Applied Acoustics **39** (1993) 291–305

Differences Between Headphones and Loudspeakers Listening in Spatial Properties of Sound Perception

A. Fischetti,[a] Y. Hemim[b] & J. Jouhaneau[a]

[a] Laboratoire d'Acoustique, Conservatoire National des Arts et Métiers, 292 rue St Martin, 75141 Paris Cedex 03, France
[b] Laboratoire de Statistique de l'Université Paris VI, place Jussieu, 75005 Paris, France

(Received 24 February 1992; revised version received 8 June 1992; accepted 11 June 1992)

ABSTRACT

This paper deals with change of subjective spatialisation induced by loudspeaker listening of dummy-head recorded musical samples. Data were analysed through multiple correspondences analysis. Three kinds of analysis were performed: (1) subjective characteristics with headphones and loudspeakers; (2) interindividual differences with headphones; (3) interindividual differences with loudspeakers.

1 INTRODUCTION

The initial purpose of this paper was an improvement of the knowledge of recording/diffusing chain effect on some spatial properties of sound perception. The quality of concert halls is usually investigated through dummy-head recorded sound samples listened to with headphones, and is supposed to give to the listener the same impression as when sitting in the room. However, sound perception could be shifted under normal loudspeaker listening. A previous experiment was performed in ESPRO (Espace de Projection), the acoustically modulable room of IRCAM (Institut de Recherche et Coordination Acoustique/Musique, Paris): multidimensional relationships between six spatial subjective characteristics, eight acoustical criteria and three geometrical parameters were presented.[1,2]

Applied Acoustics 0003-682X/93/$06.00

This study emphasizes differences between headphone and loudspeaker perception of sound spatial characteristics.

2 METHOD

2.1 Physical measurements

ESPRO is a parallelepipedic room (24 m long and 15 m wide) whose walls are composed of orientable trihedrals with absorptive–diffusing–reflecting sides. The height of the ceiling is variable from 2·5 to 12·5 m. An anechoic musical signal was diffused in ESPRO through two loudspeakers, and recorded at three positions with two stereophonic microphones placed on each side of a dummy head. Impulse responses were also measured and 15 acoustical criteria were computed. The procedure was repeated for 81 configurations of ESPRO, chosen from different combinations of

TABLE 1
Geometrical Parameters of ESPRO

Code	*H (m)*	*A (%)*	*D (%)*	*R (%)*	*A/D (%)*
1012, 1013	10	100	0	0	— (strong)
1071, 1072, 1073	10	69	0	31	— (strong)
1131, 1132, 1133	10	53	32	15	1·65 (median)
1181, 1182, 1183	10	22	78	0	0·28 (weak)
1271, 1272	10	0	0	100	0·30 (weak)
2131, 2132, 2133	7·5	53	32	15	1·65 (median)
3071, 3072, 3073	5	67	0	33	— (strong)
3131, 3132, 3133	5	59	28	13	2·10 (median)
3181, 3182, 3183	5	25	77	1	0·32 (weak)
3251, 3252, 3253	5	18	50	32	0·36 (weak)
3271, 3272, 3273	5	1	0	99	— (weak)

H: height of ceiling : high (10 m), median (7·5 m) and low (5 m).
A/D: ratio of absorptive to diffusing panels : strong (∞), median (200%) and weak (30%).
P: receiver position in the room : P1 and P3 in the axis of the room, respectively, at 7 and 16 m from the loudspeakers line, and P2 close to a lateral wall at 12 m from the loudspeakers line.
G: absorptive panels distribution gradient, which indicates that absorptive panels are either predominantly grouped or uniformly distributed in the room (three values).
R: percentage of reflecting panels (three values).
Configurations used in the study of ESPRO: First number represented height of ceiling. 1 for high (10 m), 2 for median (7·5 m), and 3 for low (5 m) ceiling. The two following numbers described walls absorption : 01 and 07 for strong, 13 for median, 18, 25 and 27 for weak absorption. Fourth number concerns position in the room. 1 for near (P1), 2 for excentred (P2) and 3 for distant position (P3).

three values of each geometrical parameter (see Table 1). Preliminary tests have shown that variations of absorptive panel distribution in the room were poorly discriminated through listening tests, while reflecting and diffusing panels produced roughly identical effects. Therefore, a set of 31 configurations were selected according to variations of position in the room (P), ratio of absorptive to diffusing panels (A/D) and ceiling height (H). Configurations of ESPRO were represented by a four-number code (see Table 1). Taking into account the correlation between the 15 acoustical criteria,[3] it appeared relevant to select only eight among them (see Table 2).

2.2 Perceptive tests

The sample under test was a 15 s duration piece of Schubert's 14th string quartet. In a first test, musical samples were reproduced through headphones. Each sample was reproduced twice, and 11 sound engineering students, not highly experienced in listening tests, were required to evaluate samples according to six subjective characteristics (see Table 2): apparent room size, depth perception (or apparent distance), lateral localisation, spatial impression, subjective reverberation time and preference. Several days later, 10 of the same subjects repeated the test with loudspeakers listening in a studio with very few reverberations.

2.3 Multidimensional analysis

Data were analysed through multiple correspondences analysis (MCA).[4] Input was a (31, n) matrix, made by the description of 31 configurations

TABLE 2
Definitions

Acoustical criteria
Direct energy (Dir) / Energy arising within the first 40 ms after direct sound (En. 40) / Early decay time (EDT) / Reverberation time (RT60) / Clarity (dB) (C80) = 10 log ($\int_0^{0\cdot 08} p^2(t)\,\mathrm{d}t / \int_{0\cdot 08}^{\infty} p^2(t)\,\mathrm{d}t$) / Ratio of direct to reverberant energy (Dir/rev) / Interaural cross correlation coefficient (IACC) / Total energy of impulse response (En.Tot)

Subjective characteristics
Apparent room size (rs): impression to be in a small or a great room (five choices) / *Apparent distance from the source or depth perception (dp)*: impression to be near or far away from the source (five choices) / *Lateral localisation*: perception of sound direction from the left or from the right (five choices) / *Spatial impression (si)*: impression to be surrounded by the sound or to listen to it like through a small window (four choices) / *Subjective reverberation time (srt)*: impression to be rather in a dull or a reverberant room (rated on a subjective scale ranged from 0 to 10) /*Preference (pr)*: rated on a subjective scale ranged from 0 to 10

TABLE 3
Multidimensional Analysis

I, J: set respectively of configurations (in lines) and variables (in columns)

$k_{i,j}$: table of logical coding which replaces real measurements ($k_{i,j} = 0$ or 1 according to whether measurement falls or not into the class). All following definitions concerning lines are analogous for columns

$k_i = \sum_{j \in J} k_{i,j}$: weight of raw i; $k = \sum_{i \in I, j \in J} k_{i,j}$; $f_J^i = \left\{ f_j^i = \frac{k_{i,j}}{k_i},\ j \in J \right\}$: profile of raw i

$f_i = \sum_{j \in J} \frac{k_{i,j}}{k}$: mass of profile f_J^i $d_{i,i}^2 = \sum_{j \in J} \frac{1}{f_j} = (f_J^i - f_J^{i'})^2$: χ^2 distance centred at f_J

Clouds $N(I)$ and $N(J)$ of profiles of lines and columns f_J^i and f_I^j, respectively, associated with masses f_i and f_j, are considered in the space structured by χ^2 distance centred at f_J

$N(I)$ develops in principal directions of expansion called principal axes of inertia: $u_{\alpha J}$

$F_\alpha(i)$ and $G_\alpha(j)$ are coordinates, respectively, of f_J^i on $u_{\alpha J}$ and f_I^j on $u_{\alpha I}$; $F_\alpha(i) = \sum_{j \in J} f_j^i u_{\alpha j}$

$f_J = \sum_{i \in I} f_i f_J^i$: barycentre of cloud $N(I)$; $f_J^i - f_J = \sum_\alpha F_\alpha(i) u_{\alpha J}$

$\rho^2(i) = \sum_{j \in J} \frac{1}{f_j} (f_j^i - f_j)^2$: χ^2 distance between f_J^i and f_J centred at f_J

$\lambda_\alpha = \sum_{i \in I} f_i F_\alpha^2(i) = \sum_{j \in J} f_j G_\alpha^2(j)$: moment of inertia of cloud $N(I)$ in direction α

Principal axes maximise variance (or total inertia) of the projection of clouds $N(I)$ and $N(J)$:
$\sum_\alpha \lambda_\alpha = \sum_{i \in I} f_i \rho^2(i) = \sum_{j \in J} f_j \rho^2(j)$

$G_\alpha(j)$ are obtained from $F_\alpha(i)$ by the transition formula: $G_\alpha(j) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_i \frac{f_{i,j}}{f_j} F_\alpha(i)$

The relation allows the clouds $N(I)$ and $N(J)$ to be plotted in same planes

For each i (analogous results are obtained for each j), the program prints $f_i F_\alpha(i)$, cor (i, α), and:
$\mathrm{Qlt}(i) = \sum_\alpha \cos^2(i, \alpha)$: quality of representation of element i

$\mathrm{in}(i) = \frac{f_i \rho^2(i)}{\sum_i f_i \rho^2(i)}$: relative contribution of element i to total inertia

$\mathrm{cor}\ (i, a) = \frac{F_\alpha^2(i)}{\rho^2(i)} = \cos^2 (f_J^i - f_J, u_{\alpha J})$: relative contribution of factor α to $\rho^2(i)$

$\mathrm{ctr}\ (i, \alpha) = \frac{f_i F_\alpha^2(i)}{\lambda_\alpha}$: relative contribution of element i to moment of inertia λ_α

For each factor α the program also prints:

$\tau_\alpha = \frac{\lambda_\alpha}{\sum_\alpha \lambda_\alpha}$: proportion of variance, and $\sum_{q=1}^{\alpha} \tau_q$: cumulated variance of ESPRO (in lines) with n variables (acoustical criteria or subjective characteristics) in columns. Correlations between variables allowed configurations to be sufficiently described by a smaller number of factors $n' < n$. MCA built and n'-dimensional space, in which configurations and initial variables were plotted (definitions about MCA are listed in Table 3). To reduce heterogeneous raw data, a logical coding was used. Each measurement scale was replaced by a partition into classes, and initial values were coded into a table of numbers $\{k_{i,j} = 0$ or 1, according to whether the value fit or not to the class$\}$. Interpretation of each axis was based on elements which give the strongest contribution to its variance. Proximities between elements into the new space were related to their correlations in such a way that the latter can be analysed by visual representation of each plane. Lateral localisation was not analysed in the present study. Three analyses were performed (Table 4); (1) subjective characteristics with headphones and loudspeakers listening; (2) interindividual perception of space with headphones; (3) interindividual perception of space with loudspeakers.

TABLE 4
Analysis

Analysis	*Variables*	*Space*
1	Six subjective characteristics with headphones and loudspeakers (average values over subjects)	Common subjective space
2	Six subjective characteristics for each of the 11 subjects with headphones listening	Interindividual space with headphones
3	Six subjective characteristics for each of the 11 subjects with loudspeakers listening	Interindividual space with loudspeakers

Variances of Axes

Axes	*Analysis 1 (Common space)*		*Analysis 2 (Headphones)*		*Analysis 3 (Loudspeakers)*	
	Var. (%)	*Cum. var. (%)*	*Var. (%)*	*Cum. var. (%)*	*Var. (%)*	*Cum. var. (%)*
1	39·205	39·205	44·934	44·934	7·330	7·330
2	15·008	54·214	19·949	64·883	6·614	13·945
3	9·782	63·995	15·215	80·134	5·055	19·000
4	8·215	72·210	10·258	90·392	4·864	23·864
5	5·679	77·889	5·403	95·796	4·392	28·256
6	5·145	83·035	4·204	100·00	4·228	32·484

Proportion of variance (Var.) and cumulated variance (Cum. var.) (see Table 3) are listed for the six first axes of each analysis.

3 RESULTS

3.1 Analysis 1: Relations between subjective characteristics and geometrical parameters for each listening system (common subjective space)

3.1.1 Structure of common subjective space
Variances of axes are listed in Table 4. The six first axes of common space accounted for 83% of total variance. Axis 1 was related to small and large values of subjective reverberation time (RT), (see Fig. 1(a)). Medium and large values of subjective RT were respectively located on the positive and negative side of axis 2. Values of preference were distributed on axis 3 (see Fig. 1(b)). Small and medium values of preference were located on each side of axis 4 (Fig. 1(b)). Figure 1(c) shows that axis 4 was also related to large values of spatial impression and axis 5 to large values of apparent room size. Values of depth perception were distributed on axis 6 (see Fig. 1(d)).

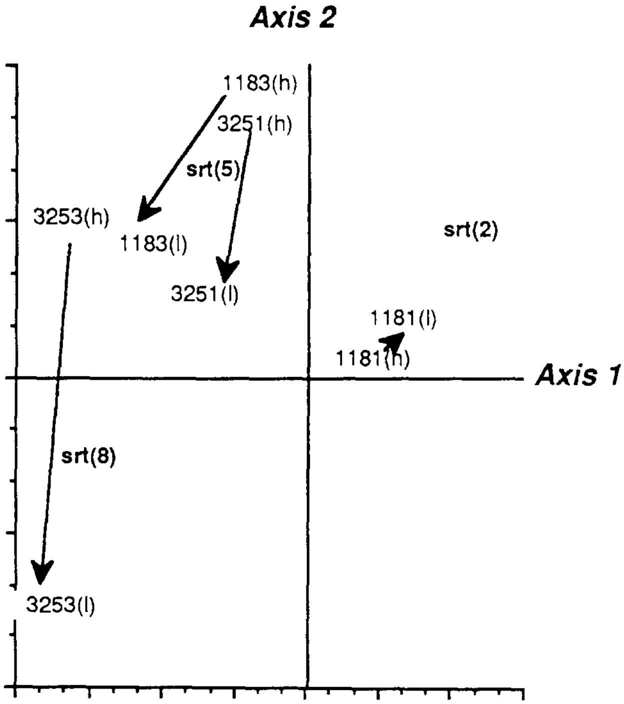
3.1.2 Differences between headphones and loudspeakers
Subjective RT. Figure 1(a) (plane 1–2) reveals that, with loudspeakers, large values of subjective RT were more related to distant positions in the room than with headphones.

Depth perceptions. Figure 1(d) (plane 5–6) reveals that perceived depth was smaller with headphones than with loudspeakers.

Spatial impression. Figure 1(c) reveals (plane 4–5) that large values of spatial impression were correlated to the same configurations (low ceiling—very diffusing) with headphones and loudspeakers. However, correlation values were greater with headphones.

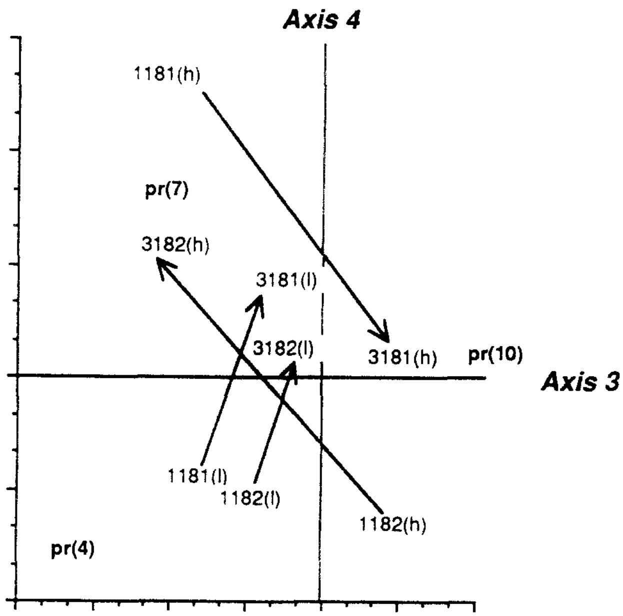
Apparent room size. Figure 1(c) shows that large values of apparent room size were rather correlated to P3—low ceiling-diffusing configurations with headphones, and to P3—high ceiling-absorptive configurations with loudspeakers listening.

Preference. Figure 1(b) (plane 3–4) illustrates the effects of ceiling height on preference. It appears that in near positions P1 and for each listening system, preference was emphasised by decrease of ceiling height. It was also the case in off-centre positions P2 with loudspeakers, but not with headphones for which high ceiling configurations were preferred. Differences between headphones and loudspeakers listening for a same configuration are represented in Fig. 1(b'). They appeared to be more

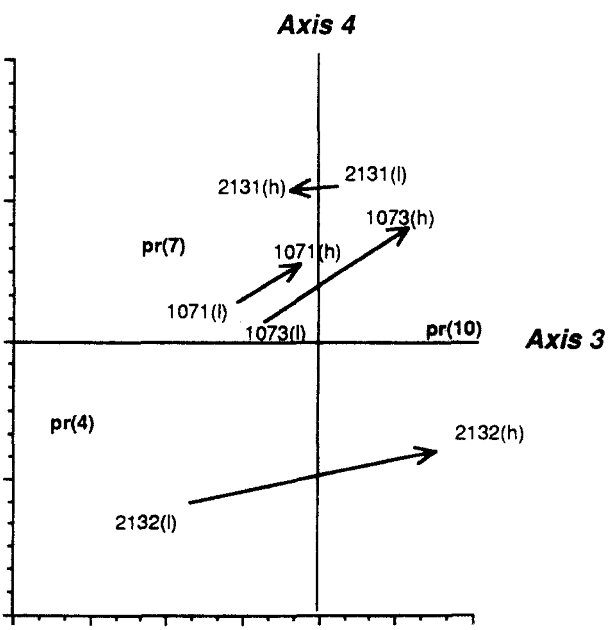


(a)

Fig. 1. Common subjective space (Analysis 1). Same configurations listened with headphones and loudspeakers are respectively represented by (h) and (l). Values of subjective characteristics were numbered from weak to strong values (see Table 2). Some values and configurations were not represented for clarity of figure. (a): Plane 1–2. Axis 1: weak/strong values of subjective reverberation time (respectively srt(2) and srt(8)). Axis 2: strong/median values of subjective reverberation time (srt(8) and srt(5)). With loudspeakers, great values of subjective reverberation time were particularly related to distant room positons (ex: 3253 (l)). (b) and (b'): Plane 3–4. Axis 3: weak/strong values of preference (resp. pr(4) and pr(10)). Axis 4: weak/median values of preference (resp. pr(4) and pr(7)). (b) Illustrates effect of ceiling height on preference. Preference was emphasised by decreasing ceiling height (ex: 1181–3181), except for excentred position P2 for headphones (1182–3182). (b') Illustrated differences between headphones listening (h) and loudspeaker listening (l). Differences were greater for distant room positions (ex: 1232). Preference was greater with headphones (h). (c): Plane 4–5. Axis 4 was related to great values of spatial impression (si(3)) and axis 5 to great values of apparent room size (rs(3)). Great values of apparent room size were related to low ceiling diffusing configurations with each listening system (ex: 3251, 3253). Great values of apparent room size were rather related to low ceiling-diffusing configurations (ex. 3183) with headphones and high ceiling-absorptive configurations (ex: 1013) with loudspeakers. (d): Plane 5–6. Axis 6: weak/strong values of depth perception (dp). All configurations were perceived as more distant with headphones (h).

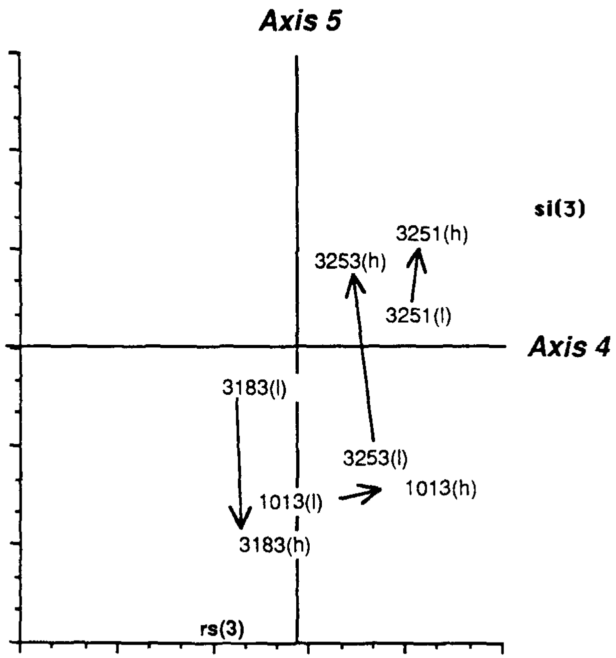


(b)

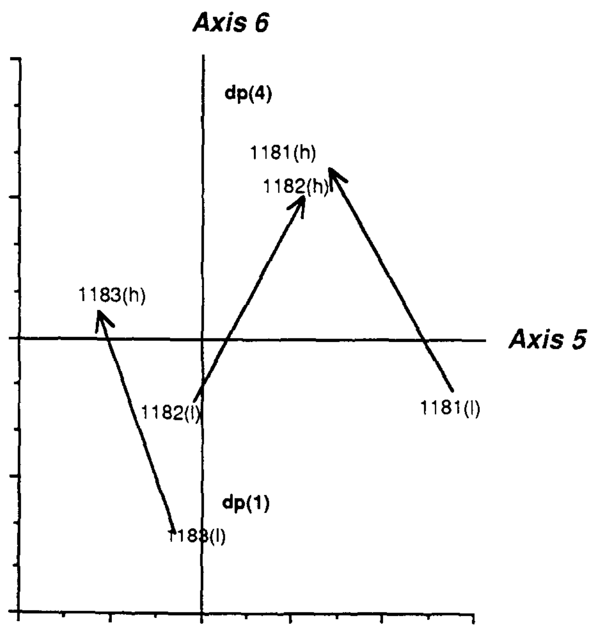


(b')

Fig. 1—*contd.*



(c)



(d)

Fig. 1—*contd.*

important in positions P2 and P3 than in near positions P1. Figure 1(b') also shows that, for any configurations, preference was greater with headphones than with loudspeakers.

3.2 Analyses 2 and 3: interindividual differences with each listening system

The three first axes of analysis with all subjects accounted for 80% of total variance with headphones and for 32% with loudspeakers (see Table 4). Axis 1 was related to, respectively, large and small values of depth perception with headphones and loudspeakers. With each listening system, axis 2 was related to large values of apparent room size. Axis 3 was related to small values of depth perception with headphones and to small values of apparent room size with loudspeakers.

3.2.1 Headphones listening

Results are presented in a previous paper.[1] Contributions of all subjects were nearly identical on axis 1, but different on axes 2 and 3: five subjects had their main contribution to axis 2 and four subjects to axis 3. Figure 2(a) and 2(b) illustrates the distribution of subjective evaluations for one subject of each group. Only one subject had no main contribution to any subjective axis.

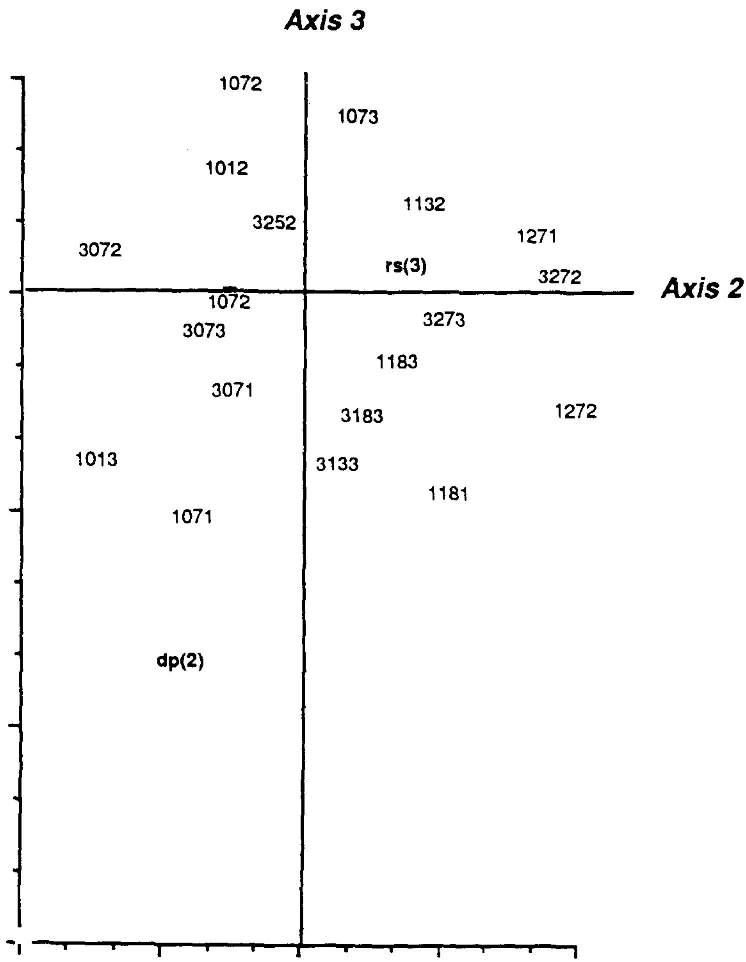
3.2.2 Loudspeakers listening

Three subjects had a main contribution to axis 1, six subjects to axis 2, and two subjects to axis 3, so judgements of subjects were scattered in planes 1–2 and 2–3. Figure 2(c) shows the distribution of subjective evaluations for a subject which contributed to axis 1.

4 ANALYSIS AND INTERPRETATION

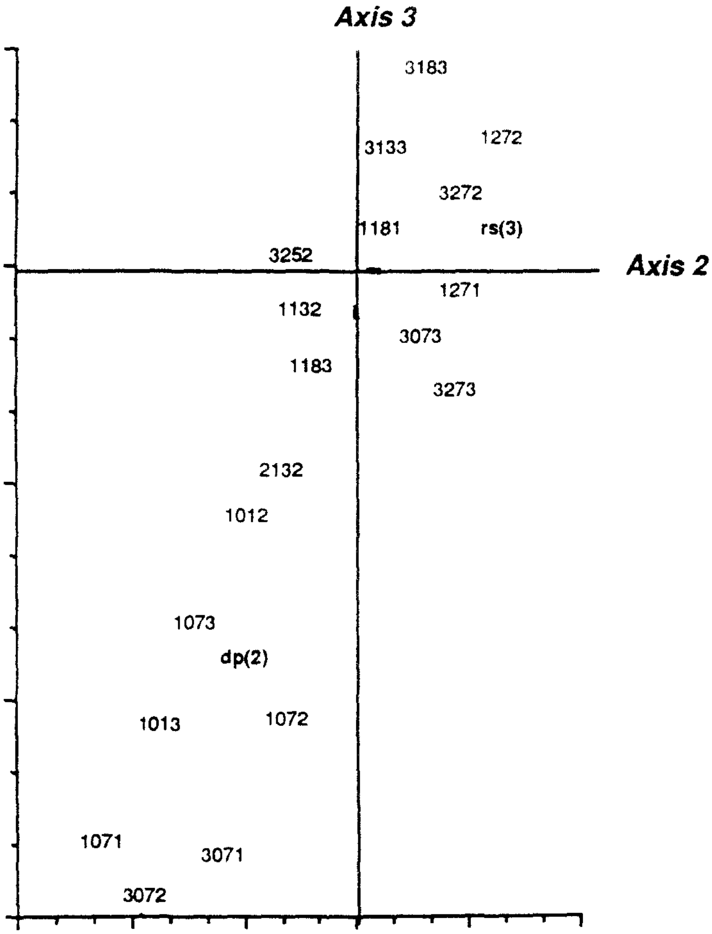
4.1 Differences between headphones and loudspeakers

The greatest differences between each listening system were obtained with distant positions, and consequently with small values of dir/rev ratio (see Fig. 1(b')). Such an observation could reveal a greater effect of the listening system for reverberant than for direct sound, which mainly differ by their spatial properties. Localisation of direct sound is related to the respective positions of source and listener, whereas diffuse sound is composed of randomly localised reflections. This stochastic character remains unchanged through the dummy head–headphones chain, when characteristics of channels are respected. However, through loudspeakers,



(a)

Fig. 2. Interindividual space. (a) and (b): Headphones listening (Analysis 2). Evaluations of configurations were mainly scattered on axes 2 and 3. Axes 2 and 3 were respectively correlated to low values of depth perception (dp) on the negative side, and to high values of apparent room size (rs) on the positive side. (a) Illustrates distribution of configurations for a subject which contributed to axis 2. (b) Illustrates distribution of configuration for a subject which contributed to axis 3. (c) Loudspeaker listening (Analysis 3). Axes 1, 2 and 3 were respectively correlated to weak values of depth perception, strong and weak values of apparent room size. Figure represents distribution of configurations for a subject which contributed to axis 1.



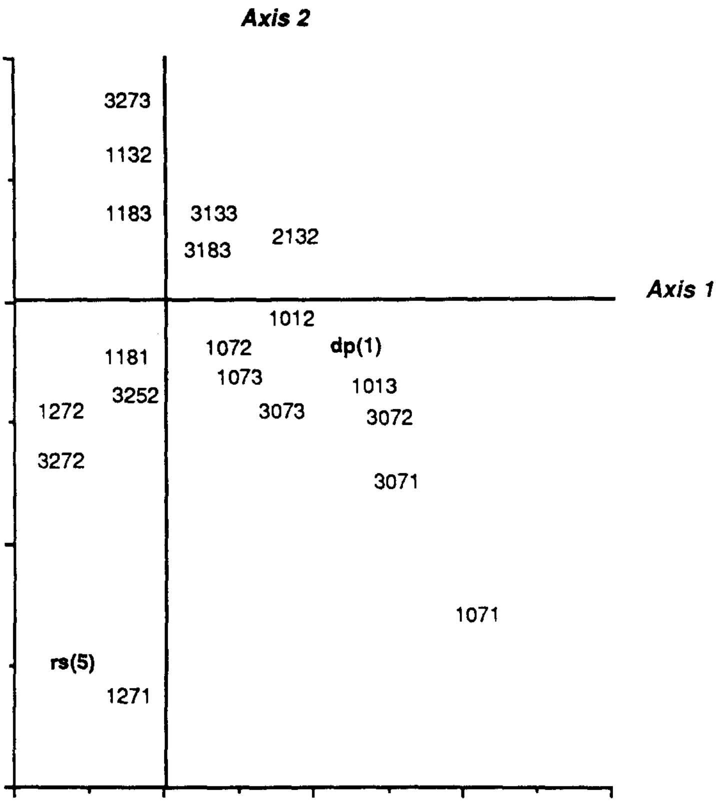
(b)

Fig. 2—*contd.*

some constraints on sound direction are determined by loudspeakers' positions. If spatial stochastic character is assumed to be a criterion of diffuse field perception, the latter might be underevaluated with loudspeakers. Results could not be analysed without taking account of this observation.

4.1.1 Subjective RT

Through loudspeakers, large values of subjective RT were particularly correlated to distant positions in the room. Subjective RT is usually known to be positively related to early decay time (EDT) and measured reverberation time (RT60), but also negatively to the dir/rev ratio.[1] If diffuse sound is assumed to be less well perceived with loudspeakers,



(c)
Fig. 2—*contd.*

subjective reverberation could be in proportion more negatively related to direct energy (and consequently to position in the room).

4.1.2 Depth perception

According to Coleman, depth perception increases with decreasing dir/rev ratio.[5] Small values of depth effect with loudspeakers could be explained by the poor impact of reverberant energy in the perception of depth.

4.1.3 Spatial impression

Values of spatial impression appeared less important when perceived through loudspeakers. This characteristic is known to be related to early lateral reflections level.[6] Localisation of reflections are correctly represented through headphones, but are very dependent on loudspeakers positions. This latter fact could be related to the small value of spatial impression.

4.1.4 Apparent room size

Through headphones and loudspeakers, the apparent room size was respectively correlated to low ceiling diffusing and high ceiling absorptive configurations (consequently, respectively high and low values of sound level). It was known that apparent room size increased with increasing reverberation time.[1] That could explain the correlation of spatial impression with reverberation level through headphones. However, it can also be noted that in real concert halls, total sound level decreases with increasing room size. If diffuse sound is assumed to be perceived less through loudspeakers, increase of apparent room size could be mainly related to sound level decrease.

4.1.5 Preference

For off-centre listener positions (P2), low ceiling configurations were preferred through loudspeakers, but not through headphones. Large values of preference are known to be related to median values of RT60 and dir/rev ratio.[1] Dir/rev was reduced by decreasing ceiling height, and consequently shifted from high to middle values in P1 and from low to very low values in P2, so preference increased in P1 and reduced in P2. However, if increase of diffuse sound is assumed to be mainly perceived as an increase of total sound level when observed through loudspeakers, preference could be more marked in any position in the room.

Greater values of preference with headphones could be explained on the basis of the two channels characteristics when sound samples are recorded through a dummy head, which produces more natural listening.

4.2 Interindividual differences

The main result concerning interindividual variations appeared in the fact that the six first axes totalled less variance with loudspeakers (32%) than with headphones (80%). Such an observation reveals greater inter-individual differences for the loudspeaker system. However, the clearest perception of diffuse sound field through headphones remains in agreement with Schroeder's[8] assertion that reverberation is a factor of consensus between subjects.

5 CONCLUSION

Some significant differences between headphones and loudspeakers in sound spatial perception are presented. The effect of the listening system was particularly relevant for distant positions in the room: (i) subjective RT was higher and depth perception lower with loudspeakers; (ii) preference was stronger with high ceiling configurations when observed through loudspeakers, but not through headphones. It also appeared that with loudspeakers and for any position: (i) large values of spatial impression were less well perceived, (ii) interindividual differences were higher and (iii) large values of apparent room size were more related to low values of reverberant level.

Results were interpreted as a less accurate representation of diffuse field spatial characteristics through loudspeakers, mainly due to the source positions. A perceptive study of concert halls by means of musical samples recorded in the room could be analysed from this point of view. Such an experiment could allow a better comprehension of the recording-diffusing chain effect on sound perception. The procedure could also be employed to compare different recording systems.

ACKNOWLEDGEMENTS

The authors gratefully thank IRCAM, who have supported the main part of this research, and particularly J.-P. Jullien and J.-D. Polack, who realised the physical measurements in ESPRO.

REFERENCES

1. Fischetti, A., Hemim, Y. & Jouhaneau, J., Relations between subjective spatialisation, geometrical parameters and acoustical criteria in concert halls. *Appl. Acoust.* **37** (1992) 233–47.
2. Fischetti, A., Jouhaneau, J. & Hemim, Y., La caracterisation perceptive des salles de concert: insuffisance des critères traditionnels et proposition d'un modèle basé sur la distribution des premières réflexions. In *Proc. 13th Int. Congress on Acoustics*, Vol. 2, Belgrade, 1989, pp. 165–8.
3. Jullien, J.-P., Correlations among criteria of room acoustics. In *Proc. 12th Int. Congress on Acoustics*, E4-9, Toronto, 1986.
4. Benzecri, J.-P., *L'Analyse des données*, Tome 2: *L'Analyse des correspondances*, ed. Dunod. Paris, 1976.
5. Coleman, P. D., An analysis of cues to auditory depth perception in free space. *Psychol. Bull.*, **60** (1963) 302–15.
6. Barron, M., The subjective effects of first reflections in concert halls—The need for lateral reflections. *J. Second Vibrat.*, **15**(4) (1971) 475–94.
7. Morimoto, M. & Maekawa, Z., Effects of low frequency components on auditory spaciousness. *Acoustica*, **66** (1988) 190–6.
8. Schroeder, M. R., Gottlob, D. & Siebrasse, K. F., Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters. *J. Acoust. Soc. Amer.*, **56**(4) (1974) 1195–201.