reports the word error rate (in percent) for each speaker, as well as the average word error rate across the ten speakers considered, and the corresponding average relative normalized decoding time. The latter quantity can be interpreted as the average amount of CPU time required to recognize one sentence in the test corpus, relative to the average amount required in the discrete parameter experiment. For every single speaker, continuous parameter modeling proves superior to discrete parameter processing. Sometimes,
Table I. Performance Statistics for 20,000 Word Vocabulary Isolated Utterance Experiments. “Discrete” Refers to Experiment (A), “Uni-Rotation” to Experiment (B), and “Multi-Rotation” to Experiment (C).

A wide margin separates the two approaches: for FMM and TFK, for instance, the error rate drops to about 63% of its original value. The average error rate drops from over 6% to under 4%, for a total improvement of roughly 25%. On the other hand, (uni-rotation) continuous parameter modelling is almost 1.5 times more intensive than discrete parameter modelling; further comments on this topic are offered in [6]. Multi-rotation processing performs quite well in terms of average error rate, which is somewhat surprising given the very small amount of training data available. This suggests that full covariance processing has the potential to give reasonable results even with such limited training.

This sheds an encouraging light on continuous parameter processing in general, and tied mixtures in particular. After speed ups are implemented to reduce the average decoding time, the next step will be to use such tied mixture continuous parameter models for (large vocabulary) connected speech recognition tasks.

ACKNOWLEDGEMENTS

We would like to thank the other members of the IBM Speech Recognition Group, and more particularly L.R. Bahl, P.F. Brown, R.L. Mercer, and M.A. Fischyn, for their help and insight shared during many profitable discussions.

REFERENCES

DEEP UNDERSTANDING OF SPEECH

Osamu TAKIZAWA and Masuzo YANAGIDA

Communications Research Laboratory
Ministry of Posts and Telecommunications
4-2-1 Nukui-kitamachi, Koganei, Tokyo, 184 Japan

INTRODUCTION

An approach to deep understanding of speech is proposed. Deep understanding here means detection and analysis of hidden meanings like puns and metaphors. The idea is based on confident phoneme identification and intentional search of semantic chaining on a frame system. This paper focuses mainly on understanding puns that are played upon words composed of the same number of syllables of different or similar phonemes.

The system is divided into the signal processing subsystem and the language processing subsystem. Conceptual design for the latter is almost completed, and the former is under development.

DEFINITION AND CLASSIFICATIONS OF PUNS

Definition of puns and their component words

A pun is defined as sharing syllable positions on a phoneme sequence among words containing the same or similar phonemes for the common syllable positions resulting in additional or hidden meanings besides the ordinary interpretation. The interpretation comes by assuming a single or a set of appropriate words that can function in the original context with semantic conformity. These words are called the "words-in-context". The additional meanings, which aren't necessarily related to the current context, comes from words with possibility of skewed, distorted, perverse interpretations. The words from which the additional meaning comes are called the "material words".

Classifications of puns

Puns can be classified in several ways. The following are some of them.

[Classification I]
Classification by the type of the discrepancy in pronunciation between the words-in-context and the material words.

- case-1 No difference.
- case-2 Phonemic difference.
- case-3 Prosodic difference.

[Classification II]
Classification by the semantic relation between the words-in-context and the material words.

- case-1 where these two sets of words are related directly to each other semantically.
- case-2 where the set of material words forms a compound word or a phrase and each element word is semantically related to each other, but not necessarily to the context.
- case-3 where these two words are related indirectly to each other semantically. "Indirectly" means "through common knowledge shared between the speaker and the listener".

403
[Classification III]
Classification by the existence of the material word(s) appearing in other places in advance.
case-1 where it exists.
case-2 where it doesn't exist.

[Classification IV]
Classification by degree of similarity in pronunciation between the words-in-context and the material words.
case-1 Similar. (Hard to distinguish each other.)
case-2 Not similar.
The puns classified as case-2 are easier to be detected than those of case-1.

PROCESSING PUNS

The Outline of Processing Puns
A machine understanding of puns here is regarded as detection of both the words-in-context and the material words. So the detectable puns are those that present discriminable differences in pronunciation between the words-in-context and the material words. This paper focuses on detecting puns of case-2 and case-3 in Classification I, all cases in classification II and III, and case-22 in Classification IV.

These puns are detected by the following procedure.
[step1] List up candidates for the words-in-context, and decide the words-in-context.
[step2] Judge to trigger the pun-procedure or not.
[step3] Select candidates for the material words.
[step4] Point out the semantic implications.

Details of the Detection Steps
[step1] This step is basically same as the conventional speech recognition systems. In order to prepare for later discussion, a "probability score" P for a candidate word is defined as follows.

\[ P = C \times \exp(n) \]

C : The product of confidence scores of the input phoneme sequence.
n : The supposed number of phonemes including silence assumed as a phoneme.

The "confidence score" assigned to a candidate phoneme is defined as the normalized reliability ratio of phoneme identification. An example is shown in Fig.1. The top line is denoted with the notations of the candidate phonemes of the highest confidence score.

\[
\begin{array}{ccc}
0 & b & 1 \\
(t 4)(r 5)(A 8)(b 4)(l 4) \\
(p 3)(J 3)(a 2)(w 2)(r 3) \\
(k 3)(w 2)(a 2)(w 2)(w 3) \\
(d 2) \\
\end{array}
\]

(unit : x 0.1)

Fig.1 An example of discrimination scores for /trab/ part.

The system makes a list of words sorted in descending order of C, where those words in conform with the context are regarded as the candidates for the words-in-context Sc.
If the number of phonemes of candidate words is larger or smaller than that presumed from analysis of input speech, the exponent in the formula above is reduced to be (the presumed number of phonemes) − (the difference). If a candidate word contain phonemes not out of the phoneme candidate list for the corresponding segment, the discrimination score for that phoneme is assumed to be 0.1 in order to avoid null product.

[step2] If Sc is found, the rank, in phoneme candidate list, of each phoneme in the phoneme sequence of a candidate word is checked whether it is the phoneme giving the highest confidence score in phoneme discrimination. If some of the supposed phonemes are not of the highest confidence score, the system tries to find other sequence Sp of words employing other phonemes of higher confidence score supposing case-2 in Classification I. If the detected prosody differs from that required by the context, the system looks for alternative word assignments supposing case-3 in Classification I.

[step3] The system investigates if Sp satisfies the conditions as the material words for puns. If Sp and Sc are related to each other with one of the relations described below, Sp is regarded as the material words, and the input phoneme sequence is judged to form a pun.

1. having common slot values.
2. having common super classes within five up-going links either in frame names or slot values.
3. having common values in association slots.

These steps are repeated forming loops on candidates for Sp's and Sc's. If no Sp nor Sc is found, the system concludes that the input phoneme sequence does not make any pun.

[step4] Having found appropriate Sc and Sp, the system displays them together with the key associative word detected in step 3 as the common value.

A Performance Example

In this section a performance example on a simple pun is given.

Person A: "Where did you make your flight reservation?"
Person B: "Japan Trouble Bureau."

[step1] The signal processing subsystem is assumed to produce output phoneme sequence for person B as /ʤænən træbl bjuː rau/.

```
gæn
Japan trouble neuron
Japan__Travel__Bureau
旅行 Bureau
Jack and Ravel
treble
```

Fig. 2 Possible candidate words for the input phoneme sequence.

"travel" /trævl/  n = 5
P = 0.4X0.5X0.2X0.4Xexp(5) = 0.47

"trouble" /træbl/  n = 5
P = 0.4X0.5X0.6X0.4Xexp(5) = 2.85

"Jack and Ravel" /ʤæk and rævl/
 n = 3-1+3-1+4 = 8 (two '1' mean 2 unmatched silences)
P = (0.5X0.8X0.2X0.3X0.1X0.5X0.2X0.2X0.4Xexp(8) = 0.02

"Japan Travel" /ʤænən trævl/
 n = 5+1+5 = 11 ('1' means 1 matched silence)
P = (0.5X0.7X0.4X0.8X0.6X0.5X0.2X0.2X0.4Xexp(11) = 12.88

Fig. 3 Examples of calculating the probability score P.
Among the possible candidate words depicted in Fig.2 "Japan Travel Bureau" attains the highest probability score P, because its number of phonemes in total is much larger than other combinations of words and it works significantly on the exponential part. So "Japan Travel Bureau" is assigned to be the preliminary and primary Sc.

As Sc isn't composed of phonemes of the highest discrimination scores, the system tries to find other sequence Sp consists of phonemes of higher confidence scores.

Then system finds the word "trouble" as the preliminary Sp. It is a common agreement that Japan Travel Bureau often makes mistakes, so a value "MISTAKABLE" is registered in the character slot (Fig.4). "CHARACTER SLOT" in the association slot means "slot values of character slot are regarded as values in association slot in that frame". Then "MISTAKE", from "easy to mistake" derived from morphological analysis of "MISTAKABLE", is found in "TROUBLE" file (Fig.5), and this filename matches with Sp. As a result the system recognize that "Japan Travel Bureau" is related to "trouble".

Finally, the system outputs statement saying that Sc is "Japan Travel Bureau", Sp is "trouble", "Japan Travel Bureau is mistakeable." and "Japan Travel Bureau and trouble are related via "mistake"."

FRAME NAME : JAPAN TRAVEL BUREAU
SUPER CLASS : TRAVEL AGENCY
PLACES OF OFFICE : TOKYO, OSAKA, KYOTO,...
ASSOCIATION SLOT : THIP, TICKET, "CHARACTER SLOT",...
CHARACTER SLOT : MISTAKABLE, KIND,...

Fig.4 A frame "Japan Travel Bureau".

TRAFFIC : PLANE, TRAIN, CAR, SHIP,...
TICKET : TRAIN, PLANE, CONCERT, MOVIE,...
TROUBLE : QUARRELING, APOLOGIZE, ERROR, MISTAKE,...

Fig.5 Samples of conceptual association files.

DISCUSSIONS

The proposed algorithm has problems as follows:
a. Efficient search algorithms are to be developed.
b. Measures to overcome missing and insertion errors in input phoneme sequences are to be designed.
c. How to decide the best pair for Sc and Sp, in case they are not decided uniquely in step3.
d. The proposed algorithm depends on highly reliable phoneme identification of input phonemes.

CONCLUSIONS

Proposed is a natural language processing scheme for detecting hidden meanings in speech, particularly for puns. Although the research in hidden meanings sounds "too much futuristic", some means to detect hidden meanings are required in order to develop a speech recognition system for conversational speech. The proposed approach may suggest a direction to cope with tough problems involved in speech recognition.

References


406
MULTISPEAKER VOWEL-CLUSTER COMPRESSION BY FORMANT FREQUENCIES NORMALIZATION

Ph. Christov
Scientific Workers Union of Bulgaria
Regular member
Personal P. O. Box 895, 1000 Sofia, Bulgaria

Introduction
In trivial spectral representation of speech

\[ t \mapsto f(t) = \sum_{n=1}^{\infty} C_n \sin((2\pi F_n) t + \phi_n), \quad t_0 \leq t \leq t_0 + T \ldots /1/ \]

the least common multiple of all harmonic frequencies \( F_n = n F_0 \)
is the pitch frequency \( F_0 \).

As far as each formant frequency \( F \) coincides with the frequency \( F_m \) of some of the vowel harmonics

\[ F = F_m, \quad F_m \in \{ F_1, F_2, F_3, \ldots F_n \} \]

it appears that the position of any vowel in the \( F_1 \) vs. \( F_2 \)space will be more or less influenced from the value of the pitch frequency \( F_0 \) in that particular vowel uttering,

Normalization & Compression
The removing of the influence of the pitch frequency canbe achieved by normalization of the harmonic and/or the formant frequencies of the speech signal by its pitch frequency \( F_0 \) (CHRISTOV 88)

\[ ||F||_{F_0} = \frac{F_m}{F_0} = \frac{m F_0}{F_0} = m \ldots /2/ \]
Expression (2) shows that the value of the formant frequency \( F \), from a three or four-digit number, reduces to the one or two-digit number \( m \) of the \( m \)-th harmonic from the Fourier Transform \( F \) of the speech signal. This means that transformation (2) should lead to a considerable compression of the \( F_1 \) vs. \( F_2 \) space along with the vowel clusters in it contained.

Example: The implication from expression (2) that formants pitch-normalization gives rise to cluster compression, is demonstrated on a vowel set extracted from the Bulgarian Central Allophones Database (CHRISTOV 87). The set consists of stressed vowels /a/ pronounced in /b-b/ context by 30 female professional speakers. The vowels are imbeded in words and uttered at the end of a carrier sentence. The 30 vowel utterances have been verified by a group of 20 listeners, which rejected two of the utterances as false. The remaining \( N=28 \) utterances have been digitalized by a rate of 20 kHz and lead to the input of an IBM 350 computer. There they have been processed by a computer program (CHRISTOV 87) which performs pitch synchronous spectral analysis followed by formant tracking. The computer output in the \( F_1 \) vs. \( F_2 \) space is shown in Fig. 1, where each ordinary point (empty circle) marks one hit, originating from a single vowel uttering. The total \( L=21 \) of points in the \( F_1 \) vs. \( F_2 \) space is smaller than \( N \), because of the coincidence of two hits in each of the seven first points. The coincidence of hits in the space of the normalized first two formants, as shown in Fig. 2, is far much higher; 82% of the hits are concentrated in only 3 points of Fig. 2, the point (3,6) alone carrying 43% of all 28 hits. Correspondingly, the total \( L \) of normalized points takes the value \( L=7 \), which is much smaller than \( L \). The observed concentration of hits and reduction the total of points after the normalization may be considered as an after-effect of the compression of the entire \( F_1 \) vs. \( F_2 \) plane containing the vowel cluster; The \((F_{max} - F_{min})\)-rectangle, \((337x639)\), in Fig. 1, in which the original vowel cluster is inscribed, takes the immensely smaller dimensions of only \((2x5)\) in the plane of the normalized formant frequencies (Fig. 2).

Discussion

The impact of removing the influence of pitch on the vowel representation is so obvious that it hardly needs any discussion:

1) Eliminating the influence of pitch (including the intrinsic pitch) on the position of the vowels in the \( F_1 \) vs. \( F_2 \) space makes the \( F_1 \) vs. \( F_2 \) representation more precise.

2) It makes possible the direct juxtaposition of male and female voices and/or voices of different dialect or ethnic groups for comparative studies of speech.
Keeping this in mind, we will rather concentrate on the side effect of normalization - the cluster compression. Apart from the theoretical considerations, the visual inspection of Fig. 1 and 2 alone inspires the conviction that formants normalization really leads to a tight cluster compression; not only the number of points in the space of the normalized formant frequencies (Fig. 2) is much smaller, but the concentration of hits in them is much higher than before the normalization (See Example).

The practical dimensions of this side effect of normalization should have highly positive consequences on the computational procedures applied to the speech signals:

1) The number of computations should be reduced because of the much smaller number, $|L_0|$, of computational objects (points in the measurement space)

2) Large amounts of working space in the computer memory (main and peripheral) should be saved during computations because of the measurement space and cluster compression.

Conclusion

The pitch normalization of the two-formant representation of vowels may prove to be an useful tool applicable to both comparative phonetic studies and machine recognition of fluent speech independently from the speaker.

References


CHRISTOV, Ph., A Large Bulgarian Central Allophones Data Base, Proc. 11-th International Congress on Phonetic Sciences, Tallinn, August 1-7, 1987; Vol. 5, pp. 232-235

Fig. 1. Isolated cluster in the P1 vs. P2 space of the Bulgarian vowel /ø/ uttered in stressed position in /b-b/ context by 30 female speakers
Legend: o - single uttering
       - two coincident utterings from two different speakers

Fig. 2. The vowel cluster of Fig. 1, after pitch normalization of the formant frequencies
Legend:
o - single uttering
i - i coincident utterings from i different speakers
i ∈ {2,5,6,12}
MACHINE SPEECH RECOGNITION BY BSDP, A KNOWLEDGE BASED
- TEMPLATE MATCHING ALGORITHM

Miroslav L Kiš, Miroslav R Matašek
VTI RV I PVO, Niška bb ETT Beograd, B.Revolucije 73
11132 Žarkovo Yugoslavia 11000 Beograd Yugoslavia

Introduction

Numerous R&D projects in the field of machine speech recognition have shown that results primarily depend on the amount of speech knowledge incorporated and utilized by the algorithm. By taking advantage of visual analysis in the framework of BSDP (Binarized Spectrograms and Dynamic Programming) algorithm, it is possible to improve usual methods (e.g. clustering) for selecting reference templates and the algorithm as a whole. Even more, artificially constructed reference templates make it possible to analyze and determine the features that provide good recognition of similar words.

Algorithm Description

The basic steps of the machine speech recognition (MSR) system based on binarized spectrograms and dynamic programming are [1,2]: a) preprocessing of speech signal (preemphasis and short time power spectrum - spectrogram determination with averaging over the frequency bands which define 25 channels); b) spectrogram binarization (determination of frequency - FRN and channel - CHN normalization). Algorithms for FRN and CHN determination are as follows. Mean value for each frame is computed and compared with value of each point power in the same frame. If a point power value is greater then the average, FRN value in this point is one, or zero, if it is not. The same procedure is used for CHN but it is performed for each channel and the mean value is multiplied by a factor 0.3 before the comparison. The next steps are: c) word template forming (binarized spectrogram which is the combination of channels 1-10 and 21-25 from the FRN and channels 1-5 from the CHN); d) determination of similarity between unknown and reference templates (by using Dynamic Programming Algorithm - DTW with Itakura's local and global constraints); and finally, e) decision making (on the basis of global distance analysis i.e. analysis of distances Dtr between unknown - test, and all reference templates). If the reference templates are formed from spoken words - other possibilities for the reference template construction will be discussed later - steps from a) to c) have to be done for each reference word.

The results of the testing both BSDP and LPC-DTW algorithms on a vocabulary consisting of ten Serbo-Croatian digits. are shown in Table 1. Each digit has been spoken by 60 male and 60 female speakers, without repetition. The training group consisted of 30 male and 30 female speakers while the others belonged to the test group. Having in mind simplicity of the procedure just described, it is impressive that BSDP has shown advantages over one of the best existing template matching algorithms.

DTW Constraints

The problems, arising from the application of DTW algorithm for noni-
near time alignment, are well known. Highly restrictive local and global constraints do not allow good matching even for templates of the same word. On the other hand, if the constraints are not so restrictive, good matching of templates of different words is possible and unexpectable errors may occur.

Simultaneous analysis of both optimal path and BS has led us to an important result [3]. Visual analysis of the CHH has shown strong stability of some features that correspond to the vocalized segments of words. This part of CHH (channels 1-5) introduces additional constraints into DTW algorithm by defining "forbidden zones" marked in Fig.1 a-d. Optimal paths in Fig.1 a,b correspond to the case when CHH is not used. When the templates of the same word are aligned (digit "3"-Fig.1 a,c), optimal path is not significantly affected by CHH introduction. On the contrary, when templates do not originate from the same word, "forbidden zones" prevent good matching, and much greater distance is obtained (digits "3", "4"-Fig.1 b,d).

Reference Template Selection

Good reference templates are the crucial part of a template matching recognition system, particularly when the speaker independence is required. BSDF approach makes it possible to select reference templates by visual analysis of the features important for word discrimination [4]. But, this is not a convenient method in the case of a large vocabulary. In [5] a procedure, for combining results of MSR visualization experiments i.e. feature analysis method described in [4] with results of clustering technique, has been proposed. In Fig.2 two clusters of word "1" are shown. Templates in cluster a) all have similar structure - like letter "X". Templates in cluster b) have slightly different structure: left-upper part of "X" is missing. Cluster centres are templates which have been spoken by speakers M095 and M051, respectively. When these cluster centres were used as references, a lot of "7"->"1" confusions occurred. By using clustering technique, we obtain some characteristic template structures but they are not necessarily the ones that provide good recognition. In Fig.3 the templates obtained by subsequent visual analysis are shown. It can be seen that the feature, significant for good recognition, is the "black island" in the first 5 channels at the beginning of word "1", what was confirmed by experiments.

It follows that clustering is a useful technique for the first, broad feature analysis. But, for good recognition, subsequent visual analysis is needed.

Feature Analysis by Using Artificial Reference Templates

Determination of speaker independent features and a method for its reliable extraction are two basic, and problematic steps in the development of a "feature extraction" MMH system. Visual analysis of spectrograms, performed by experienced "spectrogram readers", is a valuable source of such knowledge, but it requires hard, long time work. In the framework of BSDF algorithm a very efficient feature determination procedure can be carried out.

In Fig.4 some spectrograms of confusable words (digits "5", "6", and "9") are shown. Here, it should be remarked that confusions of these types occurred when both LPC-DTW and BSDF algorithms were applied. Visual analysis of binarized spectrograms has shown that, even with such information
reduction, reliable determination of the main features is very difficult.

In this regard, especially interesting and promising result has been obtained with reference templates that do not originate from spoken words i.e. templates that have been formed by manual design. These artificial
reference templates, shown in Fig.5, contain only those features that were supposed to be distinctive. By using these three reference templates for
recognition of words "5", "6" and "9", only 3.6% error rate has been obtained. Errors have occurred only on those word templates that have not had the features defined by the artificial references.

Proceeding from this result, we can define a procedure for feature
determination as follows: a) visual analysis of confusable word spectrograms;
b) initial guess of distinctive features; c) verification of the assumption
through recognition experiment by using artificial reference templates.

Conclusion

Despite extensive research work and some encouraging results the
problem of MSR is still far from the solution. It has become quite clear
that the amount of speech-specific knowledge, effectively used by MSR
algorithms, has the major impact on their performances. We are faced with
two essential problems: how to incorporate and efficiently utilize speech
knowledge in a MSR algorithm, and how to deal with variabilities that could
not be exactly predicted, since this knowledge is limited.

BSDF algorithm offers a rather natural solution. Being a template mat-
ching algorithm, BSDF keeps good characteristics of such an approach i.e. a
very effective mechanism for ignorance modeling. In addition, visual analy-
sis, of both binarized spectrograms and the entire recognition process, makes it possible to: understand problems that arise, extract relevant features and efficiently select the knowledge-based reference templates. Thus, speech-specific knowledge can effectively be used for the algorithm perform-
ances improvement. Such an approach results in a high performance knowledge-based template matching algorithm that can be easily implemented [6].

References

[1] M.R.Matanić, M.L.Kiš, "Speech Recognition by Dynamic Programming and
One-Bit Coded Spectrograms" (in Serbo-Croatian), in Proc. XXVIII ETAN

Recognition Experiment and BSDF - Binarized Spectrograms Dynamic
Programming Algorithm" submitted for publication in IEEE ASSP.

for Isolated Word Recognition" (in Serbo-Croatian), in Proc IX BIH
Symposium in Informatics, Sarajevo, Yugoslavia, March 1985.

[4] M.R.Matanić "Visualization of the ASR Experiment-a New Methodology
for Developing Speaker-Independent Automatic Speech Recognition
Systems" (in Serbo-Croatian) in Proc. XXXI ETAN Conf., Vol.6, Bled,

Recognition" (in Serbo-Croatian), in Proc. XXXI ETAN Conf., Vol.6,
Bled, Yugoslavia, June 1987.

[6] D.Mitić, V.Vignjević, M.Jovanović, M.Kiš "A Laboratory System for Ma-
chine Speech Recognition Experiments Based on BSDF Algorithm" in Proc.
**13th INTERNATIONAL CONGRESS ON ACOUSTICS**
**YUGOSLAVIA * 1989**

**Fig. 1.** Optimal paths: a) and c), test word (T) "3", reference word (R) "3"; b) and d) T "3", R "4".
- $D_{3,3} > D_{3,4}$ for a) - b)
- $D_{3,3} < D_{3,4}$ for c) - d)

**Fig. 2.** Two clusters: centers (+) and some elements, of word "1".

**Fig. 3.** Some references for word "1"

**Fig. 4.** FRN's of some words "5", "6" and "9", left to right.

**Fig. 5.** Artificial references for words "5", "6" and "9", left to right.

<table>
<thead>
<tr>
<th>Training group</th>
<th>Test group</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSDP</td>
<td>0.30 %</td>
</tr>
<tr>
<td>LPC-DTW</td>
<td>1.09 %</td>
</tr>
</tbody>
</table>
A LABORATORY SYSTEM FOR MACHINE SPEECH RECOGNITION EXPERIMENTS
BASED ON BSDP ALGORITHM

D.S. Mitić, V.Ž. Vignjević, M.R. Jovanović, M.L. Klš

VTI KoV, Katanićeva 15, 11000 Beograd, Yugoslavia
VTI RV I PVO Niška bb, 11132 Žarkovo, Yugoslavia

Introduction

Speech recognition experiments on mainframe computers are preceded by data acquisition on a large set of speakers. These experiments include long-lasting checking of noise and different speaker and pronunciation influences. But sometimes, the time gap, between data acquisition and recognition, makes experiments control and results explanation very difficult. On the other hand, commercially available real-time devices for speech recognition are application ruled, prohibiting any serious hardware or software upgrading or changing. Therefore, they are not suitable for the research. Bearing all these facts in mind, we have realized a laboratory system for real-time machine speech recognition experiments (LS-MSR).

The main purpose of the design of the system was to support further development of BSDP (Binarized Spectrograms and Dynamic Programming) algorithm [1-6]. Being based on that algorithm, the system is very suitable for efficient experimental work. During the software development, special attention has been paid to those elements that enable efficient visual inspection and correction of spectrograms, feature-analysis and construction of artificial reference templates.

Besides a description of the system configuration, its main characteristics are outlined and applicability of such system in the MSR research is considered.

LS-MSR Description

LS-MSR configuration (Fig.1) consists of: BSDP based device for MSR of isolated words (MASP), tape recorder, personal computer and a printer. It can be used in several ways: a) MASP alone, can be used for real-time speech recognition experiments or as a computer voice input device; b) MASP – PC configuration allows block of data (e.g. binarized spectrograms) transfer, PC data post processing and PC control by speech via MASP; c) MASP – tape recorder – PC hardware/software connection offers synchronized recognition of words recorded on a tape and thus analyses of experimental data can be done automatically.

MASP Hardware

MASP is a multiprocessor system designed for real-time speech processing. MASP hardware (Fig.2) [8] consists of: dual MC68000 microcomputer system (master MP, slave SP processor, shared memory SHM and control logic), input speech processing unit, speech synthesis unit, interface II to PC, interface I2 to MASP terminal and interface for the tape recorder synchronization. The speech processing unit consists of: preamplifier PA, amplifier A, band-pass filter BPF and A/D converter. The synthesis of speech is based on aliophone method (GEC SP0256AL2).
During the processing, speech signal is fed to the input speech processing unit either from the microphone or tape recorder. Speech samples are then passed to the microcomputer unit (two processors with partly shared tasks). MP supervises and controls all actions in the system. Also, it takes tasks of: communication with both the user and PC; A/D conversion; DTW performing on one half of the reference templates and final decision making. SP is responsible for binarized spectrograms forming and DTW performing - on the rest of the templates.

**MASP Software**

Flow chart of the basic MASP software routines is shown in Fig.3. In order to provide real-time operation, all software routines have been written in MC6800 assembly language. But, an ordinary user is guided by menu-selectable options which makes internal structure of the software invisible for him. After the system initialization, a MENU with modes of operation appears on both PC and MASP terminals. When user selects one of the offered modes, he is asked to perform all further actions (if there is any), necessary to complete the operation. On the other hand, an experienced user can take advantage of powerful interactive debugger to perform more sophisticated analyses.

In addition to the standard modes of operation i.e. training and recognition, MASP offers some modes characteristic of the BSDP recognition algorithm application [6]. It is possible to view and edit binarized spectrograms of reference words. By changing binary zeroes to ones and vice versa, the features, significant for good recognition, can be enhanced. Also, user can delete some references or send them to the PC for further processing and storing. These references can be retrieved and used later on, during some other experiments.

Efficient software solutions, connected with the visual analysis capabilities of the recognition algorithm [6], allow to perform some very useful experiments. For example, in Fig.4-a, binarized spectrogram of a Serbo-Croatian word (digit "1", pronounced as /yedAn/) is shown. This word was not recognized when speaker independent references were used. The reason was very weak pronunciation of the sound "d". Visual analysis of the correctly pronounced and recognized word Fig.4-b, has shown significant difference -"white whole" in the middle of CHN [6]. Thus, by taking advantage of such an analysis, problems speakers can be detected and classified and their training is simple and under the self-control.

**PC Software**

The interactive mode of operation is based on menu-selectable options. It is possible to communicate between MASP and PC, edit, display or hard copy reference word spectrograms and perform recognition results analyses. It is easy to correct or even create entirely artificial spectrograms - a very useful option for feature-analysis [6].

**Summary**

According to original and efficient algorithm [1-6] we have been able to realize LS-MSR, a laboratory system for machine speech recognition experiments - an exceptional device concerning its capabilities and possible applications. Its hardware is made up of commercially available...
components, resulting in low price and easy device multiplication.

We should outline some important characteristics of LS-MSR: real time experiments capability, MENU-driven user friendly software (no need of machine hardware, software or language knowledge); huge amount of data (a lot of speakers) experiments, owing to MASP - tape recorder connection; spectrogram visualization and editing; and reference template selection for high recognition accuracy.

Although, the experiments on mainframe computers can not be avoided during a speech recognition algorithm development, some specialized devices for real-time experiments are extremely useful. LS-MSR is highly recommended both as a tool for research and experimenting in machine speech recognition and as an "intelligent" peripheral device for existing and new generation computers - a step towards an easier man-machine communication.

References


417
Fig. 1. The Laboratory System for Machine Speech Recognition

Fig. 2. MASP hardware schema

Fig. 3. The microcomputer software flow chart

Fig. 4. Binarized spectrograms of word "1" of the same speaker, obtained on LS - MSR.
RECOGNITION OF THE PHONEMS BASED ON THE NETWORK MODEL

V. J. Chichupal

Academy of Sciences of the USSR, Computer Center,
117967, Moscow, Vavilova 40, USSR

Abstract
One of the well-known rules of phonetic labelling and segmentation of
continuous speech is based on the knowledge of speech signal parameters
distribution. In this paper, we describe an algorithm, that uses a
multilayered network in order to represent the information about
distribution of speech parameters. Every layer of the network is a
so-called "self-organized feature map" introduced by Kohonen /1/. It is
shown that such multilayered network is capable of effective represen-
tation (or coding) of speech signal parameters. The structure of the
network is designed in such a way, that the deeper a layer the more
complicated features are coded by every unit of the layer. The
integration of features is achieved in the time domain, thus at the
first input layer we deal with rather local features - the vectors of
short-time parameters, but at deeper layers each unit responds to the
time trajectory of the more local features detected at the previous
layer. The units of the last layer respond to a certain phonetic
category (vowel or syllable).

Network Model for Speech Analysis
Let the vectors \( \mathbf{x}(t) = (x_i(t)), i=1,2,...,N \) be vectors of the
values of the speech parameters in discrete time moments \( t=0,1,2,\ldots \) (feature
vectors). The network consists of several (in our experiments - 5)
layers of the units. Each layer is represented by a square matrix \( u_{ij} \)
or \( u(\mathbf{i}, \mathbf{j}) \), where \( i \) is the coordinate vector of a unit: \( \mathbf{i} = (i,j) \).
All units belonging to the same, for example, 1-th layer, get at a given
time the same input - the feature vector \( \mathbf{x}(t) \). Each unit \( u(\mathbf{i}) \) has its
own set of connection weights: \( \mathbf{w}(\mathbf{i}, \mathbf{j}) = (w_{ij}(\mathbf{i}, \mathbf{j})), i=1,...,N \). These
weights determine response of the unit on the input vector. Moreover,
all units have a particular set of weights (identification weight vector
\( \mathbf{w}(\mathbf{i}, \mathbf{j}) \)) in order to identify the type of the input feature vector. The
number of these weights is equal to the number of different phonetic
categories in speech signal.

The network formation (that is the determination of the values of the
weights \( \mathbf{w}(\mathbf{i}, \mathbf{j}) \)), \( \mathbf{x}(t) \)) is carried out during two consecutive stages of
learning. At the first stage, we define the values of the \( \mathbf{w}(\mathbf{i}, \mathbf{j}) \).
Initially these values are a small random numbers. When a feature vector
\( \mathbf{x}(t) \) reaches at moment \( t(t>0) \) the \( i \)-th layer:
1. \( \mathbf{w}(0,1) \) is defined as:
\[
\mathbf{w}(0,1) = \arg\min_{\mathbf{w}} \| \mathbf{x}(t) - \mathbf{w} \| \quad (1)
\]
2. For all \( u_{ij}(\mathbf{i}, \mathbf{j}) \), such that \( \mu(\mathbf{i}, \mathbf{i}^*), \mathbf{w}(0,1) \leq \rho \):
\[
\mathbf{w}(\mathbf{i}, \mathbf{j}) = \mathbf{w}(\mathbf{i}, \mathbf{j}) + a(t)(\mathbf{x}(t) - \mathbf{w}(\mathbf{i}, \mathbf{j}))
\quad (2)
\]
where \( a(t) = \sigma(a(t)) \leq 1, a(t) \downarrow 0 \),
and \( \rho \) - the metric in the \( R^N \).
As is known, the algorithm (1), (2) have the same features as the vector
quantization procedures.
3. Input vector \( \mathbf{x}(t) \) is fed to the next, \( i+1 \)-th layer:
\[
\mathbf{w}(i+1) = (w_{ij}(t-2), w_{ij}(t-1), w_{ij}(t))
\quad (3)
\]
During the second stage of the learning process the weights $\tilde{w}(n,t)$ are modified. The values of $\tilde{w}(n,t)$ are kept constant. Initially, all the $\tilde{w}(n,0)=0$. To find out the values of the $\tilde{w}(n,t)$ a test speech signal segmented into phoneme intervals by an expert was used. The correspondence $f_i(x(t)) \rightarrow w_i$, where $W_i$ - a phonetic category to be identified, is based on the results of the segmentation. The test signal is processed by feature extractor and the resulting feature vectors are used as input vectors to the networks described above. Let $x_i(t)$ be the input vector to i-th layer of the network at the moment t. In this case:

1. $\tilde{w}(n_0,t)$ is defined with according to (1).

2. For $u_i(n_0)$:

$$\tilde{w}_i(n_0,t+1) = \begin{cases} 
\tilde{w}_i(n_0,t+1), & \text{iff } i=f_i(x_i(t)) \\
\tilde{w}_i(n_0,t) & \text{otherwise}
\end{cases}$$

3. Compute the norm:

$$\tilde{w}_i(n_i,t+1) = \frac{\tilde{w}_i(n_i,t+1)}{\|\tilde{w}_i(n_i,t+1)\|}$$

and according to (3) calculate $f_{i+1}(n_i,t)$.

At the end of this learning procedure a set of components $\{\tilde{w}_i(n_i,t)\}$ (with fixed $n_i$ and t) represents a histogram of the feature vector trajectory as one of the possible phonetic categories described by $\tilde{w}_i(n_i,t)$.

In order to label and segment the speech signal the network operates in the following way: next feature vector $x(t)$ is fed into the network and according to (1) are defined the coordinates of the active unit $n_{i+1}(t)$ at the layer 1-th. Then according to (3) the input vector to the next, i+1 layer $x_{i+1}(t)$ is formed. There is a lot of useful decision rules for phonetic category identification. For example, the input vector $x(t)$ (and the corresponding segment of the speech signal) belongs to the i-th phonetic category if:

$$I = \max \{ \max \tilde{w}_i(n_i,t) \}$$

The Experimental Testing Of The Network

As a speech material we used a sequence of synthetic sounds, consisting of five phonemes. Two of these phonemes (number 1 and 2) simulate two-formant vowels with the following formant frequencies F1=600, F2=2200 and F1=900, F2=1500 Hz. The fifth sound (number 3) is a "fricative" and the other two sounds (the number 3, 4) are stops ("p" and "t"). There is no difference between the short-term spectra of the bursts of the sounds 3 and 4. The speech material contains all possible combinations of this sounds. As an input to the program-synthesizer the string of symbols - the numbers of the sounds is used. For this string the synthesizer produces a synthetic signal in the time domain. The signal contains both stationary and transition regions. We assume that a correct discrimination rule for sounds 3 and 4 (that have approximately equal short-time characteristics of the bursts) must be based on the analysis of a longer interval of speech signal, including the transition region to the next, "vowel" sound. For sampling rate of 10 kHz every 0.8 as the output of the 21-channel bank of filters calculated via
252-point FFT. The central frequencies of the bandpass filter are equally spaced. To eliminate small fluctuations in the amplitude of the filter outputs (resulting from the constant lengths of the FFT and processing step) the outputs are passed through the external filter:

\[ y_i(t) = \max_{t} x_i(t), \Delta t = 25.2 \text{ ms}, \]

where \( y_i(t) \) is the \( i \)-th output of the filters bank.

Before the feature vector \( z(t) \) is fed into the network, a small random noise is added to the components of \( z(t) \). This procedure leads to the non-discrete function of the features.

The network is composed of the five layers of units. Each layer having 8 x 8 units. The weight vector \( w_i(n,t) \) has 21 weights and vectors \( w_i(n,t) \), \( i=2,4,5 \) have the dimensionality of 6. The vectors \( w_i(n,t) \) consists of 6 components. The number of this components is equal to the total number of phonemes in the signal plus "nondefined" type of sound (for example, for representation of transition regions).

The results of the formation process are described in Fig. 1 and 2. Fig. 1 shows values of the weights \( w_i(n,t) \). Each unit \( w_i(n) \) is represented by a graph of the values \( e_i^2(n,t) \). Fig. 2 displays the values of the \( w_i(n,t) \) and \( w_i(n,t) \). The values obtained are in a good agreement with those of the test signal.

The segmentation and labelling was carried out on the same synthetic signal. The length of the record was about 14 s and contained approximately 150 synthetic sounds. The error rate (there was no misclassifications or omissions of sounds, but false insertions) was about 4%. The use of the information from the layers 2-5 derived from the analysis of a longer signal intervals increases the possibility of correct determination of the sounds 3 and 4.

Conclusion

It is shown that multilayered network is capable of effective representation of speech signal parameters. Speech signal analysis algorithm, described in this paper uses this capability to form criteria for phonetic labelling and segmentation of continuous speech waveform. It is shown, that the algorithm, after a simple learning procedure, acquires the ability to produce a quite reliable phonetic identification. The identification involves not only short-term features of the sound, but takes into account the coarticulation effects as well.

Reference

VERWENDUNGSMÖGLICHKEIT DER BINÄREN DARSTELLUNG
VON SPEKTRALMERKMALEN IN WÖRTERKENNUNG

H. Kubzdela
Acoustic Phonetics Research Unit, Polish Academy of Sciences
Niskowskiego 10, 61-704 Poznań, Poland

Einleitung

Das erste Glied in der Kette von Operationen aus denen der Spracherkennungsprozeß besteht, ist die akustische Analyse. Das Ergebnis dieser Analyse hat die Form einer Matrix, die als Bild einer Außergewohnheit betrachtet werden kann. In der automatischen Spracherkennung werden je nach den vorhandenen technischen Mitteln sowie dem Erkennungssystem gestellten Anforderungen, akustische Analyse verschiedener Art benutzt. Tauglich für die automatische Spracherkennung wäre solche akustische Sprachsignalanalyse, welche in ein Bild resultiert, das erstens - ein kleines Informationsvolumen hat, zweitens - nur diese Sprachsignalmerkmale zum Ausdruck bringt, welche bei Dekodierung der phonetischen Information eine große Diskriminierungskraft aufweisen, drittens - von den unregelmäßig auftretenden Merkmalen frei ist. In dieser Arbeit wird gezeigt, wie man die Äußerung in Form eines einfachen binären Bildes, das die oben genannten Anforderungen ziemlich gut erfüllt, ausdrücken kann.

Binäre Sprachbilder mit gewählten Spektralmerkmalen


\[ q_i = c \sum_{j=1}^{i-k} p_j \]

berechnet.

473
Laut dem zweiten Prinzip wird das Spektrumelement im Punkt \( i \) durch die Summe der links von ihm gelegten Elemente maskiert. In dieser Summe werden diejenigen Elemente nicht berücksichtigt, die von der Ort \( i \) näher als \( k \) entfernt sind. Die Grenze, über welche eine volle Maskierung erfolgt, liegt für jeden Schall im anderen Punkt.

Das dritte Prinzip von Bestimmung der wichtigen Spektrumbereichen bildet die verknüpften zwei ersten Prinzipien. Laut diesem Prinzip wird der wichtige Spektrumbereich durch die Punkte bestimmt, in denen die Spektrumshüllende entsprechend konvex ist und zugleich über einen nach der Formel (1) ermittelten Pegel liegt.


Vergleich und Abschätzung der binären Sprachbildarten


Es wurde eine Vergleichsschätzung der Diskriminierungskraft der zwei letzten von drei hier betrachteten Typen der Sprachrepresentation durchgeführt. Ein Grund für diesen Beurteilung waren Resultate der Vergleichungen von binären
Der Vergleich der 20 Äußerungen der zwei einander ähnlichen polnischen Worte, gesprochen von zwei verschiedenen Stimmen, einer weiblichen und einer männlichen. Die Ergebnisse dieser Vergleicherungen werden in die Tabelle 1 eingetragen. Die dabei benutzten Symbole bedeuten folgendes:

PB - Prinzip der Bildbestimmung, EM - eigene Muster, c - Koeffizient von Formel 1, Ng, Ub - die Raten der negativen und unbestimmten Erkennungsresultate.

Tabelle nr 1. Anzahlen der negativen und unbestimmten Vergleicherungsraten für jedes der betrachteten Prinzipien der Bildbestimmung, für einige Werte des Koeffizienten c und für verschiedene Vergleicherungskombinationen.

<table>
<thead>
<tr>
<th></th>
<th>c=1/32</th>
<th>c=1/25</th>
<th>c=1/16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ng</td>
<td>Ub</td>
<td>Ng</td>
</tr>
<tr>
<td>III</td>
<td>FM 15</td>
<td>10</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td>EM 0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>II</td>
<td>FM 12.5</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>EM 2.8</td>
<td>2.8</td>
<td>5.8</td>
</tr>
</tbody>
</table>


**Bemerkungen**

Beides schafft also gute Voraussetzungen zur Verwendung einer vollen Repräsentation des zu erkennenden Sprachelements bei einem knappen Bedarf nach technischen Mitteln.

Literatur


Abb. 1. Beispiele der binären Bilder bestimmten nach den Prinzipien I, II, III (a,b,c) und das komprimierte Bild von b
SPEAKER IDENTIFICATION BY MEANS OF SPEECH SPECTRA FROM KEY WORDS

W. Majewski

Institute of Telecommunication and Acoustics,
Technical University of Wrocław,
ul. P. Prusa 53/55, 50-317 Wrocław, Poland

Introduction

It is well known that speech spectra are highly speaker-dependent and because of that may be used for the purpose of automatic speaker recognition [1,2]. Long-term text-independent speech spectra are, however, inconvenient for practical applications due to a long duration of required speech samples. Thus in the present study a series of speaker identification experiments was carried out under controlled laboratory conditions in order to investigate the possible usefulness of speech spectra averaged over short speech samples /single key words/. In this paper the method of speech signal analysis, the algorithm and results of speaker recognition, as well as the evaluation of effectiveness of adopted parametric representation of speech, will be presented and discussed.

Speech Material

In the experiments 20 male speakers exhibiting no speaking defects were used. For each speaker 20 repetitions of two isolated Polish words, i.e. "logarithm" /logarithm/ and "anorony" /anorony/, were recorded directly on floppy discs by means of a microphone, amplifier and 10-bit A/D converter coupled to an IBM XT personal computer. The utterances selected for the study contained almost exclusively the voiced sounds, especially those considered as good carriers of the individual voice features. The recordings were made during one recording session in an ordinary room of average acoustical parameters /computer laboratory/.
Parametric Representation of Speech Material

Speech spectra averaged over a short specific text/single words/ were obtained by averaging short-time spectra calculated with the aid of FFT algorithm. Before FFT calculations, the sampled speech signal for a given utterance was normalized with regard to the total energy. Similarly like in our previous studies [2,3], it was adopted that each utterance of the speaker will be represented by a vector presenting the speech signal energy in one-third octave bands. The analyses were performed in Hamming windows of ca 100 ms length/N-1024 samples at 10 kHz sampling rate/every each 512 samples, and in the bands of medium frequencies ranging from 100 to 3200 Hz. Since the number of bands P=16, each utterance was represented by a 16-dimensional vector $X_{m,i}$

$$X_{m,i} = \left[ x_{m,i,1}, \ldots, x_{m,i,p}, \ldots, x_{m,i,16} \right]$$  /1/ 

where: $m=1,2,\ldots,K$ denotes the speaker,

$i=1,2,\ldots,I$ denotes the repetition of given utterance,

$p=1,2,\ldots,P$ denotes the frequency band.

In order to evaluate the discriminating power of particular components of parameter vector, the F-ratio [4], i.e. the ratio of interspeaker variance to mean intraspeaker variance was used

$$F = \frac{1}{F-1} \sum_{m=1}^{I} \left( \overline{x_{m,p}} - \overline{x_{p}} \right)^2 \sum_{i=1}^{I} \sum_{p=1}^{P} \left( x_{m,i,p} - \overline{x_{m,p}} \right)^2$$  /2/ 

where: $x_{m,p} = \frac{1}{I} \sum_{i=1}^{I} x_{m,i,p}$, $x_{p} = \frac{1}{I} \sum_{m=1}^{K} x_{m,p}$

The F-ratio values were calculated separately for each word from all repetitions by all speakers. Some vector components exhibited relatively small values of F-ratio and the possibility to eliminate those components was examined in recognition process.
Experiments - Results and Discussion

Speaker identification experiments were performed utilizing the nearest mean /N/ algorithm and decision criterion based on minimum of multidimensional Euclidean distance. The results of speaker identification for 16-dimensional vectors representing the averaged logarithmic spectra of the words "logarytmy", "anonymy" and the both words combined are presented in Table 1. The scores are presented for different length of learning sequence /LS/ and for two kinds of testing sequence /TS/: TS=LS, i.e. TS identical with LS, and for TS\#LS, i.e., for TS formed from all repetitions of the given utterance excluding the repetitions utilized to form the reference vectors.

Table 1. Speaker identification scores for 16-dimensional vectors

<table>
<thead>
<tr>
<th>LS length</th>
<th>Correct &quot;logarytmy&quot; %</th>
<th>Correct &quot;anonymy&quot; %</th>
<th>Identification both words %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TS=LS</td>
<td>TS#LS</td>
<td>TS=LS</td>
</tr>
<tr>
<td>5</td>
<td>99.0</td>
<td>87.5</td>
<td>90.0</td>
</tr>
<tr>
<td>10</td>
<td>97.0</td>
<td>91.5</td>
<td>93.5</td>
</tr>
<tr>
<td>15</td>
<td>96.0</td>
<td>87.0</td>
<td>91.0</td>
</tr>
<tr>
<td>20</td>
<td>93.8</td>
<td>-</td>
<td>90.5</td>
</tr>
</tbody>
</table>

The results presented in Table 1 are encouraging and confirm the usefulness of speech spectra averaged over specific key words for the purpose of speaker identification. The scores for both words combined are higher than for single words, especially for TS\#LS, i.e., for the cases closer to practical applications. The length of LS has also an influence on speaker identification -- to obtain reliable performance the learning sequence should be sufficiently long.

The possible reduction of vector dimension was also examined. The experiment was carried out for the best results obtained for TS\#LS, i.e., for both words combined and for LS=10 and 15. The subsequent components of parameter vector were removed on the basis of the smallest values of F-ratio. The results are presented in Table 2. The data included in this table indi-
cate that a substantial reduction of vector dimension is possible and satisfactory scores may be obtained for a small dimension of parameter vector /e.g. 88 \% for only four dimensions/.

Table 2. Speaker identification scores for different dimensions of parameter vector

<table>
<thead>
<tr>
<th>LS length</th>
<th>Correct identification /%/</th>
</tr>
</thead>
<tbody>
<tr>
<td>p: 16</td>
<td>14 12 10 8 6 4 3 2</td>
</tr>
<tr>
<td>10</td>
<td>94.5 94.0 88.5 82.0 87.5 82.0 85.5 74.0 41.0</td>
</tr>
<tr>
<td>15</td>
<td>99.0 97.0 94.0 92.0 90.0 87.0 86.0 80.0 51.0</td>
</tr>
</tbody>
</table>

Conclusions

It is concluded that speech spectra averaged over a fixed text consisting of single words may be successfully used for the purpose of speaker identification, at least under described conditions of speech production. The presented method does not require the time normalization what simplifies speech analysis procedure. The other positive feature is the possibility of substantial reduction of vector dimensions what is important for a possible real time implementation. It is worth, however, to add that for application of the presented method under less restricted and more realistic conditions it will be rather necessary to utilize speech spectra averaged over longer speech samples, consisting of more than two words used in this study.

References

LA DUREE COMME INDICE DE RECONNAISSANCE ACOUSTIQUE DE DIVERS PROCEDES DE VALORISATION EN FRANCAIS.

par Bernard Flamant

Institut Universitaire de Technologie - Université de Nantes -
Heinlex - B.P. 420 - 44600 Saint-Nazaire (France)

1. Considérations liminaires et finalité de ce travail

A partir d’un quelconque énoncé neutre PNo, diverses phrases comportant des indices de valorisation - lexicaux ou lexicaux-pontématiques - désignables comme soit PV1, PV2, PV3 ... peuvent être constituées. Bien que le contenu logique reste dans un rapport d’identité avec PNo, la valorisation introduit une focalisation sur l’un (ou plusieurs) des termes, occupant une fonction sujet ou autre. Ainsi, dans les énoncés "Pierre est parti" (P0P0), "C’est Pierre qui est parti" (P1P1), "Quant à Pierre, il est parti" (P2P2), "Il est parti" (P3P3), une constante logique lie de façon immuable les contenus de signification : dans ces 4 énoncés en effet, Pierre (sujet) est parti (et ceci, de manière tout à fait irréfutable et que soit le référent). Par contre, les variations énoncatives apportent des modulations, par l’introduction d’une focalisation plus marquée par rapport à PNo : dans PV1, "Pierre" est enchaîné dans la tournure insistant "c’est... qui..."), morèmes considérés comme éléments de présupposition (v. R. MARTIN : 1976, p. 53) ; dans PV2, "Quant à ...", tout en mettant en évidence le sujet "Pierre", détermine une présupposition antonymique (notons que la mise en relèvement se double alors d’une pontématisation et d’une reprise du sujet sous la forme d’un élément de rappel - pronominal) ; dans PV3, "Pierre" est détaché par un double processus : pontématisation + reprise du sujet dans la 2ème partie de l’énoncé sous une forme pronominalisée en relation anaphorique (similairement à PV2) avant détachement du nom (v. P. CADIO : 1988, p.10 et B. FRADIN : 1988, 26-56). Dans le corpus que nous avons établi, ce sera le C.O.D. qui fera, dans tous les cas, l’objet de la valorisation (ceci autorise aussi sa place en position finale d’énoncé). Mise à part l’introduction des procédés de valorisation (in PV1, 2, 3), les énoncés restent sensiblement identiques au plan lexical. Pour les énoncations neutres, la monosyllabie (monosyllabe afin que ne se greffent pas d’éventuels phénomènes d’accentuation secondaire lors de l’oralisation) est placé tantôt en position finale, tentant affecté de l’accent dit "rythmique" (in PNL), tantôt en position inaccentué (in P2P2). Nous nous proposons de déterminer ici l’impact des transformations énoncatives focalisantes sur l’un des paramètres acoustiques de la parole, en l’occurrence la durée. Dans des travaux antérieurs (B. FLAMANT : notamment 1985, 173-207), nous avons pu mettre en relation tout particulièrement la tournure présenteative-subordonnée "c’est... qui/moi(e)" et son incidence sur les 3 paramètres - durée, intensité, Fo - nous permettant d’observer une "concentration tri-paramétrique" (p. 203) sur l’élément valorisé dans ce type de valorisation. Certes la valorisation d’un élément peut s’effectuer sans support syntaxique, sans modification structurelle de l’énoncé ; par un fait d’insistance phonétique seul, la focalisation peut être réalisée et d’énoncés PNo ou PNL du type "Pierre est parti" ou "Il a franchi ce pont" peuvent s’élaborer des variantes prosodiques contrastant avec une dictio plus en conformité avec le constituant syntaxique PNo ou PNL (v. R. MARTIN : 1976, p. 87 ; v. aussi B. MULLIN : 433)
1980, p. 250). Cette focalisation n'est cependant que potentielle et non généralisée car ne bénéficiant pas du support syntaxique de valorisation. Dans les 3 énoncés PV1, 2, 3, la collusion sémantico-syntaxique provoque à contrario une focalisation structurale et, corollairement, une réalisation orale quasi-mont toujours valorisée. Nous avons retenu 126 énoncés courts (6-8 syllabes, donc composés a priori d'un seul élément rythmique) distribués comme suit :

- 21 énoncés du type PV1
- 71 énoncés du type PV2
- 21 énoncés du type PV3
- 42 énoncés du type PN

Les mots-cibles ont été constitués de telle sorte que les combinaisons phonétiques C - V soient les plus nombreuses possibles. Le corpus ainsi établi a été dû par 2 locuteurs masculins sans influence régionale particulière. Leur lecture enregistrée sur bande magnétique (INFR 4701 Report) a été "transformée" en tracés minographiques au Laboratoire de l'Institut de Phonétique de Strasbourg. Une observation initiale permet de constater que le 2ème locuteur (loc. 2) adopte un débit général de la parole légèrement supérieur à celui du 1er locuteur (loc. 1) : 10,54 phonèmes par sec contre 9,49 phon./sec. (en prenant en compte les 126 énoncés).

2 - Données instrumentales et observations

2.1 Valeurs relatives au débit. Suivant les types d'énonciation, le débit présente des valeurs différentes :

<table>
<thead>
<tr>
<th>types d'énonciation</th>
<th>PV1</th>
<th>PV2</th>
<th>PV3</th>
<th>PN1 et 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>locuteurs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loc. 1</td>
<td>9,67</td>
<td>9,52</td>
<td>9,10</td>
<td>11,85</td>
</tr>
<tr>
<td>loc. 2</td>
<td>9,54</td>
<td>10,53</td>
<td>9,03</td>
<td>11,15</td>
</tr>
</tbody>
</table>

Chez les 2 loc., le débit moyen de type PN est plus rapide que celui caractérisant les énoncés PV ; dans ces derniers énoncés, c'est PV3 qui est affecté du débit le plus lent, par conséquent de l'insistance phonique la plus pressante : les durées phonétiques moyennes sont à cet égard très révélatrices : dans les réalisations du loc. 1, les valeurs (en cs) sont les suivantes : 10,40 (PV1), 10,59 (PV2), 12,48 (PV3), 8,53 seulement en PN1 - PN2 ; dans celles du loc. 2, les durées moyennes présentent des clivages similaires : 10,36 (PV1), 9,57 (PV2), 11,19 (PV3), 9,11 (PN1 - PN2).

2.2 - Incidence pausale. Les énoncés PV2 et PV3 se caractérisent graphiquement par une ponctuation qui revêt une fonction à la fois syntaxique, communicative et sémantique (v. L.C. VEDERINA : 1980, p. 65 ; v. aussi N. CATACH : 1980 - entre autres), même si cette triple fonction apparaît en excédent par rapport au rôle déjà focalisateur de la position détachée du C.O.D. (PV3) ou de la position particulière de ce C.O.D. derrière la locution prépositive de mise en relief "quant à" (PV2). Nous désignerons par "pause séparationnelle" l'arrêt vocal effectué dans l'énoncé entre la partie thématique et la partie discursive. Dans PV1, ces pauses sont trop occasionnelles pour être prises en compte : 7 occurrences sur 21 énoncés chez le
loc. 1. 2 seulement chez le loc. 2 (sur 21 énoncés également). Dans PN1 et 
FN2, les occurrences pausales sont quasiment inexistantes chez le loc. 1 : 2 
sur 42 énoncés et elles sont nullles chez le loc. 2. En ce qui concerne 
les taux d'occurrence de ces pauses séparatoires dans PV2 et PV3, ils sont 
beaucoup plus, voire très représentatifs de ces types d'énonciation. En 
contre, les durées moyennes des pauses y sont importantes, tant en valeur abso-
lue (en cs) qu'en valeur relative (en % par rapport à la durée totale des 
énoncés concernés), ce qui est sans incidence sur le débit (cf. 2.1) 
qui procède directement non seulement du temps de parole, mais aussi de la 
fréquence et des durées pausales (v. v. LOCCI : 1974, pp. 146 et sq.):

<table>
<thead>
<tr>
<th>Types d'énonciation</th>
<th>loc. 1</th>
<th>loc. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>nbre d'occurrence</td>
<td>Valeurs pausales moy. abs. (en cs)</td>
<td>Valeurs pausales moy. relat. (en %)</td>
</tr>
<tr>
<td>PV2</td>
<td>21/21</td>
<td>25,45</td>
</tr>
<tr>
<td>PV3</td>
<td>21/21</td>
<td>27,5</td>
</tr>
</tbody>
</table>

7.1 - Valeurs temporelles des mots-cibles et des segments C - V

2.3.1 - Durée des mots-cibles. Lorsque le mot-cible est en position inac-
centuée (PN1), la durée apparaît très nettement déficitaire par rapport aux 
valeurs observées dans les autres types d'énonciation. Par contre, en posi-
tion finale d'énoncé (PN2), cette durée est bien inférieure à celle cons-
tatée dans les PV (loc. 1), ou bien elle lui est supérieure (loc. 2) : les 
écart restent toutefois assez modestes. Dans les PV, un clivage est observa-
ble : PV1 < PV2 < PV3 et ceci, dans les réalisations des 2 locuteurs :

<table>
<thead>
<tr>
<th>types d'énonciation</th>
<th>PV1</th>
<th>PV2</th>
<th>PV3</th>
<th>PN1</th>
<th>PN2</th>
</tr>
</thead>
<tbody>
<tr>
<td>locateurs &lt; loc. 1</td>
<td>79,74</td>
<td>32,55</td>
<td>35,81</td>
<td>29,64</td>
<td>19,65</td>
</tr>
<tr>
<td>loc. 1</td>
<td>30,67</td>
<td>32,62</td>
<td>33,81</td>
<td>34,98</td>
<td>22,05</td>
</tr>
<tr>
<td>durée des mots-cibles (en cs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3.2 - Durée des segments C - V constitutifs des mots-cibles

<table>
<thead>
<tr>
<th>types d'énonciation</th>
<th>PV1</th>
<th>PV2</th>
<th>PV3</th>
<th>PN1</th>
<th>PN2</th>
</tr>
</thead>
<tbody>
<tr>
<td>loc. 1 &lt; C</td>
<td>11,43</td>
<td>10,95</td>
<td>12,21</td>
<td>8,86</td>
<td>7,6</td>
</tr>
<tr>
<td>V</td>
<td>15,93</td>
<td>18,60</td>
<td>19,86</td>
<td>18,14</td>
<td>10,3</td>
</tr>
<tr>
<td>durée des segments C - V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loc. 2 &lt; C</td>
<td>12,40</td>
<td>12,14</td>
<td>13</td>
<td>10,36</td>
<td>8,53</td>
</tr>
<tr>
<td>V</td>
<td>17</td>
<td>18,71</td>
<td>18,83</td>
<td>22,29</td>
<td>11,88</td>
</tr>
<tr>
<td>durée des mots cibles (en cs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

c = segment consonantique initial / V = segment vocalique

Les durées du segment consonantique en énoncé neutre sont toujours in-
férieures à celles relevées en énoncé valorisé ; c'est en PV3 que les va-
leurs sont les plus fortes : l'insistance consonantique y est nette chez les 
2 locuteurs. Concernant les durées vocales, l'accent rythmique conserve 
un rôle important : les valeurs vocales y sont notables et sont même les 
plus fortes chez le loc. 2. En PN2 par contre, n'étant soumises ni à l'ac-
cent, ni à la valorisation, elles demeurent les plus faibles comme c'était d'ailleurs le cas pour les consonnes. Au sein des énoncés valorisés, un accroissement des durées des voyelles apparaît : PV1 → PV2 → PV3.

3 - Synthèse : les cliniques temporels de la valorisation

Quoique notre travail se fonde sur des moyennes qui, par nature, occultent des dispersions au plan des valeurs numériques, des tendances se dégagent :

1) Le débit présente des valeurs inférieures dans les énoncés valorisés par rapport à celles observées dans les énoncés neutres : parmi les énoncés valorisés, c'est l'énoncé segmenté (PV3) qui est affecté du débit le plus lent, en raison des valeurs de durée syllabique, mais aussi de celles concernant les pauses.

2) Un clinage au niveau des durées des mots-cibles apparaît entre les différents types de valorisation : moins la lexicalisation est importante (ou complexe), plus les valeurs de durée de ces mots-cibles augmentent.

3) C'est ainsi que l'absence de lexicalisation dans la partie initiale de l'énoncé (en PV3) provoque un accroissement de la durée du mot-cible, imputable à la fois à une augmentation temporelle des segments C-V constitutifs (sorte de phénomène de compensation) : les valeurs de durée sont toujours en excédent dans ce type d'énoncé segmenté.

4) La position accentuée reste la plus faible quant aux valeurs de durée des mots-cibles qui sont toujours en déficit - avec des écarts notables - par rapport à toutes les autres énoncations.

5) La position finale (ici finale d'élément rythmique) présente une certaine résistance temporelle : si les valeurs de durée de la partie consommatique sont toujours inférieures à celles relevées dans les énoncés valorisés, par contre, la partie vocalique peut présenter des valeurs supérieures à celles observées dans ces mêmes énoncés valorisés (c'est le cas essentiellement des réalisations du loc. 2).

4 - Références

N. CATACH (1980), "La ponctuation", in Langue française, n° 45 (février), Larousse, Paris, pp. 16-27.
A FORMANT ANALYSIS OF ITALIAN DIPHTHONDS

P. Maturi

Dipartimento di Filologia Moderna
Università di Napoli
via Porta di Nave, 1
80133 NAPOLI - ITALY

The aim of this work is to offer a description of the acoustical features of some Italian diphthongs, observed in various contexts.

For this purpose standard Italian diphthongs /ai/ and /au/ were chosen as a first sample, both because they are widely used and because they are made of vowels situated at the periphery of the vowel system, so that transitional phenomena are easier to observe.

Methods and Materials

Both diphthongs were uttered by 4 Italian native speakers from Naples area (3 males and 1 female) in the following contexts:

(a) in isolation:
   ai - au
   maic - Saul

(b) in isolated monosyllables:
   faida - nasa

(c) in isolated bisyllables:
   non dire mai di no -
   ho visto Ugo con Ray

(d) in monosyllables within a sentence:
   il daino ha fatto una
   lauta cena
   De Naurto torna oggi dal
   Cairo

(e)-(f) in bisyllables within a sentence:
   apprezzo molto la
   laicità di Maurizio.

A spectrographic analysis was made of all uttered diphthongs (14 x 4 subjects) by means of a Kay DSP-Sonograph, and the frequencies of the first two formants were measured by windows of 12.5 ms.

The formants of a diphthong can be divided into three successive phases (see Fig. 1):

(1) a phase of approach to the centre values of the first vowel (V1);
(2) a phase of transition from V1 to the centre values of the second vowel (V2);
(3) a phase of release.

This work only takes into account:

(1) formant frequencies measured at points V1 and V2, for the purpose of giving a 'static' description of the diphthongs, i.e. of determining their actual elements for a narrow phonetic
transcription; (II) the variation of frequency values all through stage (2), in order to give 'dynamic' information about the diphthongs.

Stages (1) and (3), on the contrary, are left out, as their features seem not to be specific for diphthongs, but common to the generality of vowel phenomena (e.g. in groups O-V or V-O).

Results and Discussion

The average values obtained for the two elements of either diphthong are as follows:

/ai/: V1 (a-/) F1 = 728 Hz F2 = 1429 Hz
V2 (-i-) F1 = 366 Hz F2 = 2176 Hz

/au/: V2 (a-) F1 = 752 Hz F2 = 1269 Hz
V2 (-u-) F1 = 415 Hz F2 = 783 Hz

Values for /a/ in both diphthongs fall within the area of It. [a], although they are considerably oriented towards [æ] and [a] respectively (as is shown by the difference between F2 of /a/- in /ai/ and in /au/).

Average values for /-i/ and /-u/, on the contrary, stand out of the areas of [i] and [u], actually falling within [e] and [o] areas (see Fig. 2).

So at this first level of the analysis, it clearly comes out that the most suitable transcription of the diphthongs here examined would be [æ] and [o].

A closer look into the realizations of /ai/ and /au/ in the different contexts chosen brings out the following (see Fig. 3):

/ai/: V1 coincides with [a] only in contexts A-E-G-F, while in D-E-G it is a less open vowel, [æ] or [æ];
V2 falls within the [i] area in A-E-C (though very [e]-oriented), coincides with [e] in E-D-F, and lies between [æ] and [ε] in E;

/au/: V1 always coincides with [a] except in context G, where it's realized somewhere between [o] and [ɔ];
V2 is realized as [u] only in A-E, while elsewhere it is [o].

In brief, the realizations (average of all four speakers) are as follows:

Fig. 2 - Average values of V1 and V2 in /ai/ and /au/ compared with the areas of existence of standard Italian vowels (from Ferrero et al., 1978). Dots refer to /ai/, x's to /au/.
As for the 'dynamics' of our diphthongs, a complete description of all its aspects would require much more space than it's available here. We will therefore limit ourselves to a few considerations about this point.

First of all, diagrams showing the variations of formant values in the various contexts (see Figure 4, relative to subject CA) bear evidence of a surprising uniformity in the 'itineraries' followed, in spite of the large diversity of their points of departure and arrival (V1 and V2, see above).

Even very anomalous realizations, like for example D and G in Fig.4, show a strong tendency to reach the main path, before they bend toward their final target. This main path runs through the areas of [e], [e], for /ai/, and of [ɔ], [o] for /au/.

We checked these experimental data by listening to small portions of the diph-
thongs and trying off-hand percetive tests with unaware listeners: it was easy for everybody to recognize as \( [ \alpha ], [ \theta ], [ e ] \) intermediate portions of /ai/ and as [r], [o], [o], successive segments of /au/.

The interesting results obtained and here briefly presented seem particularly useful in the field of vocal synthesis.

Experiments of vowel and diphthong synthesis carried on on the basis of similar methods gave in fact very good percetive results and confirmed our opinion about the necessity of an accurate dynamic description of all aspects of Italian phonetics (Cutugno et al 1989; also Naturi 1989).

References
-Ferrero F., Genre A., Boë L.J., Contini M., Nozioni di Fonetica Acustica, Omega, Torino, 1979;
Temporal variables and fundamental frequency following unilateral left anterior or posterior lesion
Parth M. Bhatt, Experimental Phonetics Laboratory, University of Toronto, Toronto, Canada.

I. Introduction
This paper presents the results of a spectrographic and oscillographic analysis of temporal variables and fundamental frequency in the spontaneous speech output of six francophone, adult, right-handed patients: three, with unilateral left anterior lesion and three with unilateral left posterior lesion.

II. Clinical and instrumental studies of prosodic modification
Previous clinical reports indicate that in the initial stages of severe Broca's aphasia patients are still able to produce short utterances with some degree of melodic variation (Alajouanine and Lhermitte, 1964; Botez et al., 1968; Brissaud, 1901; De Blezer and Poecck, 1984). More specific prosodic modifications due to left anterior lesion may occur in subsequent stages of the patient's evolution. A. Pick (1913) and G. Monrad-Krohn (1947) reported the first cases of foreign language dysprosody, in which a "foreign accent" appears because of direct modification of tonal or accentual systems (displacement of fixed accent in Czech and inability to produce lexical tone in Norwegian). Some clinicians now use the term dysprosody to designate all cases of foreign pronunciation (Bhatt, 1987a), whether or not prosodic systems are directly involved. The term dysprosody has also been used to describe the flat, discontinuous speech melody produced by patients suffering from severe non-fluent aphasia accompanied by agrammatism (Goodglass; 1968).

Instrumental studies of prosodic modification following left anterior lesion have concentrated on describing the speech output of subjects with Broca's aphasia. Studies by Danly, et al. (1979; 1982) showed that F0 sentence declination was present but applied to smaller domains and that major F0 falls occurred in sentence final position. Their subjects made a greater number of F0 rises than controls, but did not encode sentence length by choosing a high initial F0 peak. Rynals (1962; 1984) found a reduction of F0 range in eight subjects with Broca's aphasia and a higher average F0 in subjects with large anterior left hemisphere lesions. Cooper et al. (1984), also found that Broca's aphasics produce a flat F0.

On the other hand, prosodic systems and temporal variables are generally considered to be unaffected by posterior lesions to the left hemisphere. Patients produce fluent speech output with apparently unmodified intonational variations. Duchan et al. (1980) reported that their patient possessed a rich intonational inventory (contrastive intonation, parentheses, questions). Cooper et al. (1979) reported that their patient showed an overall F0 declination, but that the initial peak was higher than that of control subjects. They also found that their patient made a greater number of F0 resets that did controls. Danly et al. (1983) confirmed these results and reported that F0 declination was applied to smaller domains than in controls and that their patients produced a greater number of continuation rises. Cooper et al. (1984) reported that their patients produced unusually high initial and medial F0 peaks.
II. Subject population speech sample and phonetic instrumentation

The subjects for this study were six francophone, adult right-handed patients. Subjects A, B and C had suffered unilateral left anterior lesions: Subjects D, E and F, unilateral left posterior lesions. There were two female subjects, A and C, and four male subjects, B, D, E, and F. At the time of interview, Subject A was 23 years of age, Subject B, 63, Subject C, 36, Subject D, 71, Subject E, 50 and Subject F, 64. Five of the six patients had suffered cerebrovascular accidents: Subject A, a thrombosis of the internal carotid artery; Subject B, an occlusion of the middle cerebral artery; Subject C, an occlusion of the internal carotid artery; Subject D, an embolism of cardiovascular origin, obliterating the middle cerebral artery and Subject F, a thrombosis of the left sylvian artery. The sixth subject, E, suffered from a cerebral tumor, an astrocytoma.

Subjects A, B and C suffered from severe right hemiplegia and right hemisensory neglect. Subject D suffered initially from a left hemiparesis. Subjects E and F showed no signs of motor or sensory impairment. All subjects were neurologically stable at the time of interview.

The speech sample used for instrumental analysis was taken from the clinical aphasias batteries currently in use at the Salpêtrière and St. Anne Hospitals in Paris, France. The patients were replying to questions about their illness or their profession. For each subject a total corpus of approximately 300 syllables of spontaneous speech was analyzed. This speech sample was submitted to two parallel instrumental analyses. The first analysis was carried out by a digital real-time fundamental frequency analyzer (Pitch Machines FM-100) and the second by a digital real-time colour spectrograph (Pitch Machines RT-1000).

IV. Results of instrumental analysis and discussion

The results in Figure 1 below suggest that neither F0 range nor average phrase or sentence final F0 movements clearly differentiated the subjects speech output. On the other hand, speech signal and pause ratio (Figure 2) provided strong distinguishing variables. Subjects with left anterior lesion produced a much higher pause ratio than subjects with left posterior lesion.

The data presented in Figure 3 allow us to further clarify this portrait. The two sets of subjects are not strongly differentiated by articulatory rate or by average pause duration (Goldman-Eisler, 1968; Grosjean and Deschapelles, 1972). When actually articulating speech or when pausing, the two groups behaved in a surprisingly similar fashion. On the other hand, the highly significant variables appear to be the total number of pauses and overall speech rate. Subjects with left anterior lesions produced a much higher number of pauses and a much lower speech rate.

These results suggest that the discontinuity of the melodic line (the presence of numerous pauses) is the primary differentiating factor between these two groups of subjects. This discontinuity, combined with the recurrence of rudimentary intonational patterns (Bhatt, 1987a, b) provides a more accurate description of the speech output of subjects with left anterior lesions and may be responsible for the clinical impression of pitch flatness. Finally, subjects with left posterior lesion do not show a wider pitch range or a higher articulatory rate when compared with subjects with anterior lesion.
### Table 1: Frequency data.

<table>
<thead>
<tr>
<th></th>
<th>Non-final Fo (in Hz)</th>
<th>Total Range (in Hertz)</th>
<th>Range Coefficient</th>
<th>Average Rise (% of Fo)</th>
<th>Average Fall (% of Fo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>172 Hz</td>
<td>150-380</td>
<td>1.329</td>
<td>47.47</td>
<td>33.03</td>
</tr>
<tr>
<td>B</td>
<td>98 Hz</td>
<td>80-190</td>
<td>1.222</td>
<td>32.83</td>
<td>19.37</td>
</tr>
<tr>
<td>C</td>
<td>186 Hz</td>
<td>135-375</td>
<td>1.290</td>
<td>34.22</td>
<td>36.13</td>
</tr>
<tr>
<td>Avg.</td>
<td></td>
<td></td>
<td>1.247</td>
<td>38.17</td>
<td>29.51</td>
</tr>
</tbody>
</table>

| σ     | 0.0898               | 6.00                   | 7.28              |
|       | D                    | 118 Hz                 | 85-180            | 0.803                  | 20.94                  |
|       | E                    | 133 Hz                 | 105-190           | 0.638                  | 25.54                  |
|       | F                    | 109 Hz                 | 90-180            | 0.819                  | 35.82                  |
| Avg.  |                      |                        | 0.753             | 27.33                  | 23.54                  |

| "t"   | 1.680                | 4.274                  | 2.410             | (0.001)                | (0.020)                |
| p(4)  |                      |                        |                   | (0.001)                | (0.001)                |

**Figure 1.** Frequency data.

### Table 2: Temporal variables.

<table>
<thead>
<tr>
<th></th>
<th>Total signal duration</th>
<th>Speech signal duration</th>
<th>Pause duration</th>
<th>Speech signal ratio</th>
<th>Pause ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>22,552 ms</td>
<td>8,922 ms</td>
<td>13,631 ms</td>
<td>39.55%</td>
<td>60.45%</td>
</tr>
<tr>
<td>B</td>
<td>15,141 ms</td>
<td>9,296 ms</td>
<td>5,844 ms</td>
<td>61.41%</td>
<td>38.59%</td>
</tr>
<tr>
<td>C</td>
<td>21,277 ms</td>
<td>7,602 ms</td>
<td>13,675 ms</td>
<td>37.73%</td>
<td>62.27%</td>
</tr>
<tr>
<td>Avg.</td>
<td>19,657 ms</td>
<td>8,597.3 ms</td>
<td>11,059 ms</td>
<td>48.23%</td>
<td>51.77%</td>
</tr>
</tbody>
</table>

| σ     | 3,235.50             | 727.26                 | 3,261.24       | 10.76               | 11.31       |
|       | D                    | 7,997 ms               | 6,111 ms       | 1,777 ms            | 77.47%      |
|       | E                    | 11,517 ms             | 9,431 ms       | 2,086 ms            | 81.89%      |
| Avg.  | 9,581 ms             | 6,438 ms              | 3,143 ms       | 67.19%              | 33.81%      |

| σ     | 1,492.63             | 1,493.96               | 584.87         | 6.16                | 6.61        |
|       | "t"                  | 10.037                 | 9.895          | (0.001)             | (0.001)     |
| p(4)  |                      |                        |               | (0.001)             | (0.001)     |

**Figure 2.** Temporal variables.

### Table 3: Pause and speech rate data.

<table>
<thead>
<tr>
<th></th>
<th>Number of pauses</th>
<th>Average pause duration</th>
<th>Articulatory rate (syll/sec)</th>
<th>Speech rate (syll/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>98.00</td>
<td>139.09 ms</td>
<td>3.35</td>
<td>77.04</td>
</tr>
<tr>
<td>B</td>
<td>100.00</td>
<td>48.43 ms</td>
<td>2.99</td>
<td>102.60</td>
</tr>
<tr>
<td>C</td>
<td>192.00</td>
<td>98.39 ms</td>
<td>4.20</td>
<td>82.90</td>
</tr>
<tr>
<td>Avg.</td>
<td>112.33</td>
<td>96.30 ms</td>
<td>3.59</td>
<td>87.60</td>
</tr>
</tbody>
</table>

| σ     | 18.87            | 37.07 ms              | 0.494                        | 10.83                  |
|       | D                 | 39.00 ms              | 4.46                         | 195.80                 |
| E     | 50.00            | 41.72 ms              | 3.74                         | 163.80                 |
| F     | 56.00            | 56.12 ms              | 4.36                         | 184.20                 |
| Avg.  | 48.00            | 48.20 ms              | 4.19                         | 181.50                 |

| σ     | 7.48             | 5.97 ms               | 0.318                        | 13.60                  |

| "t"   | 17.712           | 10.150 ms             | 0.941                        | 26.880                 |
| p(4)  | <0.001           | <0.010 ms             | <0.500                      | <0.001                 |

**Figure 3.** Pause and speech rate data.
REFERENCES


A DIGITAL DYSPHONIC VOICE ANALYZER

P.Jospa, J.Schoentgen(*)

Institut de Phonétique de l'Université Libre de Bruxelles,
50 av F.D.Roosevelt, 1050 Bruxelles, Belgique.

Introduction

A pseudoperiodic voice source signal is the acoustic result of the vibration of the vocal folds. Their vibratory pattern can be disturbed by asymmetries in their mass or tension (Isshiki 1972, Isshiki and Ishisaka 1976) - a consequence of, for example, the presence of tumors, cysts or the paralysis of the folds.

A stand-alone device for the extraction of two parameters quantifying the cycle to cycle perturbations of the vibrating folds has been developed by our laboratory. An index of the fluctuations in the duration of a sequence of glottis cycles (jitter) and an index of the variations in a sequence of amplitudes (shimmer) have been computed on the base of a sustained vowel signal recorded on-line via a microphone.

The procedure is based on a so called short time amplitude variation criterion (AVC) outputted by a new speech signal preprocessing method (Jospa, 1982) which allows for the reliable processing of male, female and children’s voices (F0 = 40 ... 500 Hz). A study carried out previously (Schoenjen, 1981, 1988, 1989) on dysphonic voice samples demonstrated the high reliability of the new preprocessing method as compared to analysis by linear prediction (e.g. Davis, 1979) or low-pass filtering. Its superiority is especially manifest in the case of high-pitched voices.

Short time preprocessing and the perturbation parameters computation.

The AVC preprocessing method continually estimates the short time variations in the speech signal envelope in a chosen frequency band. In the case of voiced speech segments, the AVC signal displays a well-marked positive peak in the vicinity of glottal closure, and a negative minimum in the region of fast decreasing signal energy (figs. 1,4). Starting from here, it is easy to estimate the duration and amplitude of individual glottal cycles. This amounts to the detection of a sequence of positive AVC peaks higher than a dynamically adjusted threshold. The elements of AVC computation consist mainly of a signal envelope extractor in the first formant region, followed by a low-pass differentiator filter. The purpose of this latter stage is to smooth the signal envelope and to calculate its finite difference:

$$AVC = e(t + \Delta t) - e(t - \Delta t), \quad \Delta t \text{ time interval being a function of the bandwidth of the low-pass differentiator (a FIR filter).}$$

Period and amplitude perturbation parameters (PPQ and APQ) are defined as follows (Koike, 1973):

(*) Fonds National de la Recherche Scientifique
\[ N = (k-1) \sum_{i=1}^{k} \left| \left( \sum_{j=1}^{N} x_{i+j-1} \right) - kx_{i+m} \right| \]

\[ \frac{1}{N} \sum_{i=1}^{N} x_i \]

with \( N = 15 \), \( k = 5 \), \( m = (k-1)/2 \), and \( x_i \) the duration of the \( i \)th cycle, or, alternatively, with \( x_i \) the amplitude of the \( i \)th cycle as measured from the AVC. The calculations are carried out in two stages. The first stage is real-time. It consists of the sampling of the speech signal at 10 kHz, the AVC computation, the detection of a sequence of 16 relevant AVC peaks (i.e., 15 glottal cycles) and the storing of respectively 2500 speech signal and AVC samples in the device's on-board memory. The second stage computes fundamental frequency and average glottal cycle duration over 15 pseudoperiods, and also the period and amplitude perturbation quotients according to (1). FIR interpolation, to increase the sampling frequency, is carried out locally in the vicinity of the AVC peaks; the goal is to enhance measurement precision (an overall precision of around 10 microseconds is so achieved).

**Description of the analyzer**

The design of the voice analyzer is based on two commercially available microprocessor chips: a Texas Instrument TMS 32010 for high speed signal processing and computation, and an Intel's 8031/51 microcontroller for overall device control. Input/output management by the 8031, and signal processing by the TMS 32010 are carried out in parallel, resulting in a real-time computation of the AVC up to 10 kHz. The analyzer features a microphone signal input, an line input (for inputting signals already recorded on another storage medium), a keyboard and pushbuttons, a liquid crystal display, a dual-beam oscilloscope and a built-in printer. The system's software, together with a monitor, is stored in on-board EPROMS. A built-in select-button puts, under monitor control, either the EPROM resident software, or a host computer resident software. This feature is intended to facilitate future software improvements and updating.

The system's block-diagram is shown in figure 2. Figure 3 gives a global view of the analyzer prototype. The device is portable and stand-alone. It is easy to use and without any discomfort for either operator or patient. The patient sustains a vowel [a] for, typically, a couple of seconds; the operator adjusts the recording level taking into account the instructions given by the machine, and then starts both storage and analysis by pushing a single button. Analysis results are displayed in quasi real-time. At power-up an environmental noise test program is run automatically; its object is to prevent analysis in too noisy an environment. During signal acquisition, the system monitors signal levels continuously and forbids analysis on too feeble a signal or in the case of an A/D converter overflow. During this stage the system displays both the speech signal and the preprocessed signal in real-time on an oscilloscope screen. The objective is to allow for a visual appreciation by the operator of the patient's speech signal. Once stored in memory, both signals can be made to scroll forwards and backwards, or to stop on any window of 256 stored samples (25.6 milliseconds).
Clinical Evaluation

The device has been tested on 160 speakers in a clinical environment, i.e. 105 control speakers without any laryngeal complaints (56 males and 49 females) and 55 dysphonic speakers (29 males and 36 females). A total of 1600 measurements were carried out (10 measurements for each speaker). A first examination of the results confirms the reliability of the device and the relevance of the measurements obtained. A statistical analysis of the data collected so far is in under way.

Applications

As far as clinical applications are concerned, acoustic measurements are a non-invasive means of characterizing dysphonic voice signals reproducibly, quantitatively and objectively. We think that the relevance of these measurements should be evaluated in the framework of a precisely defined task. Possible clinical tasks are:

i) Follow-up, i.e. charting the evolution of a patient's voice in the course of treatment.
ii) Comparison, i.e. a voice signal is compared and classified with reference to a set of voices which are well known to the laryngologist and which therefore play the role of an ad hoc standard.

iii) Documentation, i.e. the building of a database for further reference. In the long run, accumulating data on a large number of voices would also enable statistics to be established on different speaker categories and different pathologies.

iv) Expertise in a legal framework involving accidental damage to a person's larynx.

v) Screening for vocal pathologies, especially cancer of the vocal cords.

Finally, it is suggested that the AVC signal, whose shape depends mainly on laryngeal control, may be useful as a visual feedback signal during the vocal training of the hearing-impaired.

When proceeding from tasks (i) to (v), it must be understood that demands for knowledge about voice quality in the population at large are becoming more and more pressing. However, as yet, no model is available that describes voice quality in the general population in quantitative terms. We do not believe that task (v), for example, can be tackled successfully without completing task (iii) beforehand (e.g. Kasuya et al, 1986). Our experience suggests that applications (i) to (iii) can be attempted successfully at the present time. However, screening and expertise, based on acoustic measurements alone, must wait until these limitations are better understood with reference to the population at large.

Bibliography


Isshiki N. (1972), Studio Phonologica, VI, pp 38-44.


Jospa P. (1982), Rapport d'Activités de l'Institut de Phonétique de l'Université Libre de Bruxelles, 17, pp 89-108.


Schoentgen J. (1982), Speech Communication, 1, 3-4, pp 269-282.


Schoentgen J. (1989), Speech Communication, 8, 1.
Figure 1: Top: speech signal
Bottom: preprocessing signal (AVC).

Figure 2: System's block-diagram.

Figure 3: 

Figure 4: Top: speech signal.
Bottom: preprocessing signal (AVC) (dysphonic voice).
Multichannel approach for speech intelligibility assessment

F. Kehlstadt, L. Morellini, E. J. Rathe, M. Rossi

Acoustics Laboratory, Swiss Federal Institute of Technology,
1015 Lausanne, Switzerland

A monitoring and control system with four independent signal paths and individual binaural outputs to each test person allows both the conventional methods and variety of new methods for the subjective assessment of speech intelligibility. It includes facilities for determining audiometric parameters. Both the system and its applications will be discussed.

Introduction

SESI is a subjective intelligibility assessment workstation, the function of which is to control and monitor by computer intelligibility and articulation tests.

Its purpose is to establish a reference data base of the intelligibility in transmission channels in relation to the channel's parameters, such as bandwidth, disturbing noises, time or frequency related distortions, transmission modes, etc.

This data base will be used later on as a reference for an objective measurement apparatus, i.e. one without listeners.

The workstation

The computerized element of the station is made up of a compatible PC, inside of which is a GDTB IEEE488 card that enables communication with the other peripherals of the station. A second card reproduces two 10 second signals, sampled with 16 bits at 50 kHz, which are mainly used to generate two disturbance signals (sampled noises). The stimuli are stored on a 135 MB hard disk, which can store (memorize) up to a 1000 seconds of stimuli or disturbing noises.

The second element is made up of a mixer-conditioner-amplifier rack (MCA) which comprises all the analog functions of switching, switchable gains, controlled attenuations, additions, signal measurement, amplified outputs for the audiometric headphones, as well as intercommunication system between the operator and the subjects. Most of these functions are controlled by the PC through the GPTR connection.

The MCA has the following characteristics:

- 4 signal inputs, line level;
- 1 microphone input for the acquisition of phonemes;
- 3 filtered inputs, coming from the D/A convertors in the PC (1 stimuli signal, i.e. words or sentences, and two signals of disturbing noise, i.e. speech, road noise, stationary noise, etc., both of which are meant to be either electronically added to the stimuli or reproduced in the room where the subjects are sitting);
- 6 individual stereo outputs for audiometric headphones;
• 1 filtered output (16th order) meant for the A/D converter in the PC for the data acquisitions;
• 4 auxiliary outputs (used for example either for connecting one or more external channels or disturbance systems that are to be assessed or compared, or for connecting a power amplifier system in order to reproduce noises around the subjects);
• Measurement of the sound level over a dynamic range of 140 dB in 32 points of the apparatus for the automatic testing of the switching and attenuation functions;
• 1 input for a sound pressure measurement probe for the calibration of the headphones.
• Subject-operator intercommunication system that may either be controlled by the PC or manually controlled by the operator.

The interpretation of the words emitted by the station is generally directly supplied by the subjects. They just have to use the 6 touch keyboard or specially made miniterminals (MT), the 2 X 40 characters screens of which display a maximum of 6 variations of the word perceived (only one of which is correct). The MT's, like the MCA's, are controlled through the GPIB bus.

A typical application of the station is shown in figure 1. In this case the parameters studied may for example be the disturbing noise level, relative to the stimuli of transmission or in the reception room, or the characterization of the transmission channel being studied.

Figure 1.

Audiometric test

At the beginning of each series of intelligibility test, it is necessary to measure the threshold of hearing of the subjects in order to make a suitable selection: a too greater dispersion of the value of the threshold would make it more difficult to interpret the results.

It is also necessary to relativize the intelligibility results in relation to the subjects' hearing, in order to obtain values that only depend in the end on the disturbing noises being studied and no longer on the individual subjects.

A simultaneous measurement for each frequency on the six subjects (individual adjustment of the level of this threshold with each MT) enables:
1) the test to be automated and as short as possible;
2) the subject to become familiar with the use of the MT.

**Intelligibility test**

A great bibliographic research was undertaken in order to establish the specifications of the workstation. It covers a maximum of references concerning the subjective and also objective evaluation of the intelligibility and articulation. It starts from the works of the first half of the century and goes up to current works. It is administered on a computerized data base that continues to grow as new works are published. This led us to envisage using three main procedures for the intelligibility tests, controlled partially or totally by the computer.

**First case**

The words are inserted (or juxtaposed) into an accompanying sentence, the role of which is to familiarize the subjects to the sound level while also attracting their attention. For each word "pronounced" by the station, there is a list of 2 to 6 words of virtually the same type of simple logotome or CVC (consonant-vowel-consonant), or multisyllabic word, amongst which is the correct one. A written list is then proposed to the subjects on the miniterminals, from which they choose their "correct" version by pressing the corresponding key on the keyboard. Once all the answers have been received, the next word is pronounced. In this case, the interpretation results are immediately received (keys pressed by the subjects).

The tests that can be carried out in this way are mainly the "Modified Rhyme Test" (MRT), the "Diagnostic Rhyme/Alliteration Test" (DRAT; DAIT), the "Four Alternative Auditory Feature test" (FAAF).

**Second case**

There is no written list of words close to the correct word. The classic procedure consists either in making the subject(s) phonetically write the word perceived or in asking the subject(s) to repeat the word. The result is taken down by the test director.

We chose to implement the second procedure. The oral answer of a subject is received by the director thanks to an intercommunication system. The director takes down the binary results, by using his own terminal (console) on which he can read the correct answer. The subjects are asked one at a time in writing (through the terminal) to give an oral answer.

The tests that can be carried out in this way are mainly the "Rhyme Test" (RT), the "Phonetically Balanced words" (PB words) and the "Fournier Test".

**Third case**

Comparative tests between either several different disturbing noises or several transmission channels or simulations are carried out. The subjects hear the same word or sentence, one at a time, through various switchable paths. The subjects then indicate for example the one which seems to be of the best quality. The reception is therefore carried out at the terminal of each subject.

The tests that can be carried out in this way are mainly the "isopreferences test" and the "Speech Reception Threshold test" (SRFT).
Current applications

The organigram of three tests representing the three different types of intelligibility measurements have been implemented in the workstation. Lists of words and logotomes in several languages have been sampled and stored on a hard disk, as well as their written variations when available. We can therefore carry out the following tests : RT, MRT and the Fournier test. The first experiments on the influence of the intelligibility of several simple parameters such as the signal-to-noise ratio, spectrum limitations or the compression of the dynamic range allowed us to become familiar with the workstation.

Conclusions

We wished to have a tool that was both as flexible and non limiting as possible as regards the choice of both the types of tests and the degradation applicable to the test signals. The future applications envisaged consist on the one hand in applying the other tests mentioned and translating them into the languages in which they do not yet exist and studying on the other hand the influence of the parameters such as :
- harmonic and inharmonic distortions phase and pitch distortions, etc.,
- time distortions, interruptions, etc.,
- non stationary noise or signal dependant noise, etc.
- linear or non linear phenomena and distortions of all kinds.

These degradations will be obtained by means of simulators specially developed to this effect. The evaluation and calibration of already existing or future objective forecasting methods will thus be easy to undertake.

Bibliography

"Articulation testing methods". Egan J.P. Laryngoscope Vol 59 No 9 pp 955-991 (1948)
"Test of phonemic differentiation: the Rhyme Test" Fairbanks G. JASA Vol 30 No 7 pp 596-600 (July 1958)
"Improvement method for evaluating speech transmission circuits" Munson W.A., Karlin J.E. JASA Vol 34 No 6 pp 762-774 (June 1962)
"Diagnostic evaluation of speech intelligibility" Volers W.D. Benchmark Papers in Acoustics V.11: Speech intelligibility and speaker recognition; Mones F. Hawley: Bowde, Hutchinson & Ross, Inc.
"Relation between Intelligibility Scores for Four Test Methods and Three Types of Speech Distortion" Williams C. E., Hecker M. H. L. JASA Vol 44 No 4 pp 1002-1006 (1968)
"Méthodes d'évaluation de l'Intelligibilité et de la qualité de la parole". Risset J.C. JEP Journées d'étude sur la parole 1971
"Evaluation subjective de la qualité des systèmes de transmission de la parole". Beckers J-P. Electroacoustique et musique pp 91-102
"Speech Intelligibility tests and their standardization" Krider K.D. FASE Symposium Liège (1973)
A PITCH AND FORMANT FREQUENCY TRACKING SYSTEM USING A DYNAMIC PROGRAMMING ALGORITHM BASED ON VARIOUS KINDS OF KNOWLEDGE

Shozo Makino, Toshihiko Shirokaze and Ken'iti Kido

Research Center for Applied Information Sciences, Tohoku University, 2-1-1 Katahira, Sendai, 980 Japan.

Introduction

Pitch and formant frequencies are very important in speech recognition and synthesis problems. It is desirable to build up a speech database which include such information. For the automatic extraction of such information from the speech signal we use a pitch and formant frequency tracking system which uses knowledge about the speaker and the kind of phoneme. By this we hope to get more exact data than by the conventional methods.

This paper describes a pitch and formant frequency tracking system based on various kinds of knowledge. This method will be effective especially for speech data with the corresponding phoneme labels in a speech database.

Outline of a pitch and formant frequency tracking system

The speech is sampled at 10 kHs. The first order-derivative is computed for the speech. The voiced periods are extracted using power and zero-crossing rate. Ten candidates for the pitch frequency are extracted for each analysis frame. An optimum pitch frequency set is picked up using a dynamic programming algorithm from the candidates of each frame taking into account temporal continuity and reliability of the candidates. Formant candidates are extracted using a pitch-synchronous covariance linear prediction (LP) method of which order is controlled by the value of the extracted pitch frequency. An optimum formant frequency set is extracted in a similar fashion, where frequency range of the formant is also controlled by the value of the extracted pitch frequency. Finally, the extracted formants are smoothed with a median filter. Hereafter, we will describe the detail of the pitch and formant frequency tracking method.

A pitch frequency tracking using a dynamic programming algorithm

The conventional methods for the pitch frequency extraction only take account into the temporal continuity over a short period of 30 ms. As we would like to consider temporal pattern of the pitch frequency over input speech, we introduce a global cost function for estimating continuity of the pitch frequency over the input speech. At first ten peaks are picked up from an auto-correlation function of each frame and they are regarded as candidates of the pitch frequency. When \( P_{n,i} \) and \( PH_{n,i} \) are defined as the frequency and amplitude of the \( i \)-th peak at the \( n \)-th frame, a local cost function for connection between the \( n \)-th peak at the \( (i-2) \)-th frame, the \( m \)-th peak at the \( (i-1) \)-th frame and the \( k \)-th peak at the \( i \)-th frame is given as follows:

\[
C(i,k,m,n) = -W1\cdot PH_{n,i} + W2\cdot |P_{n,i} - P_{n,m} - k| + W3\cdot |P_{n,i-2} - P_{n,m} - k|
\]  

(1)

Here, \( W1 \) is a weighting coefficient taking into account amplitude of the peak and minus sign is attached to \( W1 \) so that the peak with larger amplitude can be picked up. \( W2 \) and \( W3 \) are those taking into account temporal continuity of the peaks. The global cost function \( S \) for one extracted voiced period with \( N \) frames is represented by minimum of the sum of the local cost functions as shown in the equation (2).

\[
S = \min_{k,m,n} \sum_{i=1}^{N} C(i,k,m,n)
\]  

(2)

A53
The equation (2) is solved by the following equations (3) and (4) based on a dynamic programming algorithm.

\[ S(i,k) = \min_{m} \left[ S(i-1,m) + C(i,k,m,n) \right] \]  
\[ S = \min_{k} S(N,k) \]  

The proposed method was examined for fifty Chinese mono-syllables uttered by a male speaker because the Chinese speech has four tones and is convenient for estimating temporal continuity of the pitch frequency. The following three pitch frequency extraction methods are compared: 1. the method using the auto-correlation of the original waveform, 2. the method using the auto-correlation of the residual waveform obtained by 14-order LP analysis, and 3. the method using Cepstrum analysis. The weighting coefficients \( W_1 \), \( W_2 \) and \( W_3 \) in the equation (1) were set to \( 200, 1 \) and \( 1 \) based on preliminary experiment results. Table 1 shows ratio of correct samples to all samples, where the pitch frequency extraction was regarded as incorrect if the speech sample contains a frame in which the estimated pitch frequency was more than 10Hz apart from that given by manual inspection. In the conventional method, the pitch frequency was given as that with the largest peak of the auto-correlation function. The proposed method remarkably improved the accuracy of the pitch frequency tracking.

Based on the comparison of the three pitch frequency extraction methods, the following results were found:

a. The proposed method showed the largest improvement in the method using the auto-correlation function obtained from the original waveform.

b. The method using the auto-correlation function obtained from the LP residual waveform showed poor results around the beginning and end points of the speech because it was much affected by noise.

c. The method using Cepstrum analysis showed the highest accuracy of the three methods, especially in case of the noise. However, it required the largest computational amounts.

Formant frequency tracking system using a dynamic programming algorithm based on various knowledge

We use the similar method for the formant frequency tracking to the method for the pitch frequency tracking. At first a pitch synchronous covariance LP method is carried out every one point of speech data in the extracted voiced period. The point with the minimum residual power is regarded as a starting point of a pitch period. When frequency and bandwidth of the \( k \)-th formant candidate obtained by the pitch synchronous covariance LP method are called \( F_{k,i} \) and \( B_{W,i} \), the local cost function for connection between the \( m \)-th formant candidate at the \( (i-1) \)-th frame and the \( k \)-th formant candidate at the \( i \)-th frame is given as follows:

\[ C(i,k,m,n) = W_1 B_{H,i} + W_2 |F_{k,i} - F_{m,i-1}| + W_3 |F_{k,i} - 2F_{m,i-1} + F_{m,i-2}| \]  

where \( B_{H,i} = 0 \)  
\[ = \begin{cases} \log_{10} B_{W,i} & (B_{W,i} < 200) \\ -2 \log_{10} B_{W,i} & (B_{W,i} \geq 200) \end{cases} \]  

The global cost function over the extracted voiced period is the same as in the equation (2) and is also computed using the dynamic programming algorithm [1]. Weighting coefficients \( W_1 \), \( W_2 \) and \( W_3 \) in the equation (5) were set to \( 200, 1 \) and \( 0.1 \) based on preliminary experiment results. The input speech is classified into the following four cases according to the extracted pitch frequency:

a. lower pitch frequency of male voice (\( \leq 150 \text{Hz} \))

b. higher pitch frequency of male voice (\( \leq 200 \text{Hz} \))
c. lower pitch frequency of female voice (<260Hz)
d. higher pitch frequency of female voice (<400Hz)

In order to obtain more exact data, the following three constraints are used in computing the equation (5):

1. Formant frequency range: the frequency range of the first and second formants are restricted according to the extracted pitch frequency. If the frequency of the formant candidate falls outside of the range, a penalty of 300 is added to the cost function.

2. Number of formants: the number of formants is restricted to 5 for the case "a", 4 for the cases "b" and "c", and 3 for the case "d". If the frequency of the formant candidate is greater than a pre-defined threshold \( F_{th} \), \( F_{th} - F_{ws} \) is added to its bandwidth so as to pick up the candidates with lower formant frequency.

3. Order of LP method: the order of LP method is defined as 10 for the cases "a" and "c", 11 for the cases "b" and "d".

This method was examined for a 212 word vocabulary uttered by 3 male and 3 female speakers. At first the pitch-synchronous covariance LP method used in this paper is compared to the auto-correlation LP method. Figure 1 and 2 show the results for a word /iwa/ uttered by a female speaker with 288Hz pitch frequency. Figure 1 shows the result obtained by the auto-correlation LP method. The first formant frequency in the phoneme /a/ was not correctly extracted because of the harmonics of the pitch frequency. Figure 2 shows the result obtained by the covariance LP method. This method correctly extracted the first formant frequency in the phoneme /a/.

Figure 3 and 4 show the first and second formant frequency distribution for the speech data uttered by a female speaker. Variance of the formant frequency obtained by the pitch-synchronous covariance LP method is smaller than that by the auto-correlation LP method.

<table>
<thead>
<tr>
<th>Table 1 Percent correct of pitch frequency extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch frequency extraction method</td>
</tr>
<tr>
<td>Auto-correlation of original waveform</td>
</tr>
<tr>
<td>Auto-correlation of LP residual waveform</td>
</tr>
<tr>
<td>Cepstrum analysis</td>
</tr>
</tbody>
</table>

Figure 1 Formant tracking obtained by the conventional method

Figure 2 Formant tracking obtained by the proposed method
Utilization of phoneme knowledge

As described before, we would like to use this system in building up a speech database which includes the pitch and formant frequency information. Especially when phoneme labeling was carried out by an automatic labeling system or manual inspection, we use knowledge about the kind, context and location of phonemes. In this paper we use the knowledge about the first and second formant frequency at beginning, middle and final frames of the phoneme taking into account the context of the preceding and following phonemes. Based on the knowledge, temporal pattern of the formant frequency is generated using the second order interpolation function. Distance between the frequency of the formant candidates and the formant frequency of the generated pattern is added to the cost function shown in the equation (5). Figure 5 shows the first and second formant frequency distribution for the previously described speech data using the phoneme knowledge. Detection errors of a vowel /i/ were corrected.

Conclusion

This paper shows the effectiveness of various knowledge for the pitch and formant frequency tracking. The knowledge concerning the frequency range and number of the formant predicted by the extracted pitch frequency and the knowledge concerning kind of input phonemes are useful.

Reference

ROLE OF INHERENT VOWEL DURATION IN INTRA-SYLLABLE TIMING

V.D. Kuznetsov

Laboratory of Experimental Phonetics, M. Torez Moscow Institute of Foreign Languages, Osaka 38, Moscow 119034, USSR

ABSTRACT

This study is designed to investigate temporal pattern of CVC syllable in Russian. It is found that the durations of the fricative consonants and the vowel are mutually dependent. On the one hand, there is an inverse relationship between inherent vowel duration (IVD) and the length of the preceding consonant, and on the other hand, the duration of the identical vowels in different contexts correlate positively with the length of the surrounding consonants.

One of the reasons, why there is yet no adequate understanding of the speech timing processes, at any rate as far as Russian is concerned, is the lack of the essential quantitative data on the subject. The purpose of this study is to explore the possible role of IVD in the intra-syllable timing.

METHOD

Speech material consists of Russian CVC syllables in which symmetric consonantal environment is realized either by palatalized or nonpalatalized fricatives /j, ь, y/ and V is one of the 10 allophones /a, a, o, n, y, y, y, w, h/ (cyrillic characters are used to designate the allophones; the dots above the letters indicate the vowel in the palatalized consonantal context). Target syllables are embedded in a carrier sentence "Повтори ... опять" (Say... again). One male speaker read 3 randomized lists of 100 sentences, containing 10 tokens for each of the 10 syllables. Measurements of segment, syllable and sentence durations were carried out on the oscillograms of the speech wave. The boundary between C and V was quiet easily identified by an abrupt drop of the signal amplitude, indicating the switch-over from the noise source to the vocal one. The situation becomes more complex when segmenting the VC sequence. The trouble arises from the fact, that there is an interval where both of the excitation sources are functioning simultaneously: vocal vibration rapidly diminishes and the amplitude of the excitation increases. The boundary between V and C was placed at the point where the first signs of the fricative noise became visible on the oscillogram.

RESULTS

The first part of the data analysis is performed taking into consideration only the characteristics of the consonantal environment, that is, place of articulation and palatalization feature. In the six contexts studied the average duration of the target syllables (N=50) and the corresponding sentences are (in ms):

<table>
<thead>
<tr>
<th>Context</th>
<th>Average Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>353/958</td>
</tr>
<tr>
<td>S'</td>
<td>312/941</td>
</tr>
<tr>
<td>S</td>
<td>348/947</td>
</tr>
<tr>
<td>S'</td>
<td>310/936</td>
</tr>
</tbody>
</table>

From these values it can be seen that the pattern of the durational differences observed at the sentence level reproduces quite closely that one at the syllable level. To find out the possible causes of the differences in the duration of the syllables one must turn to the data on segmental durations. In the contexts studied the mean vowel durations are as follows:
S' = 94  f' = 85  X' = 79
S = 87  f = 83  X = 74.

The mean durations of the initial (1) and final (2) fricatives are:
(1) S' = 139  S = 135  f' = 118  f = 119  X' = 115  X = 108
(2) S' = 120  S = 126  f' = 109  f = 108  X' = 104  X = 106

These data indicate that the syllable duration differences are chiefly due — as it might be expected — to the differences in the length of the consonants. The other source of the differences is the duration of the vowels: on average vowel duration seems to correlate positively with that of the surrounding consonants.

Contrary to the traditional view that palatalized consonants are longer than nonpalatalized ones, in this study the differences between average durations of the fricatives distinguishable from each other only by the value of the palatalization feature are too small and inconsistent to be significant. The same holds true for the mean durations of the vowels uttered in the context of these consonants. It should be noted that in all cases, except the nonpalatalized velar fricative [x], the average duration of the initial consonant is from 10 to 20 ms greater than that of the final one.

A more detailed information on syllable temporal structure is presented in Fig.1, plotting mean segment duration (N=10) as a function of the quality of the vocalic syllable nucleus. From the examination of Fig.1 several points emerge. The general pattern of the IVD-contrasts (the curves marked with "V") remains largely unaffected by changes in the consonantal environment. The data on IVD obtained in the present study are in agreement with the results of our previous research, involving a more representative body of durational measurements [1,2,3].

The changes in the duration of the initial consonant and the vowel tend to be compensatory/complementary; for instance, an increase in IVD is counterbalanced by shortening of the consonantal duration. Thus, the overall length of the CV interval is kept constant. The effect of IVD on the duration of the preceding consonant is more pronounced when the consonant is palatalized (Fig.1, graphs a,c,d,e). As regards the final consonant, its duration varies little and unsystematically.

A step-wise character of the V-curves obtained in the context of the bilabial in marked contrast to the smooth continuity of the other V-curves. The two easily noticeable elbows in the V-curves in question point to the rounded vowels [o,y,o,y]. The shortening of these vowels is accompanied by an increase of the length of the preceding consonant only when the bilabial fricative is nonpalatalized. This strongly suggests that the influence of IVD is overpowered in this case by a co-articulatory effect: as is known, the stronger a sound is co-articulated, the shorter becomes its duration and vice versa.

The temporal relation between the initial and final consonants described above needs to be reexamined. First, the difference between durations of these consonants diminishes as the IVD of the syllable nucleus increases and drops to zero for vowels [a,a,0,0,0]. Second, in the light of the established IVD effect there is no reason to reject the possibility that duration of the final consonant is determined not only by its position in the syllable but by the duration of the following vowel (al), though it belongs to another word.

**DISCUSSION AND CONCLUSIONS**

The temporal interaction discovered between IVD and the length of the preceding consonant looks very much like a local distribution of the time
Fig. 1
interval assigned at the level of motor control programming for articulation of an open syllable. This interpretation supports the so-called comb model, according to which time count for formation of the sequence of time intervals is triggered not by the feedback signals of articulation but by motor command itself. However, within the framework of the comb model one can hardly explain why in the present experiment the mean duration of the CV sequence is different for different consonants. Thus, for example, the mean duration of [G'V] is 233 ms, while that of [X'V] is 194 ms. L.A. Chis-tovitch suggests that time count may be triggered not by a single event but by a group of them, each having different weight [4].

IVD effect is not limited exclusively to Russian. Several researchers, studying temporal pattern of a VCV structure with vocalic environment realized by every possible combination of the vowells [a,i], detected the same effect on consonants of different manner of articulation in English [5,6]. It has been discovered that IVD effect is already present in the speech of 3 year old American children [7].

REFERENCES
4.D
LONG-TERM SPECTRAL DIFFERENCES OF CONTINUOUS SPEECH BETWEEN SPEAKERS OF ENGLISH AND FRENCH PRODUCED NORMALLY OR WITH MASKING NOISE.

W. Weiss
University of Ottawa, Ottawa, Ontario K1N 6N5, Canada.

An acoustic measure of voice quality was proposed by Frokjaer-Jensen and Prytz (1976) as $\alpha = \text{intensity above 1kHz}$.

Wedin et al. (1978) seemed to confirm the utility of this measure with a group that had undergone voice training. However, these experiments were conducted with speakers of Danish and Swedish. We are now interested in finding out how are spectral measurements of voice different for male and female speakers of French and English who had undergone voice training. However, quite possibly the speaking voices of Voice-Trained and Untrained subjects might not be very different when the speaking situation is not one of performance. Since it is known that the presence of noise produces an increase in vocal levels (Lombard 1909), we have undertaken to examine selected spectral areas when the speakers are given masking noise in either the left or right ear. Under these differential auditory feedback conditions vocal changes are unconscious and vocal strategies can be examined functionally in relationship to the same disturbance.

PROCEDURES

46 subjects were recorded while reading the same text for one minute under three conditions: 1. Normal speech (S); 2. With right ear masked with a 75db white noise; 3. With left ear masked with a 75db white noise.

The group was composed of 1. Anglophones and Francophones, 2. Men and Women, 3. Voice Trained and Untrained subjects.

Long Term Average Spectra was applied to all recorded samples and Spectral levels were determined for Fo, F1, <1000Hz (B1K), >1000Hz (A1K); and $\alpha = >1000Hz/\text{<1000Hz}$. Analyses of variance were computed in order to ascertain differences between each pair of the 3 groups.

RESULTS

S = normal Speech

SRM = speech with right ear Masked

SLM = Speech with left ear Masked

* significant at the 0.05 level

** significant at the 0.01 level

F0e: Energy at interval 80-160Hz for men, 160-250Hz for women

F1e: Energy at interval 315-600Hz

B1K: Energy below 800Hz (90-800Hz in 1/3 octaves)

A1K: Energy above 1000Hz (1000-5000Hz in 1/3 octaves)
<table>
<thead>
<tr>
<th>Table I: Mean energy levels (Dbv) of ANGLOPHONES (N=23) and FRANCOPHONES (N=23) for three speech production conditions measured over selected (1/3 octave) intervals.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech condition</td>
</tr>
<tr>
<td>Normal Speech ANGLO: -22.65 -22.22 0.43 -18.49 -28.54 -10.04</td>
</tr>
<tr>
<td>(S)  Franco: -21.36 -21.35 0.00 -17.26 -28.86 -11.59  **</td>
</tr>
<tr>
<td>Speech with right A: -21.08 -18.37 2.71 -15.66 -23.18 -7.52  **</td>
</tr>
<tr>
<td>Speech with left A: -22.10 -18.90 3.20 -16.41 -24.31 -7.89  **</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table II: Mean energy levels (Dbv) of ANGLOPHONES MALE (N=10) and FRANCOPHONES MALE (N=12) for three speech production conditions measured over selected (1/3 octave) intervals.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech condition</td>
</tr>
<tr>
<td>Normal Speech ANGMA: -22.21 -19.40 2.81 -16.23 -25.64 -9.41  **</td>
</tr>
<tr>
<td>(S)  FRANMA: -21.92 -20.27 1.64 -16.42 -27.58 -11.16  **</td>
</tr>
<tr>
<td>Speech with right A: -20.59 -15.77 4.81 -13.25 -20.53 -7.27  **</td>
</tr>
<tr>
<td>ear masked (SRM) F: -20.50 -16.73 3.76 -13.86 -23.47 -9.61  **</td>
</tr>
<tr>
<td>Speech with left A: -21.72 -16.54 5.18 -14.27 -22.32 -8.05  **</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III: Mean energy levels (Dbv) of ANGLOPHONES FEMALE (N=13) and FRANCOPHONES FEMALE (N=11) for three speech production conditions measured over selected (1/3 octave) intervals.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech condition</td>
</tr>
<tr>
<td>(S)  FRANFEM: -20.74 -22.52 -1.77 -18.18 -30.25 -12.07  **</td>
</tr>
<tr>
<td>Speech with right A: -21.46 -20.37 1.09 -17.52 -25.22 -7.70  **</td>
</tr>
<tr>
<td>ear masked (SRM) F: -18.89 -19.01 -0.12 -15.65 -25.36 -9.70  **</td>
</tr>
<tr>
<td>Speech with left A: -22.39 -20.72 1.67 -18.06 -25.84 -7.77  **</td>
</tr>
</tbody>
</table>
**Table IV**: Mean energy levels (Dbv) of Anglophones trained (N=11) and Francophones trained (N=12) for three speech production conditions measured over selected (1/3 octave) intervals.

<table>
<thead>
<tr>
<th>Interval</th>
<th>FOe</th>
<th>Fle</th>
<th>SFIF0</th>
<th>BLK</th>
<th>A1K</th>
<th>a1AB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speech condition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANGST: -21.73</td>
<td>-20.99</td>
<td>0.74</td>
<td>-17.48</td>
<td>-25.93</td>
<td>-8.45</td>
<td>**</td>
</tr>
<tr>
<td>(S) PFAT: -21.73</td>
<td>-22.14</td>
<td>-0.41</td>
<td>-17.93</td>
<td>-29.61</td>
<td>-11.67</td>
<td></td>
</tr>
<tr>
<td>Speech with right ear masked (SRM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: -20.32</td>
<td>-17.67</td>
<td>2.64</td>
<td>-14.95</td>
<td>-21.93</td>
<td>-6.97</td>
<td>**</td>
</tr>
<tr>
<td>Speech with left ear masked (SLM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: -21.63</td>
<td>-18.35</td>
<td>3.27</td>
<td>-15.94</td>
<td>-23.48</td>
<td>-7.53</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table V**: Mean energy levels (Dbv) of Anglophones untrained (N=12) and Francophones untrained (N=11) for three speech production conditions measured over selected (1/3 octave) intervals.

<table>
<thead>
<tr>
<th>Interval</th>
<th>FOe</th>
<th>Fle</th>
<th>SFIF0</th>
<th>BLK</th>
<th>A1K</th>
<th>a1AB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speech condition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANGCU: -23.50</td>
<td>-23.34</td>
<td>0.15</td>
<td>-19.42</td>
<td>-30.94</td>
<td>-11.51</td>
<td>**</td>
</tr>
<tr>
<td>(S) FRAU: -20.96</td>
<td>-20.48</td>
<td>0.47</td>
<td>-16.26</td>
<td>-28.04</td>
<td>-11.51</td>
<td>**</td>
</tr>
<tr>
<td>Speech with right ear masked (SRM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: -21.78</td>
<td>-19.00</td>
<td>2.77</td>
<td>-16.32</td>
<td>-24.34</td>
<td>-8.01</td>
<td>**</td>
</tr>
<tr>
<td>Speech with left ear masked (SLM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: -22.54</td>
<td>-19.40</td>
<td>3.13</td>
<td>-16.84</td>
<td>-25.07</td>
<td>-6.23</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Anglophones and Francophones (Table I)

Anglophones have higher \( \alpha \) significant only in the masked conditions.

Anglophones masculin and Francophones masculin (Table II)

Anglophones have higher \( \alpha \) significant only in one of the masked conditions.

Anglophones feminin and Francophones feminin (Table III)

Anglophones have higher \( \alpha \) significant only in the masked conditions. Francophones have higher FOe in the three speech conditions and a higher BLK in the Speech condition.

Anglophones trained and Francophones trained (Table IV)

Anglophones have higher \( \alpha \) than Francophones in the three speech conditions. Furthermore they have more spectral energy above 1000Hz (A1K) in the Speech condition and in one of the masked conditions.
ANGLOPHONES UNTRAINED AND FRANCOPHONES Untrained (Table V)

Anglophones show higher $\alpha$ in only one of the Masked conditions. Francophones have higher F0e and Blk in the three Speech conditions and higher P1e in the Masked conditions.

DISCUSSION

Higher "Quality" ($\alpha$) for Anglophones means that this group has more spectral energy above 1000Hz relative to below 1000Hz when compared to Francophones. Interestingly, the distinctions reach a level of significance in the more intense rendering of the masked conditions (Table I). That means that the measure is also functional. That is, anglophones have greater possibilities in the upper frequencies when departing from low level conversational speech. This is probably a result of a larger jaw opening and a greater distance between the tongue and the palate in the English setting (Laver 1980). Given that the ear is more sensitive in the upper frequencies, Anglophones could be heard with lower vocal levels than their Francophone counterparts. This can be seen with Untrained subjects (Table V) who show higher levels below 1000Hz where most of the vocal energy is concentrated. Because of higher spectral levels in F0e voice quality of Untrained Francophones is probably breathier (Izdebski, 1980). Interestingly, in the Speech condition $\alpha$ of the two groups are identical. Whereas with training vocal levels are similar except above 1000Hz where Anglophones have significantly higher levels (Table IV). Hence, the $\alpha$ of Trained Anglophones is significantly higher in the three Speech conditions. Thus, $\alpha$ differences depends on both language and training. Whereas differences below 1000Hz is language and sex dependent Francophone Feminin exhibiting the highest levels (Table III).

REFERENCES


A CHINESE SPEECH SYNTHESIS SYSTEM

Yu Chongzhi, Gao Feng, Zhu Longming

Institute of Acoustic
Nanjing University
Nanjing, 210008, China

Abstract
A real-time Chinese speech synthesis system that bases on a line spectrum pairs (LSP), the frequency domain parameter of linear prediction analysis, is proposed. The system adopts the Chinese Shangmu and Yumu as basic synthesis units. The control of tone, intensity, and duration of syllables, the transition between Shangmu and Yumu, have been discussed. The system has been developed in micro-computer with a TMS32010 signal processing chip for real time synthesizer. The quality of synthesis speech is close to natural.

I. Introduction
In recent years, the full Chinese syllable synthesis based on the study of forment or PARCOR coefficients of linear prediction has been studied in some universities and institutes. But it still need further study to improve the quality of synthesis speech as articulation, naturalness, etc. We select the line spectrum pairs (LSP) as the vocal tract parameter for speech synthesis. Since it has some advantages over other linear prediction parameters in the properties of interpolation and quantization. It is a frequency-domain parameter closely related with forment parameters and can be linearly interpolated to form smooth transition. As a result, we can get the synthesis speech with good quality by very simple control rules and the real-time synthesis can be carried out by TMS32010 easily. The aims, which we develop the full Chinese syllables synthesis system, are studying the rules Chinese speech synthesis, and try to perform the text-to-speech devices for some applications, such as talking aids for vocal handicapped, reading aids for blind and the computer voice output device.

II. LSP Synthesis Method
LSP parameter is another set of linear prediction parameter. It can be derived by lattice PARCOR analyzes method that

\[ A_P(z) = A_p(z) - k_{pi}B_p(z) \]  
\[ A_p(z) = 1 + \sum_{i=1}^{\infty} \alpha_i z^{-i} \]  
\[ B_p(z) = z^{n_{pi}} A_p(z^{-1}) \]

where \( 1/A_p(z) \) is a stable all-pole digital filter of speech signal and \( B_p(z) \) is defined in equation (3), \( K_{pi} \) is the reflected coefficient or PARCOR coefficient. Let \( K_{pi} = \pm 1 \), a pair of \( (p+1) \)th order polynomials \( P(z) \) and \( Q(z) \) are defined as
\[ P(z) = Ap(z) - Bp(z) \]  

(4)

\[ Q(z) = Ap(z) + Bp(z) \]  

(5)

By introducing a rational function, it is shown that all the zeroes of \( P(z) \) and \( Q(z) \) lie on the unit circle alternately. Both \( P(z) \) and \( Q(z) \) have real coefficients only. Therefore, if \( e^{i\omega} \) is one of their zeroes, \( e^{-i\omega} \) is also one. The following expressions of \( P(z) \) and \( Q(z) \) are represented by parameters \( (\omega_i, \delta_i) \). For even \( p \)

\[ P(z) = (1 - z^{-2})^{\sum_{i=1}^{p_1} (1 - 2\cos\omega_i z^{-1} + z^{-2})} \]  

(6)

\[ Q(z) = (1 + z^{-2})^{\sum_{i=1}^{p_2} (1 - 2\cos\delta_i z^{-1} + z^{-2})} \]  

(7)

From eqs. (4)-(5), \( A(z) \) is expressed in terms of \( P(z) \) and \( Q(z) \) as

\[ A(z) = \frac{P(z) + Q(z)}{2} \]  

(8)

Consequently, a set of parameters \( (\omega_i, \delta_i) \) is derived from \( Ap(z) \), and \( Ap(z) \) is also reconstructed in terms of \( (\omega_i, \delta_i) \). The LSP parameters \( (\omega_i, \delta_i) \) may be converted to \( (f_i, g_i) \) whose dimension is in kHz via the relation \( f_i = \frac{\omega_i}{2\pi}, \quad g_i = \frac{\delta_i}{2\pi} \). The transfer function of LSP all-pole filter is identical to PARCOR all-pole filter.

\[ H(z) = Q/A(z) = Q/[1 + (A(z) - 1)] \]  

(9)

The expression of \( Ap(z) - 1 \) is obtained directly by using the LSP parameters. For \( p \) even

\[ Ap(z)-1=\{[P(z) - 1] + [Q(z) - 1]\}/2 \]  

\[ =\{(a_1 + z^{-1})^{\sum_{j=1}^{p_2} (1 + a_j z^{-1} + z^{-2})} - \sum_{j=1}^{p_2} (1 + a_j z^{-1} + z^{-2})\}/2 \]  

(10)

where \( a_i = 2\cos\omega_i, \quad b_i = 2\cos\delta_i \).

The flow graph of LSP speech synthesizer for \( p \) even is shown in the Fig.1. The power spectrum \( S(\omega) \) relates to the LSP all-pole filter on the basis of eqs. (6), (7) and (8) as

\[ S(\omega) = \frac{|A(e^{i\omega})|^2}{|Ap(e^{i\omega})|} \]  

Fig.1 The flow of LSP synthesis

From eqs. (10) and (11), assuming that \( f_i \) and \( g_i \) are close to each other, then \( S(\omega) \) becomes large. Conversely, a necessary condition for \( S(\omega) \) to have a strong resonance at a certain frequency band is that more than one LSP parameter should concentrate.

III. System Construction
The system is composed of four parts as shown in Fig(2)

```
PHONETIC LIB.

KEYBOARD ───→ SYNTHESIS CONTROL ───→ SYNTHESIZER ───→ D/A ───→
```

Fig.2 The diagram of synthesis system

1. Keyboard for inputting the Chinese Pin-Yin and its tone pattern (0, 1, 2, 3, or 4).
2. Phonetic data library and rules of synthesis.
3. LSP synthesizer with a THS32010 signal processing chip.
4. D/A conversation and low-pass filter to loudspeaker.

IV. Phonetic and Control Rules

The basic synthesis units are the Chinese Shengmu and Yunmu, to be similar to consonant and vowel, which are segmented from natural syllables uttered by a female speaker, digitized at 8KHz and analyzed by 12-order LSP. Analyzing carefully the transition between Shengmu and Yunmu on the LSP trajectory, we found that some Shengmu always and at a stable point [Fig.3] then to connect with Yunmu and selected the allo- Shengmu based on the conclusion. There are about 70 allo-Shengmu and 39 basic Yunmu adopted in the system for high quality synthesis speech. Tone patterns (pitch and energy envelop), Shengmu and Yunmu (length, LSP and resonant) are stored in phonetic library as the basic parameters.

Usually, a syllable contains three parts which are Shengmu, Yunmu and transition between Shengmu and Yunmu. The length of Yunmu is the main part in duration of a syllable, which is controlled by tone pattern in an isolated syllable and changed under some rules when a faster synthesized speech is desirable, but while other parts are unchanged. The transition that means how to connect the Shengmu and Yunmu is very important in speech intelligibility and naturalness in the system. Its length is determined by the concatenation of Shengmu and Yunmu, is about 10 usec (or 1-3 frames). Since the LSP is a frequency domain parameter, the linear interpolation can be used for parameter transition and the result show that it is satisfactory for speech quality. The intensities of Shengmu and Yunmu are controlled separately according to tone pattern in a syllable.

V. Conclusion

On the LSP theory, we set up a real-time synthesizer in the micro-computer with a THS32010 signal processing chip and obtained better synthesis speech quality. Specially, an isolated syllable quality is close to natural because the linear interpolation method is adopted in the transition between Shengmu and Yunmu. Similarly, the linear interpolation method can be adopted in the connection between two syllables. For the aim of long synthesis speech (as a sentence) closing to natural, further research need to be done.
Fig.3 Shengmu /ch/ end at a stable point

VI. Reference
EXPERIMENTAL DETERMINATION OF ARRAY MINIMAL SIZE FOR HULL FAR FIELD COMPUTATION FROM NEARFIELD MEASUREMENTS.

Christophe Journeau, Thierry Rohan & Jean Pierre Pasqualini

CERDAN BP 77 83800 Toulon Naval, France
PRINCIPIA Recherche Développement, Place S. Laffitte, 06560 Valbonne, France
SYMINEX, 2 Boulevard de l'Océan, 13275 Marseille Cedex 9, France

Introduction

Many problems arise when one wants to measure directly the farfield of a large radiating structures: it requires large volumes of water and great distances and often implies nonidentical boundary of the environment; large distances also mean low signals and thus bad signal to noise ratio. Awareness of these problems has stimulated interest in efforts to determine farfield characteristics from data obtained in the nearfield of the source.

The aim of our study is to determine the minimal size of the planar measurement array needed to compute accurately the farfield from nearfield data. Bobber [1] stated that a Trotter array height and width must be twice that of the source studied. Yaghjian [2] derived an a posteriori criterion based on the ratio of the radiation due to the array boundary to the total radiation.

Since our approach is mainly experimental we will firstly describe the measurements we made. Then we will present the farfield computation technique we used and the results we obtained for various array sizes. In order to examine more thoroughly the physical nature of this problem, a wavenumber spectrum analysis will be shown. Finally an extrapolation technique will be tested and criteria presented.

Experimentation

A thorough experimental study of the nearfield of a vibrating hull has been made by CERDAN. It is described in reference [3].

Our test structure was a cylindrical hull 3 m long and 0.6 m in radius immersed 7 m deep under our test barge. It was excited by a shaker mounted on an internal lattice structure connected to the hull through more than 20 points.

To measure the acoustic nearfield, we used a 5 m long vertical array made of 60 hydrophones 86 nm apart. This line array was shifted in translation in order to describe a 4.3 m wide rectangular mesh of 32 lines (cf Figure 1).

The far field was finally measured.

All these data were processed through an LMS - HP 1000 system and Fourier transformed with a frequency range of 0-7500 Hz and an increment of 17.6 Hz.

Planar Wave Spectrum extrapolation technique

To predict the farfield from nearfield data, we chose - among many techniques - to use the plane wave spectrum representation of fields which is exposed in the papers of Wang [4] or Maynard et al. [5].

The pressure being known on an infinite planar surface, the Helmholtz equation enables us...
to compute the pressure at point M using the following equation:

$$p(x,y,z) = \frac{1}{4\pi} \int \int_{S(M')} \frac{\partial G(M,M')}{\partial n} \, ds(M')$$

(1)

where $G$ is the Green’s function satisfying the homogeneous Dirichlet condition on the measurement plane. If we take the two-dimensional Fourier transform of (1) we obtain:

$$P^*(k_x,k_y,z) = P^*(k_x,k_y,z_0) \tilde{G}^*(k_x,k_y)$$

(2)

The star denotes the two dimensional Fourier transform defined by:

$$P^*(k_x,k_y,z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x) \, e^{j(k_x x + k_y y)} \, dx \, dy$$

(3)

The two-dimensional Fourier transform of the normal derivative of the Green’s function $G$ is such that (2) becomes:

$$P^*(k_x,k_y,z) = e^{-j k_z z_0} P^*(k_x,k_y,z_0)$$

where $k_z = \sqrt{k_x^2 + k_y^2}$

(4)

The farfield pressure can be derived in spherical coordinates $r, \theta, \phi$ with the large $r$ approximation by the following formula:

$$P_\infty(r,\theta,\phi) = \frac{e^{j k_0 r}}{r} \tilde{p}(k_x,k_y)$$

(5)

So, each Fourier component $P^*(k_x,k_y)$ where $k_x^2 + k_y^2 < k_0^2$ correspond to a uniform plane wave propagating in the direction $\theta, \phi$.

**Results:**

Figure 2 presents the nearfield which was measured on the planar array at 2600 Hz. The position of the hull is shown by the central rectangle. This plot highlight the complexity of the acoustic nearfield.

Figure 3 present the comparison of the measured farfield and of the computed farfield for various heights of the array.

From the study of many of these plots, we deduced that an height of 2 m (1.75 diameter) is enough to compute farfields in the horizontal plane with 3dB errors.

**Wavenumber space analysis.**

In order to understand the physical phenomena, we decided to study the measured field in the wavenumber space. Figure 4 represent the two-dimensional Fourier transform of the nearfield shown on figure 2. The central ellips represent the radiation circle (the scales being slightly different for the $k_x$ and $k_y$ axes).

We can remark that the levels outside the radiation circle are very small even though the distance between the hull and the plane is of about 0.4 wavelength.

The values inside the radiation circle represent the far field values in the corresponding direction divided by $\cos\theta$.
Now we mask by zeros the measurements made in the outer part of the array letting only a 2.1 m x 2.1 m square array (cf Figure 5). We have also calculated its Fourier transform which is quite different with the one we obtained with the complete array. For some farfield directions we observe more than 6 dB difference.

**Extrapolation technique**

In order to obtain good farfield predictions with a minimal set of sensors, we experimented a technique dedicated to extrapolate the nearfield in the areas where we had no sensors. Our assumption was that the radiation can be considered as almost axisymmetrical. Candel et al [6] proved that in this case we can represent the propagation of the acoustic waves with zeroth order Hankel function as propagator. So using this propagator, we reconstructed the acoustic field created by the hull on the missing points of our initial array. This enabled us to obtain a good agreement with the first computation with a much lower number of sensors.

These results can be seen in Figure 7. We have already made the one-dimensional Fourier transform along the vertical y-axis. We consider the ky=0 component (i.e. the sum of the pressure along each
gating in the horizontal plane. On the first graph we represent the value measured on the initial array and on the smaller array with and without extrapolation. The second graph present the Fourier transform of the first one. The triangles indicate the acoustic wavenumber $K_n$. We can note that the predictions are far better with the extrapolation tech-

Using this extrapolation technique we find that with an array of the size of the cylindrical hull it is possible to compute the farfield within 3 dB.

Summary

We have studied the nearfield of a submersed cylindrical hull in order to predict the farfield radiated noise. Using the Planar Wave Spectrum extrapolation technique it has been possible to obtain very good agreements as long as the spacing between the sensors respects the Nyquist criterion (i.e. is less than half a wavelength) and the size of the array is 1.75 times the size of the radiator.

Using a wavenumber space analysis to investigate a Hankel extrapolation technique, we found that if the array is of the same size than the radiator we can extrapolate the missing sensors and then use the PWS technique to obtain results with less than 3 dB error.

References

[6] SM Candel, C Chassaignon Radial extrapolation of wave fields ... JASA 76(6) p 1823 (12/84)
A COMPARISON OF A MAGNITUDE ESTIMATION TECHNIQUE AND A PAIR COMPARISONS TECHNIQUE FOR ASSESSING TEXT-TO-SPEECH SYNTHESIS

Chaslav V. Pavlovic, Christel Sorin, Jean Pierre Rouaglière and Jean Pierre Lucas.

Centre National d'Etudes des Télécommunications, BP 40, 22301 Lannion, France and (first author only) Department of Speech Pathology and Audiology, University of Iowa, Iowa City, Iowa 52242, U.S.A.

Introduction

This study was designed to cross-validate the magnitude estimation (ME) procedure and the pair comparison (PC) procedure for assessing the quality of text-to-speech synthesis. To this goal the psychophysical scale values of the speech stimuli obtained from the ME procedure were compared to the values obtained from the pair comparisons (PC) procedure.

In the PC procedure the subject indicates his preference for one of two stimuli. The ME procedure requires the subject to make direct numerical estimations of the sensory magnitudes produced by different stimuli.

Methodology

Sixteen subjects participated in the PC experiment and another 16 subjects participated in the ME experiment. All subjects had normal hearing acuity. Four different synthesis procedures and three different prosodic rules were used to generate the stimuli. Some 12 sentences were generated with each of the 12 combinations of the synthesis procedures and prosodies. However, only five sentences were employed for any given combination of the synthesizer and prosody. The long-term average rms levels of all sentences were adjusted to be equal.

The synthesis procedures, all by French diphones, were developed at "the Centre National d'Etudes des Télécommunications" (CNET) in France. They are known as the LPC, MPLPC, FFT, and KDG. These terms will be retained in this study.

Two of the prosodies were generated automatically from text. One of them (prosody "L") uses the rules designed for simulating the intonation contours observed during reading. The other prosody (prosody "P") uses a different set of rules elaborated for simulating the prosody used by radio and TV announcers. Both prosodies use the same syntactico-prosodic parsing of the text. For the third prosody (prosody "N"), the duration and F0 values used for synthesis were, for each sentence, manually extracted from the direct recordings of these sentences by a male speaker and further applied to the concatenated strings of diphones.

A stimulus was defined as a combination of one text sentence, one synthesizer and one prosody. The signals were delivered binaurally. In the PC experiment, the comparison pairs were restricted to stimuli having the same type of prosody. With this restriction, all other possible A/B and B/A pairs of stimuli were presented in random orders. In the ME experiment all possible stimuli were presented in random orders. In both experiments the subjects were instructed to judge the "quality" of the system that produced the sentence.
Results

Both experiments resulted in multiple observations for each prosody-synthesizer combination. For the ME experiment these data were reduced to one value for a given prosody-synthesizer combination in two different ways. The "absolute" results were calculated by finding the geometric mean across subjects and synthesizer-prosody repetitions. To obtain the "relative" results the raw results were first normalized by dividing each individual listener estimation by the geometric mean of all that individual's estimations. Then, the result for a given synthesizer-prosody combination was calculated as the arithmetic mean across subjects and prosody-synthesizer repetitions. Because the correlation between the absolute and relative results was very high (0.99%) only relative results are discussed further in the paper. These results are graphically represented in Fig. 1. The curve with the squares refers to prosody L, the curve with pluses to prosody P, and the curve with the diamonds to prosody N. An analysis of variance indicated that both the effect of the synthesis procedure and of the prosody were statistically significant. No interaction between these effects was evident. The absence of an interaction was further confirmed by very high correlation coefficients among the results obtained with different prosodies. The linear correlation coefficients on the group results between the prosodies L and P, L and N, and P and N were, respectively, 0.93, 0.98, and 0.99.

A post-hoc t-test adjusted for multiple comparisons indicated that system LPC is worse than any of the others and that system MPLPC is worse than system FFT. However, no significant differences were demonstrated between system KDG on one hand, and either system FFT or MPLPC on the other. In respect to prosodies, the test indicates that prosody P is of lower quality than either prosody L or N, and that prosody L is of lower quality than prosody N.

To obtain the PC scale values, first, the proportion of time one synthesis system was preferred over another was calculated for each of the prosodies. The psychological scale values of the stimuli were obtained from these proportions (p) according to two different models: the Thurstone Case V model, and the Bradley-Terry-Luce, or BTL, model. According to the former, with the assumption of equal and uncorrelated variances of two stimuli on the psychological scale, their psychological distance is proportional to the z-value that corresponds to the proportion p. According to the latter, the psychological distance is given as 

$-\ln(1-p)/p$. For both models the scale value of a stimulus was calculated as the mean distance from the other stimuli. These values are on an interval scale only. Therefore, the true zero does not exist, and the values could be subjected to a linear transformation. The linear correlation coefficient between the values resulting from the two models is extremely high (0.999). Therefore, only the values of one of the models (BTL) will be used further in the paper.

In the PC experiment comparisons between different prosodies were not made. Therefore, it was not possible to evaluate different prosodies relative to one another. The scaled values of the stimuli are represented by squares in Fig. 2 (prosody L), Fig. 3 (prosody P), and Fig. 4 (prosody N). These are the original values after they were transformed in a linear fashion so that the results agree with those of the ME experiment for the worst (LPC) and the best (FFT) system. The results of the ME experiments are also indicated (symbols "+").
Discussion

Both methods indicate that the systems rank LPC - MPLPC - FFT from low to high, as demonstrated in Figs. 2 - 4. The only slight exception to this is the ranking for prosody P, in the PC experiment, where systems MPLPC and FFT were rated equally.

The mean ratings across prosodies obtained by the two methods (adjusted to agree at the extreme points as discussed earlier) are given in Fig. 5. Again, the squares refer to the ME experiment and the pluses to the PC experiment. The agreement between the methods, particularly in the ordinal sense, appears good. It appears though (Figs. 2 - 5), that the PC intervals among the three best rated systems, in respect to their distances from the worst system, are somewhat smaller than it is the case with the ME intervals.

Both procedures seem to indicate that system KDG falls between systems MPLPC and FFT (Fig. 5). However, the differences KDG-MPLPC and KDG-FFT appear small in comparisons with the variance in the results. Thus in the ME experiment the differences do not reach the 0.05 level of statistical significance. In the PC experiment the ranking of system KDG is variable: best for prosody P (Fig. 3), second best for prosody N (Fig. 4) and across prosodies (Fig. 5), and third best for prosody L (Fig. 2).

S.S. Stevens divided perceptual criteria into two general classes. Continua on which the discrimination is mediated by an additive physiological process he termed "prothetic." Continua on which discrimination is mediated by a substitutive physiological process he termed "methatetic." Essentially, the former continua is quantitative, while the latter is qualitative. For the methatetic continua, the PC scale is linearly related to the ME scale, while for the prothetic continua the PC scale is linearly related to the logarithms of the ME scale.

The above analysis was performed assuming the methatetic character of speech quality. It would be of interest to compare the ME and PC results assuming the prothetic character of the speech quality continua. However, the stimuli used in this study were all relatively close to each other, so that they fell on a relatively linear part of the logarithmic curve. Thus the correlations between the PC values and the ME values are almost identical to those obtained between the PC values and the logarithms of the ME values. Depending on the prosody, the former correlations are within the interval 0.955 to 0.997, while the latter are within the interval 0.974 to 0.998. Therefore, the results of this study change only minimally if a prothetic nature of the speech quality continua is hypothesized.

A final comment concerns the relative efficiency of the two methods. The mean lengths of the two procedures were approximately equal (roughly 40 min). However, the amount of information provided by the ME experiment was substantially larger (comparisons between prosodies were possible). To yield the same amount of information, and retaining the present round robin tournament structure, the PC procedure would have had 3.8 times more comparisons, and thus would have been 3.8 times longer than the ME procedure.

In conclusion it seems that there is a good general agreement between the results of the two experiments. Therefore, the two methods appear to validate each other.

477
Acknowledgements

This research was made possible by a grant from the EEC Esprit project.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5
MAGNITUDE ESTIMATION TECHNIQUE IN EVALUATING TEXT-TO-SPEECH SYNTHESIS

Mario Rossi, Chaslav Pavlovic, and Robert Espesser

LA 261, CNRS, Parole et Language, Institut de Phonétique, Université de Provence, 13621 Aix en Provence, France. (2nd author also at University of Iowa, Iowa City, Iowa 52242, USA).

Introduction

As text-to-speech systems develop it becomes necessary to evaluate whether a change in the synthesis procedure has an effect on the listener's attitude toward the system. We explore here the possibility of directly scaling intelligibility and user's satisfaction (i.e. acceptability) with the magnitude estimation (ME) technique.

Subjects in ME experiments make direct numerical estimations of the sensory magnitudes produced by different stimuli. We investigate if these judgments vary with the stimulus set size and range, or if they depend on the subject's familiarity with the test material. We further investigate whether the ME scales are practice invariant. Finally, we evaluate the relationship between the "objective" measures of speech understanding (proportion of words understood) and the "subjective" measures (MEs).

Method

Ninety eight college students participated as subjects. For obtaining MEs, sentences of between five and eleven words were randomly extracted from a novel. These sentences will be referred to as "meaningful" or semantically-correct sentences. The same sentences were also used for testing speech recognition. However, because sentences have a high degree of redundancy the speech recognition scores may be subject to a ceiling effect. To insure that even in this case an objective intelligibility measure was available, sentences "without meaning" (semantically-anomalous sentences) were created and also used. They had nearly the same syntactic structure as the semantically-correct sentences.

Seven different text-to-speech synthesis procedures (called in this study "systems") for the French language were used. Five of them were defined by using the diphone synthesis algorithm developed at the "Centre National d'Etude de Telecommunication". The parameters that were varied were the sampling frequency (10 kHz or 16 kHz), prosody (one used in reading or one used by TV announcers), and voice (male or female). These systems are referred to as the CNET group. The differences between the stimuli provided by these systems range from very small (systems differing in one parameter only) to reasonably large (systems differing in all three parameters). In order to determine whether the rating of a system depends on its distances from other stimuli used, the remaining two systems were chosen to be substantially different from the CNET group and from each other. The intention was to see if their presence or absence from the stimulus set affected the ratings. These two systems were the system Infovox SA 201/PC and the system L'icophone.

The speech was presented in three conditions of distortion in order to vary its comprehension from "easy" to "medium" and to "difficult." The first type of distortion (D1) was in effect represented by the internal imperfection of the synthesis system itself, and was thus the speech in quite condition. The second distortion type (D2) and third distortion
type (D3) were represented by different band-pass noises. The long-term-
average rms values of the seven synthesizers were adjusted to be equal to
one another. The signal was delivered to a pair of TDH-39 earphones. The
frequency response of the system was orthophone, and the overall
level of speech corresponded to normal conversational level.

The subjects were divided into seven groups of 14 individuals, with
each group assigned a different task (Table 1). That is, the subjects
from one group judged only intelligibility, the subjects from another
group judged only acceptability, etc. In this way the carry-over from one
type of judgment to another is prevented. Group 1 (tested with
meaningless sentences) and group 2 (tested with meaningful sentences)
were used to obtain objective speech intelligibility scores. Group 3 gave
MEs of intelligibility. These three groups provided data for comparing
different speech intelligibility measures. Group 4 gave MEs of
acceptability. All groups discussed thus far were presented with the
complete set of experimental conditions. That is, the subjects listened
to all combinations of seven synthesizers and three distortions (21
conditions in the total). They were all exposed to an equal but
relatively short practice period ("nominal practice"). Like group 4,
groups 5, 6, and 7 gave MEs of acceptability. However, they differed to
group 4 in that they were presented with only eight experimental
conditions. These conditions consisted of combinations of the two smaller
distortions (D1 and D2) and four systems of the CNET group. Therefore,
both the stimulus set size and range for these groups were substantially
smaller than those of group 4. In other respects group 5 did not differ
from group 4; thus the differences in their results, if any, could be
attributed to the effects of the stimulus set size and range. Group 6
differed from group 5 in that the subjects in group 6 obtained an
extensive practice in all test conditions. Group 7 was different from
group 6 in that the subjects in group 7 were tested with the familiar test
material; i.e., with the same sentences used in the extensive practice
period. Thus the differences between groups 5 and 6, and 7 and 8 should
yield information on, respectively, the effect of practice and the
familiarity with the test material. In the test and nominal practice
phases of the experiment for each condition two sentences were presented.
For all groups, in the test phase the sentences did not repeat, and, with
the exception of group 7, the subject heard them for the first time.

In the speech recognition tasks the subject were instructed to repeat
as many words as possible. In the ME tasks the subjects were instructed
to assign a number to either the intelligibility of the speech or to their
satisfaction with the communication situation. In case of the latter they
were instructed to imagine that they are listening to a story over the
radio.

Results and discussion

The objective intelligibility tests were scored as proportion of
words repeated correctly. The proportion of correct responses for a
condition was calculated by finding the arithmetic mean across subjects
and condition repetitions. In the case of magnitude estimations the
results were calculated as "absolute," or "relative." The absolute value
for a condition was calculated by finding the geometric mean across
subjects and condition repetitions. To obtain the relative value, the raw
estimations of each individual were first divided by the geometric mean of
all that individual's estimations. The arithmetic mean of these normalized
estimations across subjects and condition repetitions represented the relative value of the condition. The correlation between the relative and absolute judgments was high (0.99).

Fig. 1 illustrates the correlation between the speech intelligibility of the semantically-correct sentences and the speech intelligibility of the semantically-anomalous sentences. The linear regression line is also inserted. The high correlation (0.95) suggests that the speech synthesis does not influence listeners' semantic processing abilities independent of the influences already present at the segmental and prosodic levels. Thus meaningless but syntactically correct sentences could be used for testing global speech intelligibility. In easy communication situations they are less subject to the ceiling effect than meaningful ones.

Fig. 2 illustrates the linear correlation (0.93) between the proportion of words repeated correctly on the meaningless sentences and the relative MEs of intelligibility. Likewise, Fig. 3 depicts the linear correlation (0.90) between the proportion of words repeated correctly on the meaningful sentences and the MEs of intelligibility. Considering that the objective scores were subject to the 100% ceiling effect, these correlations could be considered very high. Thus it appears that the MEs are valid measures of speech intelligibility.

Table II is the correlation matrix between the relative acceptability estimates of the seven synthesis systems in the three different conditions of distortion. Apparently, the correlations are small. The acceptability of the systems changes dramatically with communication conditions. Thus the system should be evaluated for any given situation of use.

The difference between groups 4 and 5 was in the set size and range, between groups 5 and 6 in the amount of practice, and between 6 and 7 in the familiarity with the test material. The mean absolute results of groups 4 to 7 were, respectively, 8.3, 9.3, 8.5, and 6.5. An analysis of variance indicated that the differences between groups 4 and 5, and between 5 and 6 did not reach statistical significance. The difference between groups 6 and 7 was significant. Therefore, neither the groups' size/range effect nor the practice effect was demonstrated. However, there seems to exist an effect of familiarity with the test material.

In this analysis a hypothesis was made that the absolute judgments are on an absolute scale, i.e., that different groups of subjects give the same mean numbers for the same stimuli. This is indeed the case with the loudness judgments. To explore the questions of group size/range, practice, and familiarity effect, without confounding the analysis with this hypothesis, the relative results are analyzed. The correlation coefficients between groups 4 and 5, groups 5 and 6, and groups 6 and 7, are respectively 0.97, 0.96, and 0.76. These values do not represent a statistically homogeneous group of correlations. The high correlations between groups 4 and 5, and groups 5 and 6, suggest again the absence of the size/range and practice effect. The only moderate correlation between groups 6 and 7 indicates, as do the absolute results, that familiarity with the test material is important for the judgments of acceptability. Therefore, if the same listeners are used repeatedly in testing synthesizers, it is necessary to create a mechanism for generating a quasi-infinite pool of test material.

Acknowledgments

This research was made possible by a grant from the EEC Esprit project.
Table 1: Description of experimental groups

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A - stimulus recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task B - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task C - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task D - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task E - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task F - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task G - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task H - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2:

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task B - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task C - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task D - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task E - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task F - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task G - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Task H - recognition</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Figure 1

Figure 2

Figure 3
ACOUSTIC STRUCTURE OF HINDI WORD RHYTHMICS
(ANALYSIS, SYNTHESIS, PERCEPTION)

Irina M. Roumyntseva

Institute of Linguistics, Academy of Science,
Semenko 1/12, Moscow, USSR

The main purpose of this experimental study was the investigation of Hindi
word prosodic organization and of its rhythmic structure internal acous-
tic mechanisms, i.e. what acoustic characteristics and in what set and pro-
portions create word rhythms and what characteristics violate it. We were
interested in the rules of the absolute and relative structuring of the
rhythmic features, a possible compensation of one feature by another, the
limits of variations of intersyllabic relations by one or another acoustic
feature, the extreme possible values of relations which keep a realization
within the limits of the language system. We solved these problems on the
word structures of the minimum syllabic composition (disyllables).

The research methods used in the work are experimental ones, viz. the ana-
lysis of natural words, psycholinguistic tests for their perception, repro-
duction of the words by synthesis (machine modelling of speech).

The general volume of experimental material in intonographic analysis was
represented by 2568 realizations of 200 words pronounced and recorded on
magnetic tape by 7 native speakers (5 with standard literary Hindi and 2
with eastern dialectal Hindi). The technique of full statistic treatment of
the experimental data was applied. We got the information about time,
amplitude and pitch distributions of vocalic and consonantal elements of
the words.

The analysis of time characteristics of words of different prosodic struc-
ture showed that:

1. The quantity (duration) of vocalic elements of syllables in literary
Hindi is functionally valuable, they distinguish words. However, this dis-
tinction is not only and even not as much phonetic but rhythmic also:
words are distinguished by their different prosodic structures. The pure
phonematics of segmental quantity can be found only in Hindi monosyllables,
in which the different length of vowels is associated with different meanings
of words.

2. Words with homonymy of sound segments and with difference in vowels
quantity (jati "going" - jati "caste" - jatī "ascetic") represent in Hindi
of its literary variant the field of distinctive action of word procody,
the action, compared, with one in such English pairs of words as 'object
- object', etc.
3. In literary Hindi there is a clear tendency to a considerable lengthening of second vowels in disyllables of the CVCV, CVCV and ČVCVČ structure. In relative values the temporal prevalence of the second vowels is expressed by the average value 1,3 in the ČVCVČ model, 1,3 in the CVCV model and 1,2 in the ČVCVČ model.

4. The temporal prevalence of the second vowel in each of the considered structures does not exceed a certain critical zone. Any increase of vowels duration within the limits of this zone does not violate the initial etymological models.

5. In eastern dialectal Hindi the second vowels in disyllables also have a tendency to be longer than first ones, but this lengthening by no means can be compared with those quantitative increases of second vowels in literary Hindi. Word prosody in dialectal Hindi often violates the initial temporal proportions of words; all the prosodic models, in fact, turn out to be reduced to one CVCV.

6. The psycholinguistic tests for perception showed the obvious functional value of not only relative but also of absolute quantitative parameters. In average tempo of speech the quantity of long vowels in literary Hindi is characterized by duration from 200 ms up to 480 ms, of short vowels from 60 ms to 160 ms. The linguistic consciousness of a native person (in systems with the opposition of long vowels to short ones) apparently specifies the standard values of sounds quantities with which the concrete realizations are compared. Auditors react on the duration of vowels very sensitively and always reject the realizations of short vowels, exceeding 160 ms, and of long vowels under 200 ms.

7. The quantitative parameters of consonantal elements also take part in the creation of rhythmical models of words. On the average the consonants of short syllables are realized with shorter duration in comparison with the same consonants of long vowels. There is a tendency to the realization of syllables of different pronouncings with approximately equal time, though the length of syllables' elements vary.

8. The duration of close syllables with short vowels approaches the duration of open syllables with long vowels. However, it does matter what elements (vowels or consonants) are responsible for the increase syllable duration. And if there is a danger of confusion (a syllable with short vowel due to the quantity of its consonantal elements becomes equal in duration to a syllable with long vowel, the distinction is accomplished only by the length of vowels.
The main deduction of the intensity analysis of words was the following: the general intensity of syllable is determined by the intensity of its segments, the basic of which is vowel, in coordination with the time parameters of this syllable. As a rule, a syllable of longer duration is marked also by a larger amplitude. On the whole, the amplitude characteristics of syllables and their distribution within a word, play no independent role in rhythmic organization of word, these characteristics are attendant ones.

The analysis of fundamental pitch showed that this characteristic as well as amplitudes is of no independent significance for the prosodic structure of word.

The verification of these results was accomplished by the method of speech synthesis. The formant synthesizer SPPI - 75 was used.

The main results of Hindî words synthesis are:

1. The artificial modelling of Hindî words confirms the trustworthiness of the intonographic analysis of natural Hindî words. Consistent and many times repeated verification of every acoustic parameter for its rhythmic value reveals the acoustic mechanisms of Hindî word-building.

2. The synthesis of Hindî disyllables of different prosodic structure showed, that the main rhythmic factor is the duration (length) of vowels and syllables. In each prosodic model the certain temporal proportions of vowels and syllables function, the violation of which makes a word unrecognizable or of perverted rhythmic appearance.

3. The temporal (quantitative) proportions of vowels are the minimum sufficient parameter for Hindî word rhythms. The distinction of rhythmic models can be accomplished by this one parameter.

4. The temporal syllabic proportions of consonants increase the prosodic reliability of words and make rhythmic models more stable and plastic. The same function bear amplitude and pitch characteristics. In Hindî rhythmic system at word level a clear tendency to the coordination of all rhythmic parameters according to the \( t-A-F_0 \) formula is seen, which means that syllables of longer duration are usually marked by larger amplitudes and higher fundamental pitch. Synthetic realizations with such a distribution of acoustic parameters are usually qualified as the best, most natural and fully expressed. The natural appearance of speech (it's prosodic correspondence to the orthoepic norms) is also supported by the qualitative distinctions of short and long vowels and by their natural distinctions in pitch and intensity.
5. Synthesis more than analysis reveals the significant role played by quantitative standards of short and long vowels in linguistic consciousness of native people. Hindi people very sensitively react on different increases and decreases of vowels duration. This is, apparently, a typological feature of linguistic systems, operating with the phonemic opposition of short vowels to long ones.
APPLICATION OF KNOWLEDGE-BASED TECHNIQUES AND
STOCHASTIC MODELLING FOR RUSSIAN SPEECH RECOGNITION

E.B. Kanitcheva, A.L. Credchenko

Academy of Sciences of U.S.S.R., Computer Centre,
Vavilova str. 40, Moscow 117967, U.S.S.R.

Abstract

Speaker-independent speech recognition system for continuously spoken
Russian sentences is described. Recognition process is as follows: (1) au-
tomatic segmentation of speech into broad acoustic-phonetic categories
with the help of acoustic-phonetic rules, and labeling of these segments,
(2) attracting of the high linguistic levels (syntactic, semantic and
pragmatic rules). If uncertainty is appeared we also use the verification
module for more precise classification of phonetic categories by means of
stochastic modeling. Our experiments are conducted on dictionary that
consist of continuously spoken strings of digits and simple phrases con-
cerning of telephone dialog domain. Sensible phrases are recognized by
use of linguistic limitations.

Introduction

At present, the hybrid methods are often used for decision of speech
recognition problem. A new approach to the phonetic decoding of continuous
speech is suggested in (3). This approach consist of integrating a stochastic
model and knowledge-based expert system.

One of the main stages in automatic speech recognition is the segmen-
tation and labeling stage. The difficulty of segmentation on the phoneme
(linguistic) components is in the fact that precise segmentation of such
kind is impossible on the grounds of simple acoustic criterions. It is ne-
cessary a detailed knowledge of connections between phonology, articula-
tion and acoustic. Ambiguous of phonetic segmentation (labeling) leads to
the necessity to correct the ill-formed segments on the high levels of
signal processing.

In our work for lexical access we use the approach based on a broad
acoustic-phonetic categorization of the input utterance. Segmentation of
speech into broad phonetic classes decreases the number of alternate hy-
potheses and increases the realiability of the segmentation process. Then
depending on phonetic category the more precise classification is done
by means of Hidden Markov Models of each phonetic units. The verification
module evaluates the likelihood that a particular phonetic units has
been uttered.
Knowledge-based models have proven reasonably efficient, particularly when one can incorporate in the rules the human expertise accumulated through the examination of large number of cases. One way of capturing such an expert knowledge consists of considering the activity of a phonetitian while reading a speech spectrogram(1).

The attempt to interact several methods on the different speech recognition levels is described in our paper.

System overview

The purpose of the first stage of processing is to divide the input signal into broad acoustic-phonetic classes and to classify the segments into one of six categories: vocalic segment, voiced/voiced plosive, fricative-like segment, transient segment, silent.

The principles of the segmentation are based on the analysis of set of parameters. In current system we use the zero-crossings of the acoustical signal and its derivative, the total energy, waveform proper, et al. Information about periodicity was provided by a pitch detection algorithm. In addition to this parameters we also computer the number of intervals between two zero crossings which are less than 100 ms and greater than 200 ms. This measurements are used for determining the structure of the noise-like component of some consonants, and are based on results described in(2). This set of data and rules applied to speech parameters was used to locate segment boundaries and assign broad category labels to the resulting segments. The output of the algorithm is a sequence of segments with broad category labels. The parameters for each segment are stored in the database.

This method of segmentation was tested in a speaker-independent mode and the results were compared with manual segmentation. The experimental results show that standard deviation of differences between manually labeled and automatically obtained boundaries ranged from 28 ms to 47 ms.

Sequence of segments with broad category labels is a result of the first stage of processing. On second stage we attract high linguistic levels.

Knowledge attracting about high linguistic levels - syntax, semantic and concrete domain of speech (pragmatic) - decrease a number of mistakes in the time of lexical succession forming and utterance understanding. These knowledge are formalized in the form of production
rules. They can be progressively modified. Using of syntactic knowledge permits progressively modified. Using of syntactic knowledge permits to estimate a grammar acceptability of given word succession. Using of semantic knowledge confirms a sensibility of this word succession. Sentence can be ideal with point of view of syntax but absolutely senseless. Pragmatic knowledge, i.e. knowledge about context, are realized by different rules. Some pragmatic rules are determined and supported by words which characterize a sense of dialog. Another rules help to refuse from some specific interpretation of sequence sense. There are either rules which connect sense of separate frase with common sense of dialog. In given system linguistic limitations on sentence pronunciation are reduced to the requirement to diminish a branch coefficient in the time of utterance forming.

Knowledge are represented in a single structure. All available knowledge are precompiled into a large network which contains all the phonological variations of the various sentences allowed by the grammar of the language. The recognition of a sentence consists of finding the optimal path in this network.

If uncertainty is appeared we use module of segment verification for more precise classification of phonetic categories by means of stochastic modeling. We use 15 phonetic units, each of which is modeled by a discrete Hidden Markov Model. The speech signal is digitized at 10kHz with 10 bit quantization. We compute 10 LPC cepstral coefficients once every 10 ms. We also use some above parameters. Their values are preserved for each phonetic class. A clustering algorithm determines a codebook of 128 prototypes. The training set is made of 20 phonetically balanced sentences of the Russian language. We initialize our training procedure by using a labeled database to train a model for each phonetic units using the Forward-Backward algorithm (4). The probability of an observation sequence coming from a certain RMM is computed using the Viterbi algorithm.

Conclusion

In the future, we shall concentrate the attention on the incorporation of additional knowledge about speech. It will certainly help significantly in improving the segmentation and phonetic labeling of speech. Knowledge-based technology provides a powerful tool for improving our knowledge about speech decoding process.
References

RUSSIAN SPEECH RECOGNITION AND TRANSLATION INTO ANOTHER LANGUAGES

A.L. Gredchenko


Abstract

A knowledge-based system for continuous Russian speech recognition and translation into another languages is developed. The work of the system includes three stages: 1) recognition of speaker-independent continuous Russian utterance; 2) translation of recognized utterance into English, French, German, Spanish and Italian languages; 3) synthesis of translated utterance. The main attention is given to first stage. Speech recognition is examined as a process that uses a widening linguistic knowledge base, particularly, widening information about different levels of language - phonetic, syntactic, semantic and pragmatic. The developed software simulates a thinking of expert-linguist that reads a speech spectrogram and also a thinking of interpreter. Expert specific knowledge and also interpreter knowledge are formalised in the form of production rules and thus can be progressively modified.

Introduction

The use of speech recognition and speech synthesis technologies together with machine translation now offers the possibility of automatically translating speech. The increasing opportunities for spoken communication between countries having different languages suggests that there is a clear need for such a system. In Japan especially there are great pressures to provide an automatic translation facility to aid overseas communications.

Nippon Telegraph and Telephone Corporation (NTT) has announced the beginning of a ten-year project aimed at developing a telephone system capable of translating Japanese into English and vice versa as the conversing parties speak to one another[1].

NEC speech translation system was demonstrated on the exhibition EXPO'85 [4]. The demonstration system comprised a speaker-dependent, continuous speech recognition system capable of recognizing up to 120 words; a multilingual machine translation system designed to automatically translate a maximum of 500 Japanese or English phrases into English, Spanish and Russian or Japanese, Spanish and Russian, respectively; and a voice synthesis system used to convert the translated phrases (English and Japanese only) into spoken output for broadcasting over the loudspeaker system.

Difficult problems specific to speech translation lie in the handling of recognition errors and the parsing of disfluent speech[5]. At present, the design of speech recognition systems appears as one of the most difficult tasks in artificial intel-
telligence. On the one hand, the data to be processed are often erroneous or incomplete; on the other hand, it is necessary to take into account varied large knowledge sources, in order to accurately account for an utterance. Knowledge-based techniques provide an efficient framework for the formal description of knowledge and development of systems.

The expertise of phoneticians can be formalized into classical production rules of various kinds:
- phonetic class identification rules,
- exclusion rules,
- contextual rules (the most common ones),
- meta-rules expressing the choice of various strategies according to the actual situation.

This paper describes such rules-based speech translation system.

System overview

System consists of:
- knowledge base (KB) of expert-linguist who reads the spectrograms and parameter matrix (which reflect integral properties of short segments of speech signal) and also interpreter KB;
- inference engine (program that finds logical corollaries from all rules);
- phoneme segmentation block (receipt of utterance squeeze pattern);
- decision reception verification block.

The performances of speech recognition system are highly dependent on the choice of efficient representation schemes for the various knowledge. There are two extreme solutions:
- the first is to define an unique structure in which all available knowledge will be integrated;
- the second is to keep the various knowledge totally independent, and hence increase the modularity of the system.

Representing knowledge in a single structure is the solution that was chosen in described system. All available knowledge are precompiled into a huge network which contains all the phonological variations of various sentences allowed by the grammar of language. The recognition of a sentence consists of finding the optimal path in this network.

Management is simplified essentially thanks to common form of representation of all knowledge. Each "anchoring point" correspond to one or several utterances. Such knowledge (correspondence to utterance) are accumulated automatically after handling of great number of utterance squeeze patterns of different speakers.

Search begins from deriving of "anchoring points". These "anchoring points" are parts of utterance. Trustworthiness of word or succession of segments is very high for these points. After these small interpretations are extended, extending of best interpretation is effective enough that is it leads directly to utterance recognition. Furthermore, the parts of utterance with low probability are not examined.
Speech is conducted electorally spending as little as possible useless efforts. Expert knowledge permit to discriminate the useful data on early stage, suggesting perspective ways of their using and helping to avoid unjustified attempts as early as possible by cutting of impasse branches.

While accumulating knowledge, it is constantly necessary to take care of system coordination. Any new knowledge has to be coordinate with early available. Sometimes an appearance of a new knowledge requests reconstruction of knowledge base. Special management procedures is developed for it.

The metarules for the solution of conflict situations connected with the application of several rules simultaneously or for the removing of the uncertainty in the priorities is also developed.

In the described system the recognition algorithms work with more coarse speech units – morpheme, word, turn of speech.

Japanese specialists consider that deletion of a phone, syllable, or even a word is often not noticed when a subject listens to a stretch of meaningful discourse, suggesting that human perception of continuous speech would not always rely on segmentation and recognition of smaller units, but would often start with recognition of larger units such as words or idioms[2]. In fact, the acoustic manifestations of such basic units as phones and syllables in connected speech are so much at variance that their correct recognition is almost impossible even by our more advanced techniques. We may thus expect to achieve higher rates of correct recognition if we also adopt larger units such as words, morpheme and idioms along with phones or syllables in the automatic recognition of connected speech.

Lexical data and rules of making more precise of sense are in the data base of second stage. Rules implement a role of editor.

Third stage is directed to decision of problem of high-quality naturally sounding speech generation. Instead of synthetic methods of speech creation the preliminary record of speech answer on foreign language is used in the system. The record is pronounced at first by prepared speaker, after it is transformed into discrete form and is entered in data base. For speech output creation the system defines at first words and phrases that form which is demanded for output speech utterance. Then speech answer dispatcher (one of subroutines of machine engine) realizes a search of discrete speech patterns in data base, analyzes and lines up established words and phrases by rules in the chain without undesirable intermediate pauses. Then the speech is transformed in analog signal and final output sound signal is passed to loud speaker or headphones.

At the expense of semantic-syntactic limitations the translation problem is facilitated because first stage does a main part of work and proposes already sensible utterance to second and third stages.
Conclusion

Speech translation can be potentially applied to virtually any spoken communication across international boundaries. Suitable applications include the translation of spoken communication in specific business areas such as holiday bookings and international car-hire. There is also a need to improve communications between international telephone technicians [13]. Speech translation technology can be applied to the teaching of foreign languages where it could provide valuable assistance to students by automatically correcting their attempts at spoken translation.

References
EFFECTS OF SPEECH INTELLIGIBILITY UPON PERFORMANCE

Georges R. Carinther and Leslie J. Peters

U.S. Army Laboratory Command, Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005-5001

INTRODUCTION

The ability of personnel to communicate accurately is a critical factor in the successful completion of many tasks. Degradation of speech intelligibility (SI) may affect the interpretation of verbal communication leading to misunderstandings, errors and the increased risk of task-related accidents.

A specific example is that of communication in military tracked vehicles which may produce problems affecting system performance. Understanding commands or instructions can, in many instances, be of critical importance. Often it is almost impossible to transmit commands or orders due to noise, distractions, or hearing loss. Case histories are constantly being cited of instances where a tank commander was unable to direct the driver to take certain action, or the gunner misunderstood a command and fired at the wrong target.

PURPOSE

The purpose of this study was to quantify, as a function of speech intelligibility, the performance achieved by a tank crew when conducting gunnery exercises. The SI measure used was the Modified Rhyme Test (MRT).

PROCEDURE

General

The study was conducted in a tank simulator using thirty normal hearing tank crews consisting of a commander and a gunner. Prior to the experiment, the crews were trained in a quiet ideal environment until they consistently achieved a baseline MRT speech intelligibility score of at least 96%.

The experiment required gunnery scenarios to be conducted and performance measures (for each scenario) to be recorded at nominal SI levels of 100, 75, 50, 25 and 0 percent. An electronic chopping circuit, inserted in the earphone, was adjusted to create each of the five desired levels of SI. The commander and the gunner read a single MRT list (50 words) to each other to verify the desired SI level prior to the conduct of each scenario. The scenario was then conducted at that level of intelligibility. After the scenario, the intelligibility test was repeated, with the reported MRT score being the average of the two SI tests.
Following a rest period of approximately one hour, the crews repeated the procedure using a different SI level and a different scenario. This continued until all five scenarios had been presented in counterbalanced order.

**Performance Measures**

The specific measures used to evaluate performance as a function of SI fell into the following four categories:

**Mission Time**
- Time required to identify the target
- Time required to complete the mission

**Mission Completion**
- Percent of targets correctly identified
- Percent of targets hit

**Mission Errors**
- Percent communication errors
- Percent of times wrong target was hit

**Gunner Accuracy**
- Percent of times target was hit by the first round
- Aiming error

The scenario consisted of ten missions each requiring that the commander instruct the gunner to shoot at 1 to 3 targets (21 total targets). Four targets (tank, truck, helicopter, or troops) appeared during each mission and it was the commanders task to instruct the gunner to shoot at the appropriate target with the appropriate weapon. A closed set of commands such as: GUNNER - SABOT - TANK, would alert the GUNNER that he was to locate the enemy TANK and shoot a SABOT round at it. Changes and corrections were made throughout the missions to increase communication intensity. It was the gunner's task to shoot only at the specified target, not at the other "friendly targets", and to instruct the driver to move quickly behind a berm after each mission to avoid being "killed" by the computer.

The commander was directed to repeat instructions not understood by the gunner or use other words, realizing that excessive delay and exposure might get him "killed". If he was unable to properly direct the gunner, he could proceed to the next mission.

**RESULTS**

The mean speech intelligibility levels actually obtained for the 30 crews, at the five nominal experimental SI levels were 93.5%, 73.6%, 52.1%, 26.3%, and 7.1%.

A multivariate inferential statistical analysis found a significant difference for performance measures at the different levels of speech intelligibility. A follow-up contrast of 100% at 100% was not found to be different from 75% nor was the difference between 75%
and 50% found to be significant. However the target ID time for 50% vs 25% and 25% vs 0% were both found to be significant. This pattern held for time to hit the target and for overall time.

**Mission Time**

The time required to complete various aspects of the mission were measured as a function of SI. The mean time required to identify a target and to complete a mission are shown in Figures 1 and 2.

![Fig. 1. Time required to identify target.](image)

![Fig. 2. Time required to complete mission.](image)

**Mission Completion**

The percent of tasks correctly accomplished were measured as a function of SI. The percent of targets correctly identified and the percent of targets hit are shown in Figures 3 and 4.

![Fig. 3. Targets identified.](image)

![Fig. 4. Targets hit.](image)

**Mission Errors**

The percent of errors made during the conduct of the mission were measured as a function of SI. The percent of concepts not correctly communicated to the gunner and the percent of times the wrong target was shot are shown in Figures 5 and 6.
Gunner Error

Gunner aiming error was measured and found to vary only slightly, since once the proper target was located, SI had little effect on accuracy. The percent of time a target was hit by the first round decreased from 90% to 42% as speech intelligibility decreased from 93.5% to 7.1%.

CONCLUSIONS

The results of this study have quantified some effects of speech intelligibility level upon the performance of simple tasks controlled by speech communication requiring a closed set response. It was shown, for this task, that the greatest change in performance was caused by difficulty in identifying the target. Once the target was identified, the time to fire and hit the targets remained fairly constant among SI levels. It was found that performance remained fairly constant until SI was reduced below 28%.

REFERENCE

SOME PSYCHOLINGUISTIC ASPECTS OF TEXT PERCEPTION AND UNDERSTANDING

G.V. Ivanova, A.M. Shakhnarovich
Institute of Linguistics, Academy of Sciences,
Semashko 1/12, Moscow, USSR

One of the main points of the investigation in the field of Artificial Intelligence Systems (AIS) creation is their linguistic supplement, especially in connection with acoustic dialogue "Man-Computer" on natural language. It means that the problem of speech perception and understanding is the most important one. Different approaches to the construction of speech perception models in AIS are discussed in the report.

Actually, the question, what processes take place during text understanding and production, reduces to analysing the fact, what we are doing, while creating a text in order to express our knowledge about the reality and what we are doing, while communicating a text in order to understand it.

No doubt that a speaker, producing a text, constructs it so that an addressee could understand it. No doubt either, that an addressee, while perceiving a text, has an aim to understand it. However, a lot of problems remain unsolved. Here we can mention some problems connected with the process of text understanding, its inner mechanisms and factors which can influence the better (the most complete) text understanding. To answer these questions one should analyse the structure of communicational act. Distinguishing the place of a text in this act help to answer the question on the steps of the communication's participant as regards text as an instrument of communication and as an object of joint activity of the partners (an addressee and an addressee).

In the model of communicational act there are no those components, that can be called as "inner", however, the role of which is very important. In particular, the role of "inner" components of communication is extremely important for the orientation in the situation of the communication, for the actualization of the encyclopaedic knowledge.

The main problem here is connected with the absence of the information about reality within a computer. A computer does not have a system of qualification and classification of elements of the objective world either. This system is a result of an interaction with this world either with the help of direct, spontaneous way, or with the help of symbolic acts. Of course a "man-machine" dialogue may take place with the help of any form language, and not with the help of a natural one. But any formal language still remains a derivative one from a natural language. And it is the natural language, that contains all the information about the real world that serves to overcoming the computer isolation from the external world.

Texts in the act of communication are considered to be as something already given, though they are changing all the time structurally and meaningfully in the process of communication. These changes are caused by constantly changing external and internal factors. Here we can also mention encyclopaedic knowledge which can be actualized in different moments of communication with different degree of its apprehension.

Text understanding takes place, when the following connections are actualized, or established: "text-reality" and "text-receptent". These connections are established in the process of comparing the meaning of the text, its contents with the receptent's experience. The individual experience may be fixed with the help of a set (system) of patterns (the common patterns of reality fragments), to be able to classify the world's
elements and to use the knowledge in communication acts.

Text understanding and thus the reality, that is reflected by it can be held on the basis of qualification system, which is created by a person in the process of his individual developing under the individual strategy. Thus, the model of this qualification system is the following: perception - recognition - classification. The basis of this model is knowledge base of an individual, determined by the whole spectrum of his semantic memory.

A number of experimental series were undertaken, the subjects being preschoolers in a Moscow kinder-garten. The study of speech in children allowed us to follow speech processes in their evolution.

Numerous works show that the acquisition of Russian morphology in children is based on the development of orientation in the sound shape of words. Primarily the child is guided by the general sound properties of the morpheme, and it is only later that he begins distinguishing separate phonemes in it. Thus the child's original vocabulary comprises root words rather than items consisting of separate sounds. This implies that a child segments out syllables both in hearing and speaking. Therefore, in experiments with children we take rhythmic structures, groups of words as perception pivots.

The experimental data have been processed and analyzed by comparing the actual contents and intonation structure of the original text on the one hand and the subjective contents and intonation structure modelled by the subject in the reproduction, on the other. Content segmentation as made by a professional announcer has been compared to that in children, who acted intuitively. The degree of similarity between the respective structures testifies to the acquired level of the text understanding. Besides, it helps us to reveal some intuitive rules of text segmentation by meaningful components.

The results of the experiments showed that the main sense fragment, identified by kids, were not isolated words but groups of words. All groups of words were divided into three main classes according to their occurrence in speech.

Special words groups, recognized and identified at once were noted as nucleous word groups. The word groups recognized and thus pronounced by the most amount of participants were referred to as the significant. These word groups were not identified by every participant of our experiment, as it happened in the case of nucleous word groups. The third group was consisted of the rest word groups, that did not occur in the speech of kids, that were not important for the understanding the whole text.

Thus, according to our classification we suggested a new model of sense segmentation and speech understanding. This model is based on the identification of the nucleous word groups, that are recognized at once.
SPEECH COMMUNICATION IN NOISE AND REVERBERATION: AN EXPERIMENTAL APPROACH

L. Legros*, R. Gamba*
+Université Toulouse II, S. Allêes Antonio Machado.
31058 Toulouse CAdex. France.
*Cabinet d'Acoustique Gamba-Tissayre.
Chemin de Mânel. 31400 Toulouse. France.

INTRODUCTION

The estimation of speech intelligibility in some specific conditions need psychoacoustical tests or physical investigations. Any way, the first ones are very useful to verify the physical criteria used. We performed a method, based upon a Human Voice Simulator (electroacoustic device) in such a way to evaluate the quality of communication in several severe conditions: high reverberation and high background noise level. The same method was used to test, in actual conditions, the influence of the use of individual protective device.

We describe quickly the electroacoustic device used. Then, some experiments are described, and the results shown.

THE METHOD

The purpose is to simulate a human speaker, in a way to perform tests in exactly the same conditions. To do so, we especially had to study the quality of the loudspeaker, in its directivity characteristics. This was done with 9 loudspeakers in an anechoic chamber. Our choice was with a spherical one, reaching a diameter of 20 cm. (not very surprising according to the size of the human head!). Next, the phonetic material is very important. So we have set up 10 lists of 34 words with 3 phonemes each, each list being phonetically balanced with French language. Here is such a list: bile, dors, meurt, sage, gaine, honteux, mule, cale, rive, neuf, sol, crain, phase, chatte, pays, heurte, dupa, ami, aidé, gel, doute, suc, pelle, aveu, danse, ponte, brun, sueur, tir, immé, quinte, sol, lit, auvent.

To take into account the acoustical characteristics of the room, and specially reverberation, we include each word in a short meaningless sentence, in which the word is the last one, such as: The first word is "bile". According to (1), the phonetic balance is estimated by a $\chi^2$ test. Each list has a $\chi^2$ value less or equal to 3.1.

THE EXPERIMENTATIONS PERFORMED

The rooms used

Three rooms were used, with reverberation times showed in the Table 1 below:

<table>
<thead>
<tr>
<th></th>
<th>125 Hz</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR0</td>
<td>0.72</td>
<td>0.76</td>
<td>0.84</td>
<td>0.66</td>
<td>0.88</td>
<td>0.72</td>
</tr>
<tr>
<td>TR1</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.2</td>
<td>2.2</td>
<td>1.7</td>
</tr>
<tr>
<td>TR2</td>
<td>9.9</td>
<td>9.9</td>
<td>8.3</td>
<td>7</td>
<td>3.9</td>
<td></td>
</tr>
</tbody>
</table>

The tests. In each room, 3 background noise levels were studied: 65, 80 and 90 dBA. Leading to speech levels from 90 dBA to 100 dBA. The lower levels of speech were tested with the Human Voice Simulator, and the higher ones with a high efficiency loudspeaker (94 dB/LW/Lm).

Half of the subjects were tested with ear protective device, and auditory fatigue; they were generally submitted to 111 dBA of white noise during 15 minutes, so they reached Temporary Threshold Shift (T.T.S.) of about 15 to 20 dB from 1 kHz to 8 kHz.
THE RESULTS
The Ear Protective Devices.
They were used for high background noise levels (beyond 90 dBA). It was shown that they never decrease speech intelligibility scores, but even increase them in some cases (high level and hearing loss).
The Auditory Fatigue.
This phenomenon do not affect intelligibility scores. This is a very interesting result for workers who use to be in noisy environment.
The Background Noise Level.
Up to 95 dBA, the noise levels studied do not show significant modifications in scores.
The Signal-To-Noise Ratio.
The numerous Signal-To-Noise ratios studied have permitted to perform the computation of a regression upon all the experimental points. The relation which is fitting the best is:

\[ Z = \left(1 - 10^{-\frac{A}{n}}\right) \times 100 \]

where: \( A \) is a variable related to the Signal-To-Noise ratio \( q \) and \( n \) are constants; their values are collected in Table 2, in the case of high background noise level: 95 dBA.

Table 2: Coefficients of the regression relation:

<table>
<thead>
<tr>
<th></th>
<th>TR0</th>
<th>TR1</th>
<th>TR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>6.80</td>
<td>7.50</td>
<td>14.95</td>
</tr>
<tr>
<td>q</td>
<td>0.46</td>
<td>0.43</td>
<td>0.40</td>
</tr>
</tbody>
</table>

The results show that the \( q \) coefficient is rather constant, while the \( n \) coefficient is merely variable. We presume (some more experiments are running) that the first is closely related to the loudspeaker characteristics, while the second one is depending on the room.

CONCLUSION
The present study, with its numerous tests (760 in total), can be seen as a preliminary study leading to the elimination of some variables in the case of communication in particularly severe conditions.
- ear protective device is rather useful for intelligibility.
- auditory fatigue and noise level have not bad influence.
- the loudspeaker and the room are taken into account in a logarithmic regression.

It is now necessary to study the relationship between the \( n \) and \( q \) psychoacoustical parameters and one or several physical parameters, in the case of severe acoustical environment.

ACKNOWLEDGMENTS
The study was supported by the French Company for Electricity (Electricité de France).

REFERENCES
MODELLING OF THE SYSTEM "ATTENTION--MEMORY--PERCEPTIVE PHONETICS"

Simeon Jordanov MRCHEV
ul. "Jordan Hishiev" 21A, 8600 JAMBOL, Bulgaria

Recent developments in perceptual phonetics (cf.e.g.[12]) - part of the study of human perception of speech, are associated with advances in the fields of psycho-linguistics, knowledge engineering and applied AI, pattern recognition (cf.e.g.[1,3,5,8]). The key to establishing the units of the phone-structure of natural language is the data acquired by perception, although articulation and acoustic correlates should not be ruled out. Different articulation movements may be comprehended as versions of one articulation in case they lead to the same sound (as a unit in the study of acoustic phonetics, where classes of realizations equivalent with respect to roles in the phonemic structure of spoken words are called sounds, e.g. two sounds are identical irrespective of intensity, duration, etc.).

Perceptual Phonetics (PF) may be considered a member of two different sequences of areas of study: articulation phonetics -- acoustic phonetics—perceptual phonetics, AND perceptual phonetics—perceptual morphology—perceptual syntax. The latter two fields determine the so-called "hearer's grammar". Human memory with its internal structure (cf.[14,7]) given on Fig.1 serves as memo-ware (we use the neo-logism "memo-ware" in the already traditional way) of the perceptual phonetics. In particular: STM and EIM cover temporality FP functions, and interacting OSTM and LTM provide a "memo-computational basis" for comparison, comprehension, interpretation, etc. of speech data by the hearer (guided by his "hearer's grammar"). Auxiliary functions in this respect are performed by ISTM (repetition and assessment of incoming perceptual information), RSTM (consolidation), SLTM and MLTM (intuition, unconscious reproduction of the memo-ware-information).

The perceptual basis of natural language (the linguistic in nature system of processing the incoming utterances) has a general part (the dialect) and an individual part. Both are stored in LTM and exist as systems of standard templates (or simply templates) together with rules for comparison in the course of perception. The general perceptual basis is the common ingredient of normal representatives of a given language community. Under a "template" we understand a quasi-denotation of the stored in LTM memo-units (standard patterns) which are compared with incoming segments and supra-segments of the speech flow. Under "rules for comparison" we understand the stored in LTM procedures (algorithms?) and their consequent applications by which the comparison of data and templates is carried out.

There are templates in LTM: 1. That correspond to sound images in EIM (templates of speech sounds, intonation schemes, rhythmic structures, etc.); 2. That do not correspond to any sound image (templates for different phonemic features, diffusion, bemolity, etc.).

505
Among templates of segments one may distinguish templates for distinct phonemes and templates for phonemic compounds (such as e.g. syllables). There are no special templates for phonemic variants, as well as no templates in the perceptual basis of a language not corresponding to phonetic units of speech.

We assume a zonal organization of templates in LTM with the following structure: 1. There are some primary (atomic?) phonetic units expressed by some domain of the space of values of certain parameters, each of them corresponding to a single measurable physical characteristic; the rest of the units being compound and corresponding to composite systems of domains of the parameters; 2. To compound templates there correspond parameter sets of two types (type A and type B). The units of type A are compound units for which the set of characteristics is the same for any two opposite units and these are distinguished only by the integral value of the compound parameters, e.g. the compound characteristics of accented and unaccented syllables: they both have the same set of parameters, such as duration of the vocal part, duration of the consonant part, intensity of the syllabic peak, frequency of the basic tone, etc.; they have non-intersecting regions of values of the compound parameters (so that syllables with or without accent could be told apart). The units of type B are distinguished from opposite units by the existence of a parameter which is absent in the representation of the counterpart (any phoneme is an example of this type of unit); 3. Thus in contrast to units of type A the identification of units of type B may be based on a specific set of characteristics and not on an integral compound characteristics (as happens in case of units of type A). Based on experimental data, a hypothesis is put forward in [2] that the compound parameters of units of type B can themselves be composed by units of type A. In particular, distinct differential characteristics of the phonemes, occuring in different units, can be established by summing up the values of its components; 4. The templates in LTM of the phonetic units which correspond to sound images in ELM can be represented as zones of identical perception (ZIP). These ZIP's correspond to regions in the space of parameter values in which any two realizations are identical. So any change of the values of the parameters within the limits of the region leads to perceptually indistinguishable realizations. Such a view on the functioning of the templates is founded on ignoring in the perceptual basis of the language of the variations which are small. On the other hand identical reaction to physical features that are "near" enough is physiologically natural. In this respect it resembles the law of "all or nothing"; 5. An immediate neighbourhood of a ZIP is the zone of similarity to the template (ZST). The ZIPS of distinct phonetic units do not intersect, moreover they have non-intersecting closures in the topology, generated by the notion of nearness, while the ZST may well have non-empty common parts and this is one of the explanations for ambiguous perception; 6. For units that do not have a corresponding sound images
the existence of a zone of identical reactions can also be conjectured as well as zones of similarity; 7. The categorical character of speech sounds' perception is rejected, i.e. we do not need the notion of different speech sounds being comprehended in two completely different ways: "categorial" and "non-categorial"; 8. The boundaries of the zones (in particular of ZST) are quite unstable. This could explain the process of change of the phonetical background of a language. The instability of the boundaries have been established by experiments and it seems to be a result of different extralinguistic factors. A very substantial shift in the boundaries can be observed when a specific psychological attitude is adopted during the experiment - a fact that leads sometimes to assimilative or contrastive perceptual illusions, and for this matter should be taken into account when determining templates' boundaries by phonetic experiments.

The rules for comparison with the templates in the perceptual basis of a natural language are intimately connected with the notion of specific weight of a parameter (within the limits of LTM) and with the units of primary perception. As an example let us consider a compound parameter X for an unit of type A. Let X be a linear form of four variables $a_1, a_2, a_3, a_4$:

$$X = \sum_{i=1}^{4} a_i k_i$$

Here the $k_i$'s represent the specific weights of the parameters in X. They are determined not by the individual features of the sound analyzer of a particular person, but by fundamental characteristics of the perceptual basis of the language. To find the values of the coefficients (or the boundaries of their possible changes) is one of the central aims of PP.

The values of the parameters in a compound parameter of type A can compensate each others for example, the perception of an accent in a verbal stimuli with a fixed frequency of the basic tone can be shown to obey a relation of the following kind (here we take a Russian word "sushu"):

$$X = \frac{t_y}{T} + \frac{t_c}{T} + \frac{A_2}{A_1+A_2}$$

Note that here: $t_y$ is the duration of the vocal part of the syllable; $t_c$ is the duration of the consonant part of the syllable; $T$ is the full duration of the two syllabic word (sushu); $A_1$ and $A_2$ are the intensities of the syllabic peaks in the first and the second syllable respectively; $f_{bt}$ is the constant frequency of the basic tone.

507
It has been experimentally shown (cf.12) that when $X<1.5$, almost 80% of the participants in the experiments hear the accent on the second syllable (the values of the component parameters being irrelevant). It should be mentioned that: 1. Mutual compensation occurs proportionally to the specific weights; 2. This mutual compensation is possible at all only because the values of the parameters are such that the reaction to them does not lead to autonomous feeling them as different qualities, i.e. they should be sufficiently big to be recorded by the analyzer, but also not too big as to be perceived independently.

Under units of primary perception we understand templates for such segments and supra-segments of the speech flow that are operative in establishing the "sounding" of an utterance. In experiments with uncommon combinations of consonants, the stimuli have been comprehended with big distortion. This fact shows that the units of primary perception are not the phonemes, i.e. in the perception of unusual combinations of consonants comparison is carried out not with the templates of some phonemes, but with templates of their combinations. If in the set of templates in the perceptual basis of the human mind there is no suitable template (exactly fitting) them the sound image is mapped to all the nearest such templates (in the topology) and to all combinations of them until a suitable combination is found and a satisfactory similarity is established. Of course, another possible explanation is that phonemic templates are indeed the templates of primary units and un the perception of a sounding word a simultaneous correction is taking place. But data from 12! and 15! supports the view that this is not the case and that the units of primary perception are not the phonemes, but certain their compounds, in particular - the syllables. One more reason for this is the fact that in experiments with perception of syllables the reaction time for single phonemes is much greater than the reaction time for syllables themselves. Thus one is bound to insist that the real formative units are the syllables.

For Fig. 1 and Fig. 2 see 19-18! On Fig. 2 we present (in accordance with data about PP some of which was discussed above) a flow-chart of the hypothetical phonetic verbal perception, based on the system of human memory (for a discussion of the model of human memory cf. 17! and for some interconnections between memory and perceptions cf. 14! and 18!).

In conclusion we mention some work done on the topology in the spaces of perception parameters: the notion of nearness most adequate seems to generate a topology which is not a Hausdorff one, but what we can claim at most is that this topology is a $T_1$-topology. It is a matter of further investigation to decide whether some kind of proximity space would not fit the picture better.
Internal structure of human memory: Instantaneous memory (SIM - sensory instantaneous memory; EIM - eidetic instantaneous memory); Short-term memory (STM - immediate short-term memory; OSTM - operative short-term memory; BSTM - buffer short-term memory); Long-term memory (LTM - long-term memory; SLTM - super-long term memory; MLTM - meta long-term memory).

REFERENCES:


7. S.J. Mrchev, Modelling of Human Memory, in: 8th All Union Conf. on Neuro-cybernetics, May 1983, Rostov-on-Don, USSR


9. S.J.Mrchev, A Model of the Perceptive Phonetics, Attended by the Human Memory, 12th International Congress on Acoustics, 24-31 July 1986, Toronto, Canada

10. S.J.Mrchev, A Model of the Perceptive Phonetics, Attended by the Human Memory, Montreal Symposium on Speech Recognition, 2 21-22 July 1986, McGill University, Montreal, Canada


15. S.J. Mrchev, A System of Mathematical Models, Simulating the Human Memory-I: Modelling of Human Short-Term Memory, 11th IMACS World Congress SYSTEM SIMULATION AND SCIENTIFIC COMPUTATION 5-9 August 1985, Oslo, Norway

16. S.J. Mrchev, A System of Mathematical Models, Simulating the Human Memory-II: Modelling of Human Long-Term Memory, 11th IMACS World Congress SYSTEM SIMULATION AND SCIENTIFIC COMPUTATION, 5-9 August 1985, Oslo, Norway

17. S.J. Mrchev, Modelling of Human Instantaneous Memory: I & II, XIV International Conference on Medical and Biological Engineering & VII International Conference on Medical Physics 11-16 August 1985, Espoo, Finland

NASOPHARYNGEAL TRACT TRANSFER FUNCTIONS MEASUREMENTS WITH WHITE NOISE EXCITATION

E. Castell & P. Badin

Institut de la Communication Parlée
I.N.P.G. / E.N.S.E.R.G. - UNIVERSITE STENDHAL
Unité Associée au C.N.R.S. n° 368
46, Avenue Félix Viellet, 38031 GRENOBLE Cedex, FRANCE

Abstract:

To understand acoustic phenomena involve in the production of the nasal vowels, we have built an experimental setting to obtain vocal tract transfer functions. Our setting is an improved version of FUJIMURA & LINDOVISte’s method [1]. The vocal tract is excited through the skin of the glottis by a white noise and the transfer functions are obtained by averaging the F.F.T. spectra [2]. Results are rather complex, but we show that formants at 250 Hz and 1000 Hz of the transfer function of the nasal velar consonant could be imputed to the first resonances of the nasopharyngeal tract : this would agree FENG's predictions [3]. Other poles could be explained by sinus cavities.

Introduction:

Study of acoustic production of nasal vowels is a delicate point because the respective influences of the nasal and paranasal cavities in speech production are not clearly understood. We think that the systematic analysis of transfer functions, measured in well controlled productions, can help us to explain the phenomenon.

MAEDA [3], carrying on works of DELATTRE [4], has drawn attention to the importance of the first low formant (250 Hz) as a nasality characteristic. FENG et al [5] [6], considering the nasopharyngeal tract, have shown the possible influence of the small area at "limen nasi" upon this first resonance. In order to reconfirm these results by direct measurements, we have built an experimental setting to obtain vocal tract transfer functions [2].

Experimental setting and method:

FUJIMURA & LINDOVIStes [1] excited transcutaneously the vocal tract at the glottis with a pure tone swept in frequency, and measured the sound level at lips or at the nostrils. In our improved version of this method, the vocal tract is excited with white noise. The subject applies a small loudspeaker membrane externally at the level of the thyroid cartilage, checking that there is no noticeable sound leakage at the junction between the loudspeaker and the skin. The loudspeaker is then fed with a white noise signal with a flat spectrum (less than 3 dB fluctuation) between 20 Hz and 20 kHz (fig. 1). The output signal is picked up, at a distance smaller than 1 cm at the lips or at one of the nostrils by rather directing Electret microphone, and recorded through a P.C.M. Coder into a BetaMax video tape recorder. The signal is as well sent back to the subject by means of headphones; this allows an interesting feature : a control thanks to auditory feedback [5].

![Vocal Tract Transfer function measurement](image)

Figure n° 1. Experimental Setting.
Naso-pharyngeal tract transfer functions:

In our study, we are interested in one of the nasality characteristics: the low pole (around 250 Hz), which can possibly have its origin in a HELMHOLTZ resonance - as suggested by FENG & al. (5) (6). The Helmholtz resonator consists of the naso-pharyngeal tract - acting as the body - and the nostrils - acting as the neck of the resonator.

The frequency of the resonance depends on the volume $V$ of the resonator, the length $l$ and the area $A$ of the neck according to:

$$F_0 = \frac{c}{2\pi} \sqrt{\frac{A}{l \cdot V}}$$

In order to find the relation between the low pole frequency and the limen nasal area, we have varied artificially this area with a series of small plexiglas tubes (2). The transfer functions of the naso-pharyngeal tract are then measured. The naso-pharyngeal tract is considered as a simple tube constituted of the larynx, the pharynx and the nasal cavities.

Results:

For nine subjects, we have systematically measured and drawn the transfer functions for each nostril, for four possible limen nasal areas and for each of the vowels [a], [i], [u] and [o]. We have obtained a minimum of six measurements for each case. Most transfer functions display a structure much more complex that our simulated functions. Therefore, some measured spectral peaks and valleys seem to have fairly constant values. For each subject, the formants and zeros frequencies are very consistent. The standard deviation do not go up ten percent.

Figure n° 2. Standard naso-pharyngeal transfer functions.
So we can draw standard transfer functions for each subject. It divides up into two classes. One class shows a simple transfer function with only six poles and one zero between 0 and 5000 Hz. The other is more complex: up to nine peaks and up to three zeros in this interval (see figure no. 2).

A striking observation is that the lowering of the first resonance frequency of the naso-pharyngeal tract due to the decreased area at ilmen nasal is much less than we expected: whereas the Helmholtz formula leads to a decrease by more than 100 Hz, when reducing the diameter from 4.5 mm to 2 mm, we have hardly measured a decrease bigger than 30 Hz (see figure no. 4).

Discussion:

The formant Fn1, Fn5, Fn7 and Fn9 (respectively around 250, 1000, 2000 and 3000 Hz) are the resonances of the naso-pharyngeal tract if we consider this tract as a simple tube. Measured values agree with FENG's simulations (figure no. 3) [5] [6].

Fn2, Fn3 and Fn4 (respectively around 400, 600 and 800 Hz) could be explained by sinus cavities. We have often found zeros between these formants. Fn9 (around 1700 Hz) is the result of the disymmetry of the two nostrils.

The lower limit of the first formant could be explained by wall vibration effects, in the same way described by ISHIZAKA & FLANAGAN [7] and FANT [8]. Considering that the Helmholtz formula gives the resonance for the hard wall case, we have applied the approximate formula proposed by FANT, with a tentative value of wall vibration frequency Fw around 200 Hz:

\[ F_0 = \sqrt{Fw^2 + Fh^2} \]

We can see the results in the figure no. 4, where we have draw the HELMHOLTZ formula, the FANT's correction and the measured values. We can notice that FANT's correction and measured values draw the same curve: Fn1 seems to be a HELMHOLTZ resonance. However, the volume of the resonator for the corresponding Helmholtz frequency (around 150 Hz) should be very large - more that 800 cm3. The simple naso-pharyngeal tract can not be this resonator. The sinus cavities can not also explain this resonance. We must perhaps consider other coupling between head cavities and vocal tract.

Anyway, we can now ensure that the first low formant of the nasal vowels and of the nasal velar consonant is not a glottal formant - as suggested by MAEDA [9], because all records of measured transfer functions are obtained with the glottis completely closed.

Figure no. 3. FENG's simulation of naso-pharyngeal tract transfer function.
Conclusion:

A renewed method for direct acoustic measurements of vocal tract transfer functions has been described. This method supplies very consistent results. This should provide us with a quantity of fresh data on the vocal tract, useful for vowels as well as for consonants.

The very low first formant of the nasal vowels and of the nasal velar consonant seems to be a HELMHOLTZ resonance. Nevertheless, because of the wall vibration effects, we cannot consider that this formant is the resonance of the only naso pharyngeal tract. The sinus cavities cannot not also explain this very low frequency: the volume of these cavities is too small. How important can be the role of the head cavities in the explanation of this phenomenon?

References:


514
ECHO ACCEPTANCE CONVERSATION TESTS

D. PASCAL & D. CORVAISIER

CNET Lannion A, Département CMC
Route de Trégastel, BP 40, 22301 Lannion Cedex, France

INTRODUCTION

Problems of echo appearance with short delays (about 30 ms) will become increasingly important for the near future as mixed configurations of all digital telecommunication networks will interwork with analog ones. The echo signal is provided by the mismatch in the 4w/2w hybrid transformer where the speech signal is reflected towards the part using the numerical set. This phenomenon can be perceived as a delayed sidetone by the subscriber. To investigate the effect of this echo phenomenon on the perceived quality of a telephone connection, an interactive evaluation procedure (laboratory conversation test) involving two participants, a talker and a listener, was used (International Telegraph and Telephone Consultative Committee, CCITT, 84a).

EXPERIMENTAL PROCEDURE

For these specific tests, conversational situations are simulated in a controlled environment. Pairs of subjects are invited to hold a conversation with the help of proposed conversation tasks. At the end of the communication, they are individually asked to assess the perceived quality of the connection. In this experiment, we simulated configurations where only one of the two subscribers received an echo signal while speaking (talker echo). The second subscriber only had to keep the conversation going. Meanwhile, no one was told about that dissymmetry and both of them assessed the speech quality.

Experimental conditions

Parameters which varied in the experiment were the delay of the echo signal (16, 24, 32, 40, and 48 ms) and its attenuation level OELR (Overall Echo Loudness Rating: 14, 19, 24 dB). As regards the range of delays, 50 ms is considered to be the threshold for extreme disturbance, and CCITT recommends the use of echo control devices beyond this threshold. The range investigated of echo attenuation levels covers values usually found in telecommunication networks, i.e. from 15 dB (strong echo) to 30 dB (weak echo). Eleven conditions were selected from the matrix of parameter levels as shown in Table 1. An echoless connection was added to provide a quality reference. On the whole, twelve distinct conditions were tested.

<table>
<thead>
<tr>
<th>DELAY</th>
<th>0 ms</th>
<th>16 ms</th>
<th>24 ms</th>
<th>32 ms</th>
<th>40 ms</th>
<th>48 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>OELR:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 dB</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td></td>
</tr>
<tr>
<td>19 dB</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td></td>
</tr>
<tr>
<td>24 dB</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td>tested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no echo</td>
<td>tested</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: the 12 test conditions

Extraneous conditions

To control nuisance variables (Kirk, 82): 1) Some were kept constant, so, the Overall Loudness Rating (OLR) of the bidirectional link was fixed at 8 dB. This value determines the level of the voice
received by each subscriber, and it corresponds to the optimal listening comfort for a telephone conversation. The permanent sidetone was fixed at 12 dB, that is the lowest level for SideTone Masking Rating in the range recommended by the CCITT: from 7 dB (rather strong sidetone), up to 12 dB (rather low sidetone). Such a weak value was selected to avoid the masking effect of sidetone on the echo signal as much as possible. A 44 dB(A) white noise was injected in each link, corresponding to the maximum authorized level of the circuit noise and the environment noise of each testing place was defined by a Hoth's spectrum with a level of 45 dB(A).

2) Other nuisance variables were included as factors in the experimental design.

Experimental design
The experiment was based on a double graeco-latin square including 24 pairs of subjects and 12 circuit conditions. It was divided into 2 sessions, each one composed of 6 conversations. An extra conversation was added to the first session as an initial training condition (OELR= 14 dB, delay= 16 ms). Thus, each pair of subjects held 13 conversations in all, starting with the same reference quality.

A graeco-latin square experimental design makes it possible to take into account at least 3 sources of undesirable variability: the individual subject behaviour, the conversation task, and the order of the telephone connections. Hence, it allows their effects on the whole experiment to be balanced out. The modalities of the subject (in fact, pair of subjects) and "presentation order" variables are assigned to the rows and columns of the square respectively. The different telephone configurations are assigned to the cells of the principal latin square, and the modalities of the conversation tasks to the cells of the latin square which is orthogonal to the primary square. So, each condition is tested with all conversation tasks, which are different for the same pair of subjects, in each successive configuration.

Method of assessment
After each conversation, the subjects had to individually express their opinion about the speech quality by filling in three 5-point discrete scales (category judgment): 1) a quality scale (Q1) determined by the usual following categories: Bad, Poor, Fair, Good, Excellent; 2) an acceptability scale (Q2) limited by two opposite adjectives (Unacceptable ... Acceptable), and 3) an impairment scale (Q3) to assess the perceived quality degradation induced by the echo phenomenon.

After assigning to each category a number from 1 to 5, the first quality scale will result in the well known mean opinion score MOS (ACR method, CCITT, 84b), the second one in an analogous "mean acceptability score", and the third scale in a degradation mean opinion score DMOS (DCR method, CCITT, 84c). This last scale is specified by the five categories: Unperceptible (score:5), Perceptible but not annoying (4), Slightly annoying (3), Annoying (2), Very annoying (1). This impairment scale is taken out of the DCR method but its application to conversation tests differs from the procedure recommended for listening tests in that no explicit good quality reference is introduced prior to each evaluation.

RESULTS

After putting aside the first set of subjects' votes which were associated with the training conversation, mean values were calculated for each judgment scale, and plotted against echo delays with echo attenuation as the parameter. The MOS and the DMOS diagrams are given in figures 1 and 2. The mean score diagrams relating to the acceptability criterion (Q2 results) is not reported here as it is highly correlated to quality criterion (Q1 results) and more generally to any global opinion criterion such as naturalness and pleasantness of the voice (Guillot, Monfort & Pascali, 85; Pascali, 88). The degradation of quality is considerable even with the smallest echo phenomenon: the MOS loses 0.4
point from the reference condition (no echo) to the nearest (24 ms, 24 dB OELR) and "only" 1 point from this latter to the worst (40 ms, 14 dB OELR). Comparing the results obtained from the quality scale and the impairment scale (e.g. comparing MOS and DMOS curves and related matrices for specific comparisons between means) shows that the DMOS scale leads to a wider distribution of the scores than the MOS scale, as well as to more discriminating evaluations. The MOS curve seems to reveal that the echo attenuation effect increases with the delay. This is due to the partial masking of echo level (-14 dB to -24 dB) by the direct sidetone (-12 dB), and therefore only to the influence of the echo delay (CCITT, 88). But, it is mainly due to the insufficient discriminability of the ACR method.

The DCR method enables us to draw more interesting conclusions. The degradation of quality results from two parameters, namely echo attenuation and echo delay, so it might be worthwhile to modelize the integration process of stimulus information within the subject's perceptual apparatus. The data plotted in Figure 2 follow the parallelism prediction of the averaging model for stimulus integration (Anderson, 74). The integration rule is an arithmetic mean and equal weighting is assumed within each physical dimension (w\text{delay} and w\text{attenuation}). Furthermore, the vertical spacing of the three attenuation curves constitutes an interval scale of the quality degradation. To statistically validate this perceptual model, a variance analysis was performed on the DMOS scores assigned to the nine test conditions building a true factorial plan. It confirms the lack of interaction between the two physical dimensions and leads to the conclusion that the weight of echo attenuation is twice as high as that of echo delay. Multiple comparison tests (Tukey HSD test at .05 level) show that, at a given echo delay, a 5 dB difference in echo attenuation results in an insignificant DMOS evaluation difference, while a 10 dB difference leads to significant results. Yet, a 16 ms delay difference along a single attenuation curve gives an insignificant result for the 24-40 ms range. Outside this range, a 16 ms step is significant at a subjective level. Furthermore, the twelve test conditions may be divided in two mutually exclusive categories; one is composed of the no-echo, 24dB*24ms and 14dB*16ms conditions which are labelled as the "acceptable" quality category (DMOS > 3.5), the other one is composed of the two worst echo attenuation levels (19 & 14 dB) associated with delays greater than or equal to 32 ms: these conditions are considered as unacceptable (DMOS < 2.5). These results confirm again, the preponderant influence of echo level with regards to echo delay.

CONCLUSION

This experiment was carried out to investigate the effect of short delay echo on a telephone conversation between two participants. The main conclusion is that the echo level is perceptually more important than this echo delay. Echo delays greater than or equal to 32 ms should be avoided in combination with a high echo level, whereas up to 40 ms can be acceptable with a low echo level. If the echo delay cannot be restricted due to technical difficulties, reducing the loudness of the echo signal is more perceptually efficient.

DMOS scores obtained under similar conditions in a listening situation reveal to be 5 dB worse as it is well known that listening procedure is more severe than conversation one. Due to the influence of the high quality reference actually provided for comparison to the subject, the lowest echo attenuation curve is 1 point lower in the listening test than in the conversation test. Nevertheless, an internal reference formed on the basis of individual experiences is always present in the subject's mind at the time of evaluation. Therefore, without any introduction of a "true" reference, an assessment explicitly formulated in terms of degradation or perceptive difference still benefits by the discriminating properties of the pair comparison method in general and the DCR method in particular. Hence, the degradation quality scale is much more informative than the usual quality scale without any additional cost to experimental procedure.
REFERENCES


Figure 1: MEAN OPINION SCORE

Figure 2: DEGRADATION MEAN OPINION SCORE