The Impact of Sensory Evaluation of Sound - past, present and future.

Sensometrix 2012, Agrocampus, Rennes, France

12 July 2012

Nick Zacharov
nick.zacharov@genelec.com

GENELEC

Outline

- Domains of sound
- Short historical review – some key works
- State of the field
  - Some examples over a decade or so
  - Where are the standards?

- Successes / challenges and the future

Domains of Sound

The facets of sound

- Large number of facets of application below the level of sound, each having their own specific needs

- Three of the most common include:
  - Audio (reproduced sound)
  - Telecommunication (speech)
  - Product sound (noise)

The audio application

- Is a fairly common case
  - How to faithfully reproduce sound
    - Mostly for the purposes of entertainment
      - Hifi
      - Home Theatre / surround sound
    - The challenge lies is defining an absolute point of reference
      - How a producer heard the sound?
      - How the original was experienced in a concert hall?
    - **Focus on enjoyment**

The (tele)communication challenge

- At face value a trivial case
  - Replicate face to face conversation
    - **Focused on communication**
    - Speaker identification secondary
    - Sound quality tertiary
    - Other artefacts come later
      - Noise, distortion, etc…

Product sound

- Product sound is the overall domain where the sound is a secondary, perhaps unwanted, aspect of the product
  - Car noise
  - Wind turbine noise
  - Sounds of white products
    - Vacuum cleaner, washing machine
  - **Combination of functional sound and controlling annoyance**

Focus for today

Is on audio (sound reproduction) and (tele)communication applications

Short historical overview

Some key works

GENELEC

Some background

- Fairly well agreed that sensory science roots are in the field of food science during the 50’s
  - E.g. 9-point hedonic scale (Peryam and Pilgrim 1957)
- This has lead to the key developments in the field

David R. Peryam, Ph.D. 1915 - 1992
"The Father of Sensory Science"

And for sound..?

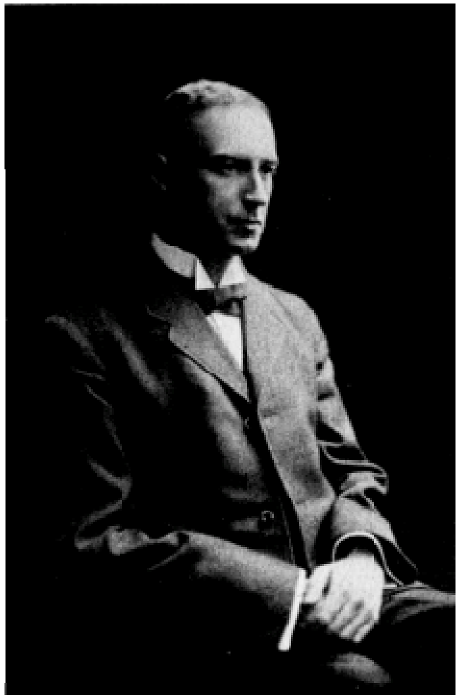
- Things started fairly early on
  - 1900 with concert halls

  - The question of interest was *“why do concert halls sound so different”* and what are the perceptual attributes that are needed for *“good”* acoustics

The pioneering work of Sabine

- Motivation
  - What makes a hall sound acoustically acceptable (1900)?

- Sabine identified some key *characteristics*
  - Loudness
  - Distortion of complex sounds: interference and resonance
  - Confusion: reverberation, echo and extraneous sounds



Wallace Clement Sabine
1868-1919

The father of concert hall acoustics

Reverberation

- Understanding the role of reverberance led to the development Sabine's equation

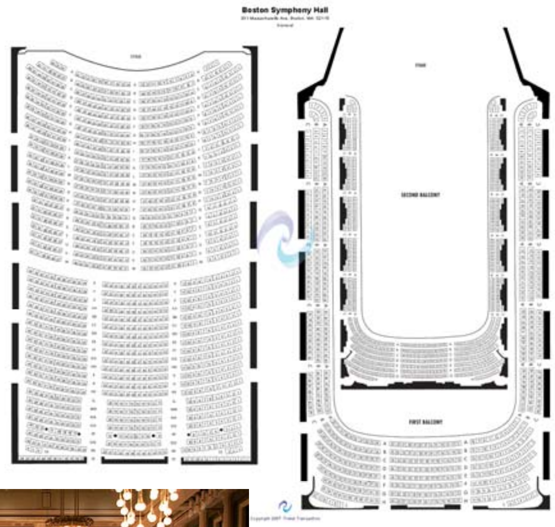
$$T = 0.161\ V/A$$

T = the reverberation time

V = the room volume

A = the total absorption area

- And subsequently some classical concert halls
  - E.g. Boston Symphony Hall



The early years (1962-1975)

- Focused on concert hall sound quality
  - **Beranek** (1962)
    - Interview based attribute development of 54 concert halls (live)

    → 8 continuous scales developed
  - **Hawkes and Douglas** (1971)
    - 14 provided attribute scales, in 4 halls at 4 locations (live)
  - **Nakayama et al.** (1971)
    - MDS analysis of a single hall, using 1-8 speakers
    - → 3 latent dimensions interpreted
  - **Wilkens** (1975)
    - Binaural recording and reproduction of 6 concert halls

    → Rated on provided 16 scales → 3 latent dimensions

1975 - 1995

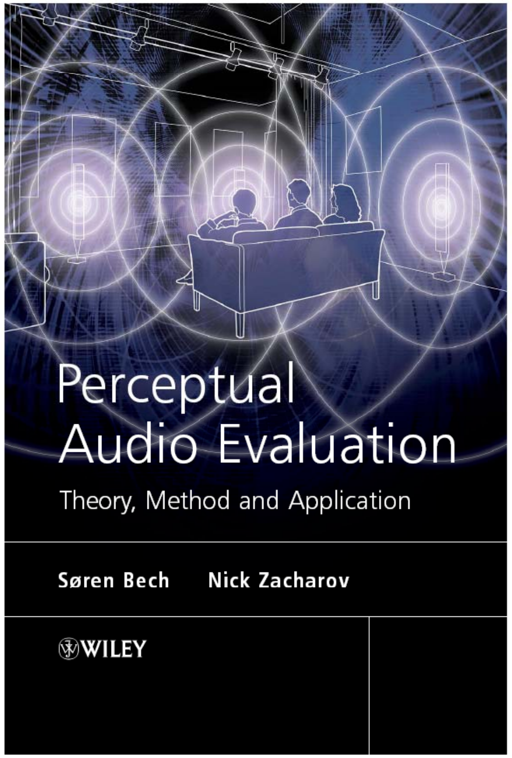
- More concert hall studies
  - Schoeder (1974), Lavandier (1989), Kahle (1995)

- The start of sensory sound reproduction research
  - Staffeldt (1974), Gabrielsson (1974,1979), Toole (1982, 1985 1986),

- The beginning of communication quality research
  - Voiers (1977, 1983), Quackenbush (1988)

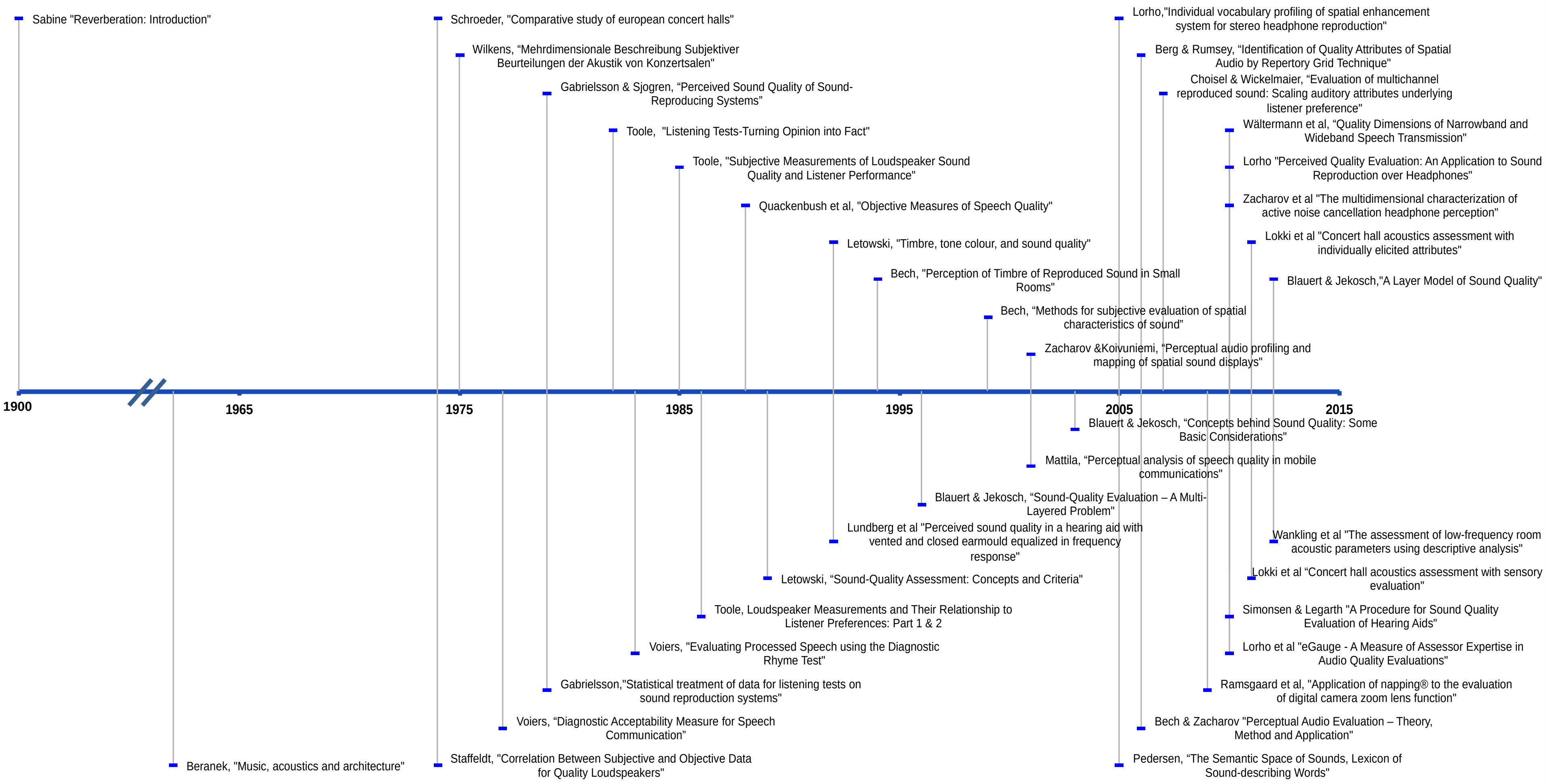
1995 - now

- Shift towards spatial sound reproduction
  - Both headphone and loudspeaker
  - Zacharov (2001), **Lorho (**2005, **2010)**, Berg & Rumsey (2006), Choisel and F. Wickelmaier (2007), Wankling (2012)
- More indepth mobile telecommications focus
  - **Mattila (2001)**, Wältermann (2010)
- Even more on concert halls
  - **Lokki et al (2011)**
- Some multimodal research
  - Audio-Visual perception, Audio-haptic, etc…
- Other applications
  - Hearding aids, active noise contrpl, Sound branding, emotions, etc.



2006

A sparse timeline of perceptual evaluation in audio



State of the field

GENELEC

Summary of early concert hall works (1962-1975)

Author	Stimulus characteristics	Perceptual evaluation method	Descriptors	Perceptual dimensions
Beranek (1962)	- Live music - 54 concert halls	- Semantic description based on inteviews with musicians, conductors and music critics. - Quantitative grading with provided attributes developed by the author	18 qualitative terms reduced to 8 continuous scales: *Intimacy, Liveness, Warmth, Loudness of direct sound, Loudness of reverberant sound, Balance and blend, Diffusion* and *Ensemble*.	
Hawkes and Douglas (1971)	- Live music - 4 concert halls - 4 different positions	- Provided attributes - Quantitative grading - 4 (unspecified) listeners - Factor analysis	14 continuous scales (+ 2 hedonic scales) selected from Beranek (1962) study[1]	5 latent dimensions interpreted as *Reverberance, Balance and Blend, Intimacy, Definition* and *Brilliance.*
Nakayama *et al.* (1971)	- Live music recorded with 8 microphones in a concert hall - Sound reproduced with 1 to 8 loudspeakers in an anechoic room	- MDS (plus preference rating) - 10 naïve[2] listeners	n.a.	3 latent dimensions interpreted as *Depth of image sources, Sensation of fullness* and *Sensation of clearness.*
Wilkens (1975)	- Binaural recording of 3 classical music selections in 6 concert halls - Static binaural replay over headphones	- Provided attributes - Quantitative grading - 40 listeners - Factor analysis	16 continuous scales (+ 3 hedonic scales)[3]	3 latent dimensions interpreted as *Strength and extension of source, Distinctness or clarity* and *Timbre of the total sound*.

Lorho 2010

Concert hall attributes

Sabine	Beranek	Wilkens	Lavandier	Kahle
~1900	1962	1975	1989	1995
Loudness	Reverberance	small – large	absent – présent (absent – present)	puissance (strength)
Distortion of complex sounds: interference and resonance	Loudness	pleasant – unpleasant	faible – énergique (weak – strong)	révébérance (reverberance)
Confusion: reverberation, echo and extraneous sounds	Spaciousness	unclear – clear	brouillé – net (muddy – clear)	balance générale (overall balance)
	Clarity	soft – hard	lointain – proche (far – near)	contraste (contrast)
	Intimacy	brilliant – dull	sec – révébérant (dry – reverberant)	puissance dans les graves (low frequency strength)
	Warmth	rounded – pointed	plat – contrasté (flat – contrasted)	puissance dans les aiguës (high frequency strength)
	Hearing of stage	vigorous – muted	coulant – heurté (flowing – halted)	pâteux (pasty)
		appealing – unappealing	dur – doux (hard – soft)	heurté (halted)
		blunt – sharp	neutre – intime (neutral – intimate)	
		diffuse – concentrated	sec – vivant (dry – lively)	
		overbearing – reticent	creux – chaud (hollow – warm)	
		light – dark	pauvre – brillant (weak – bright)	
		muddy – clear	impression d'espace (spatial impression)	
		dry – reverberant	largeur de la source (source width)	
		weak – strong		
		emphasized treble – treble not emphasized		
		emphasized bass – bass not emphasized		
		beautiful – ugly		
		soft – loud		

Staffeldt (1974)

- Early application to loudspeaker sound quality
- Aimed to subjectively and objectively characterize the (single) loudspeaker sound quality
  - Quite unique work at this time
  - Applied non-parametric statistics for the subjective data analysis
    - Bradley-Terry model
  - Subjective-Objective correlation was less successful

chamber quintet by P. Heise. Each music program had a duration of four minutes, and it was repeated as many times as needed by the testee (on the average four times for each paired comparison). The reason for using three

Table I. Testform for listening tests.

1 General preference I	19 Strong treble
2 Natural	20 Weak treble
3 Well-balanced	21 Rumbling
4 Nuanced	22 Hollow
5 Blurring	23 nasalized
6 Presence	24 Metallic
7 Distant	25 Peaked
8 Open	26 Shrill
9 Closed	27 Hard
10 Brilliant	28 Soft
11 Dull	29 Broad
12 Full-bodied	30 Narrow
13 Thin	31 Good directivity
14 Smooth	32 Poor directivity
15 Strong bass	33 Clear room-impression
16 Weak bass	34 Not clear room-impres.
17 Strong midrange	35 General preference II
18 Weak midrange	

The qualities 29 - 34 concern the stereofonic sound image.

Staffeldt (1974) – (2)

- Subjective-Objective correlation performed on frequency response data
- Grouping was logical
- Overall the analysis was less successful, as only the timbral data considered

21 rumbling	
22 hollow	
15 strong bass	16 weak bass
12 full-bodied	13 thin
Group 3	Group 10
26 shrill	
27 hard	28 soft
17 strong midrange	18 weak midrange
24 metallic	
Group 4	Group 11
6 presence	7 distant
14 smooth	
Group 5	Group 12
19 strong treble	20 weak treble
25 peaked	
Group 6	Group 13
29 broad	30 narrow
33 clear room-impression	34 not clear room-impression

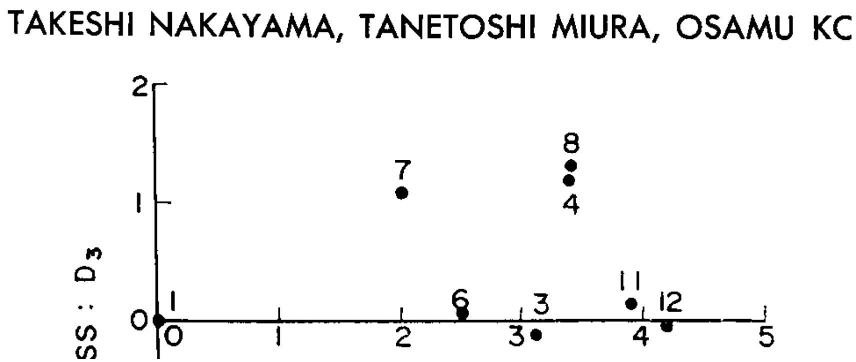
Qualities which are positioned opposite each other are contrasts.

what lower distortion of loudspeaker 5 is too weak a basis for drawing any conclusions.

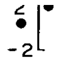
However, the sound pressure responses measured in the listening room could be related to the judgments under the two dimensions. This is most easily seen by transforming the sound pressure responses into a Zwicker graph with subjective coordinates [11]. The ordinate is here the

Nakayama (1971)

- Early MDS based analysis
  - Focused upon the number of speakers required for faithful sound reproduction
    - Very Early steps into multichannel sound



Nakayama 1971 – (2)

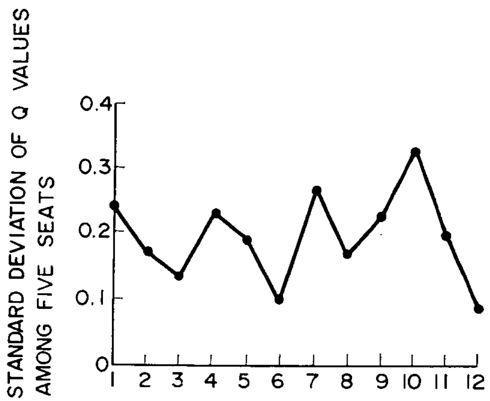


DEPTH OF THE IMAGE SOURCES : D_1

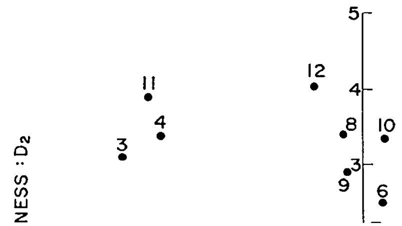
Fig. 13. Three-dimensional scales (D_1–D_3 plane).

As shown in Fig. 10, the three-dimensional sc
mit about 77 percent of the total variance of
ilarity matrix to be explained. The positive eig
ues indicate the variance explained in the Euclid
ce, and the negative eigenvalues show the variar
the similarity matrix that could not be explained
t space. The eigenvalues obtained from typical

SUBJ



When the reproduction condition is changed, there is a change in the area for equal preference value Q. Fig. 9 shows the standard deviation of the variation of Q values among the five listeners' positions shown in Fig. 4. The interpretation is that the smaller standard deviation corresponds to the wider equal-preference area.



Gabrielsson (1974)

- Analysis of loudspeaker, headphone and hearing aid sound quality
- Attribute rating
    - 60 selected adjectives from sound engineers, audiologists, hearing impaired subjects
        - Factor analysis
- Similarity rating
    - Subsequent INDSCAL analysis

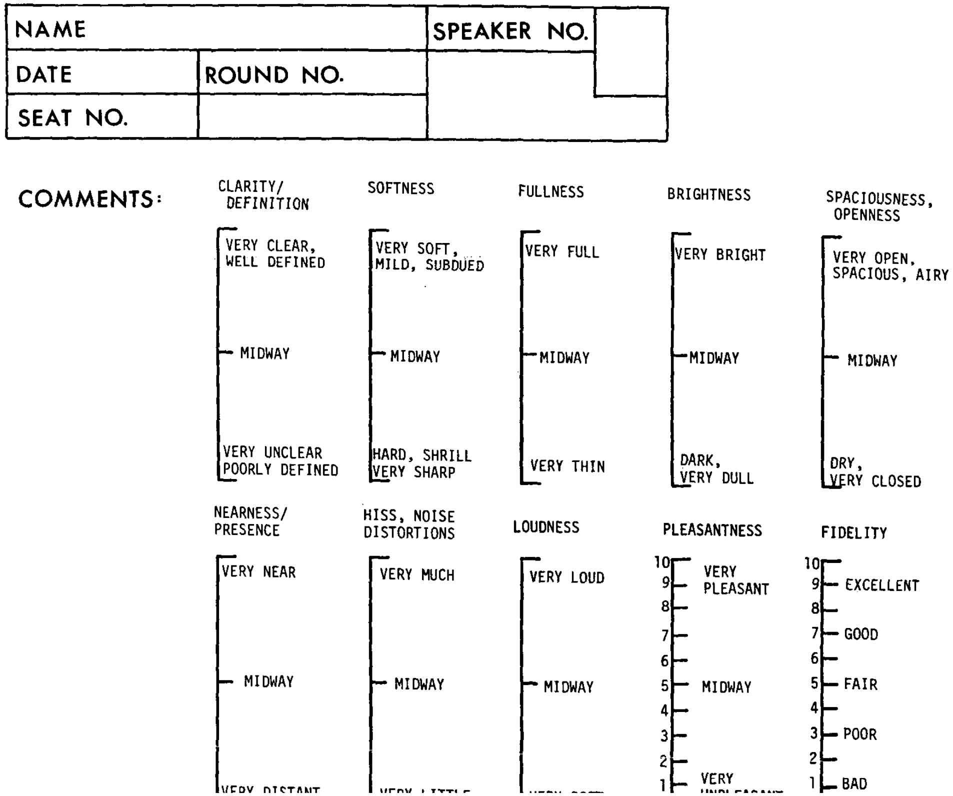
Grötig ("mushy/thick")	0.75	0.55	0.25
Gäll ("shrill")	0.25	0.88	−0.29
Hård ("hard")	0.27	0.92	−0.09
Ihålig ("hollow")	0.82	0.46	−0.16
Instängd ("closed/shut up")	0.95	0.18	0.10
Jämn ("uniform/smooth")	−0.67	−0.63	−0.04
Klar ("clear")	−0.87	−0.40	−0.16
Kontrastrik ("rich in contrasts")	−0.94	−0.17	0.02
Kraftig ("strong/loud")	−0.66	0.47	0.47
Kylig ("chilly")	0.43	0.68	−0.50
Ljus ("bright/light")	−0.23	0.38	−0.83
Luftig ("airy")	−0.91	−0.20	−0.16
Matt ("faint/feeble")	0.96	−0.02	−0.11
Mjuk ("soft")	−0.39	−0.85	0.11
Mullrande ("rumbling")	0.17	0.40	0.88
Mustig ("juicy/succulent")	−0.79	−0.11	0.48
Mörk ("dark")	0.08	−0.21	0.93
Naturtrogen ("true to nature")	−0.84	−0.47	0.01
Nära ("near")	−0.82	−0.35	0.17
Punktformig ("confined to a point")	0.78	0.30	−0.18
Påträngande ("obtrusive")	0.26	0.90	0.17
Ren ("clean/pure")	−0.85	−0.41	−0.08
Rumskänsla ("feeling of room")	−0.90	−0.29	0.01
Skarp ("sharp")	−0.16	0.85	−0.35
Skrapande ("scraping")	0.61	0.60	−0.03
Skrikig ("screaming")	0.40	0.87	−0.11
[illegible]	0.75	0.53	0.19

Gabrielsson (1974) – (2)

- 8 latent perceptual dimensions identified:
  - Clearness / distinctness
  - Sharpness / hardness – softness
  - Brightness – darkness
  - Disturbing sounds
  - Fullness – thinness
  - Feeling of space
  - Nearness
  - Loudness

Toole (1985)

- Further analysis on loudspeaker sound quality
  - Both mono and stereo systems
- Based upon provided attributes
  - Based partially upon Gabrielsson



NAME		SPEAKER NO.	
DATE	ROUND NO.		
SEAT NO.			

COMMENTS:

CLARITY/ DEFINITION: VERY CLEAR, WELL DEFINED — MIDWAY — VERY UNCLEAR POORLY DEFINED

SOFTNESS: VERY SOFT, MILD, SUBDUED — MIDWAY — HARD, SHRILL VERY SHARP

FULLNESS: VERY FULL — MIDWAY — VERY THIN

BRIGHTNESS: VERY BRIGHT — MIDWAY — DARK, VERY DULL

SPACIOUSNESS, OPENNESS: VERY OPEN, SPACIOUS, AIRY — MIDWAY — DRY, VERY CLOSED

NEARNESS/ PRESENCE: VERY NEAR — MIDWAY

HISS, NOISE DISTORTIONS: VERY MUCH — MIDWAY

LOUDNESS: VERY LOUD — MIDWAY

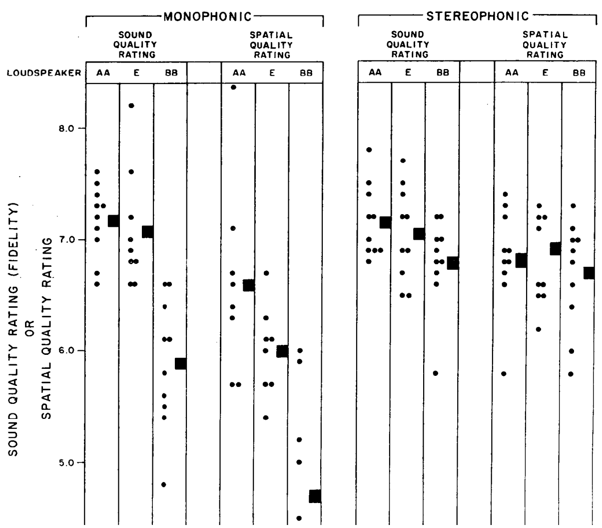
PLEASANTNESS: 10 9 VERY PLEASANT 8 7 6 5 MIDWAY 4 3 2 1 VERY

FIDELITY: 10 9 EXCELLENT 8 7 GOOD 6 5 FAIR 4 3 POOR 2 1 BAD

Toole (1985) – (2)

- An early work that studied to influence of assessor performance and difference between panels

LOUDSPEAKER SOUND QUALITY AND LISTENER PI



est the hypothesis, the data were simplified by ing the hearing threshold levels above and below Iz, and these new data were plotted against the ponding mean standard deviations for the indi- listeners (Fig. 8). The correlation coefficients est-fit" straight lines confirm that the more re- indicator of listener variability is the hearing t frequencies *below* 1000 Hz.

the hearing level should be a factor in subjective ents of sound quality was not entirely unexpected. he strong association should be with the hearing t the lower frequencies is surprising, especially he relationship is strongly developed over the of hearing level less than 20 dB, a range that in metric terms is regarded as representing acceptably l hearing [51]. Perhaps the possession of hearing adequate for speech communication, the con- nal criterion of normality in hearing, is insufficient especially critical task of judging sound quality. erally speaking, hearing loss at low frequencies mpanied by at least the same loss at higher fre- ies, although this is not invariably the case. In ular, if the hearing loss is purely conductive (that ess attenuation in the outer and middle ear), the

identifying conductive hearing loss as a signif terminant of performance in listening tests is sible without more comprehensive audiomet but hearing loss at low frequencies is a strong i [51].

The two listeners with low-frequency heari in the vicinity of 30 dB are obviously function a handicap. It is interesting that their perfo are somewhat better than would be predicted b fit line plotted through the remaining data p may be significant that these are among the perienced listeners in the group, and they ta in their judgment ability. Perhaps constant pra allowed them to overcome their handicaps pa

One common contributor to conductive hea is age, and it is reassuring (only in a scientifi to see in Fig. 9 a moderate positive correlation judgment variability and the age of the listen

2.2.3 *A Measure of the Error Due to Nuisa Variables*

Fig. 10 shows the mean standard deviation o

Letowski (1989)

- A good understanding of the perceptual facets
  - Heuristic versus diagnostic
  - Subjective versus objective

  - The value of the work was not fully embraced and

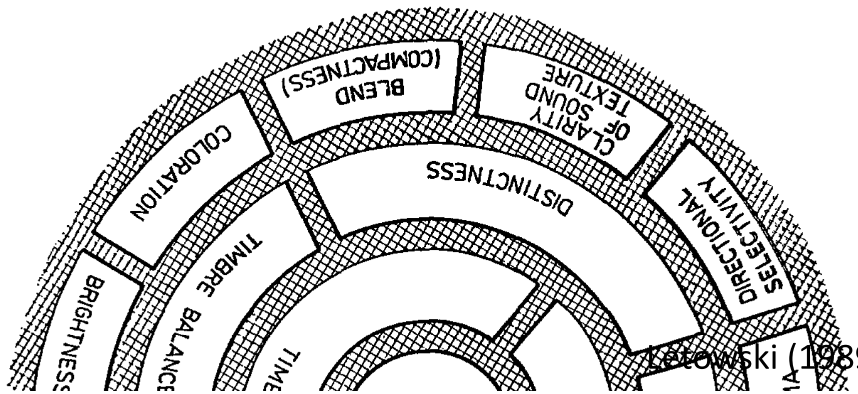
SOUND QUALITY ASSESSMENT: CONCEPTS AND CRITERIA

Letowski mural

- First attempt of structuring perceptual attribute of sound (audio)

imensional (parametric) assessment. Timbre constitu
which are assessed on the quantitative (domin
rity) scales, while quality components are assessed
ence scales [26].

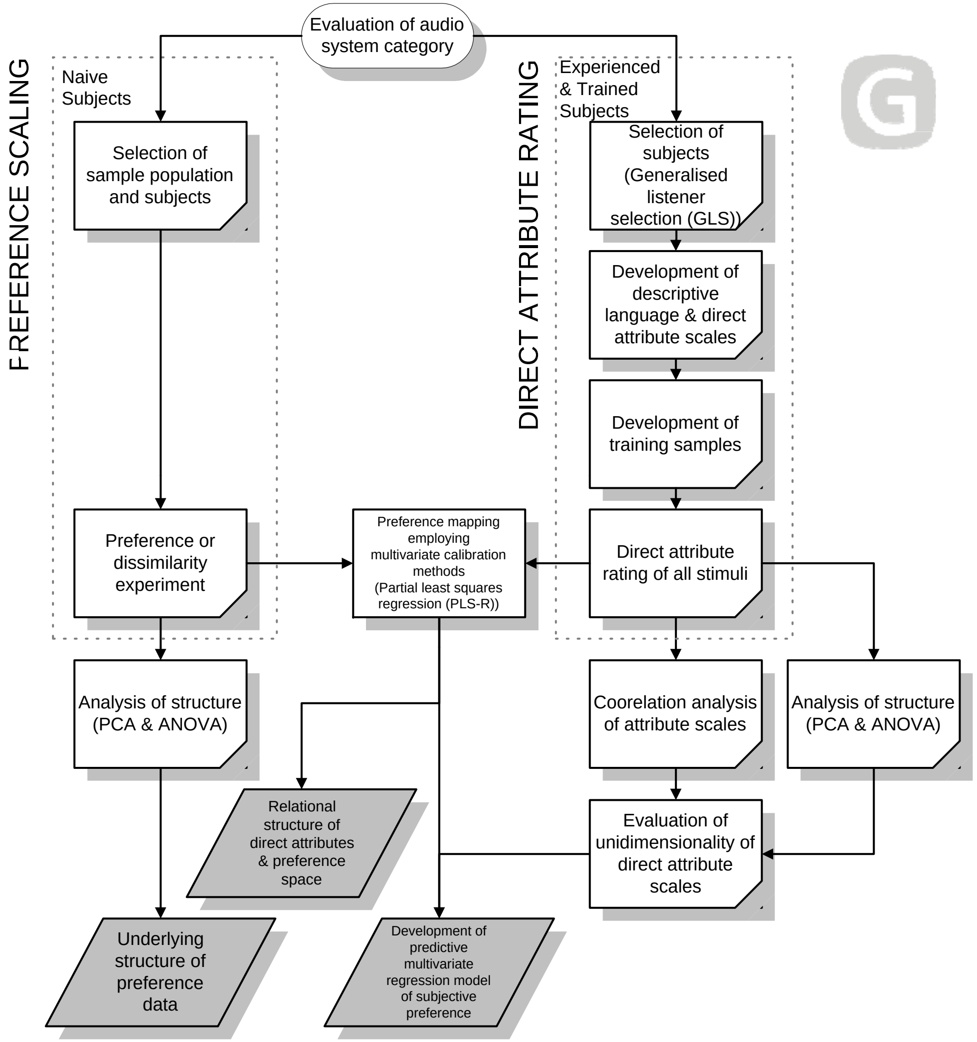
is one more practical reason to differentiate
and sound quality. Sound quality extends beyond ti
orates an impression of spaciousness. In such cases



Letowski (1989)

Zacharov (2001)

- Adaption of descriptive analysis and preference mapping for audio
- ADAM – audio descriptive analysis and mapping



Zacharov (2001) – (2)

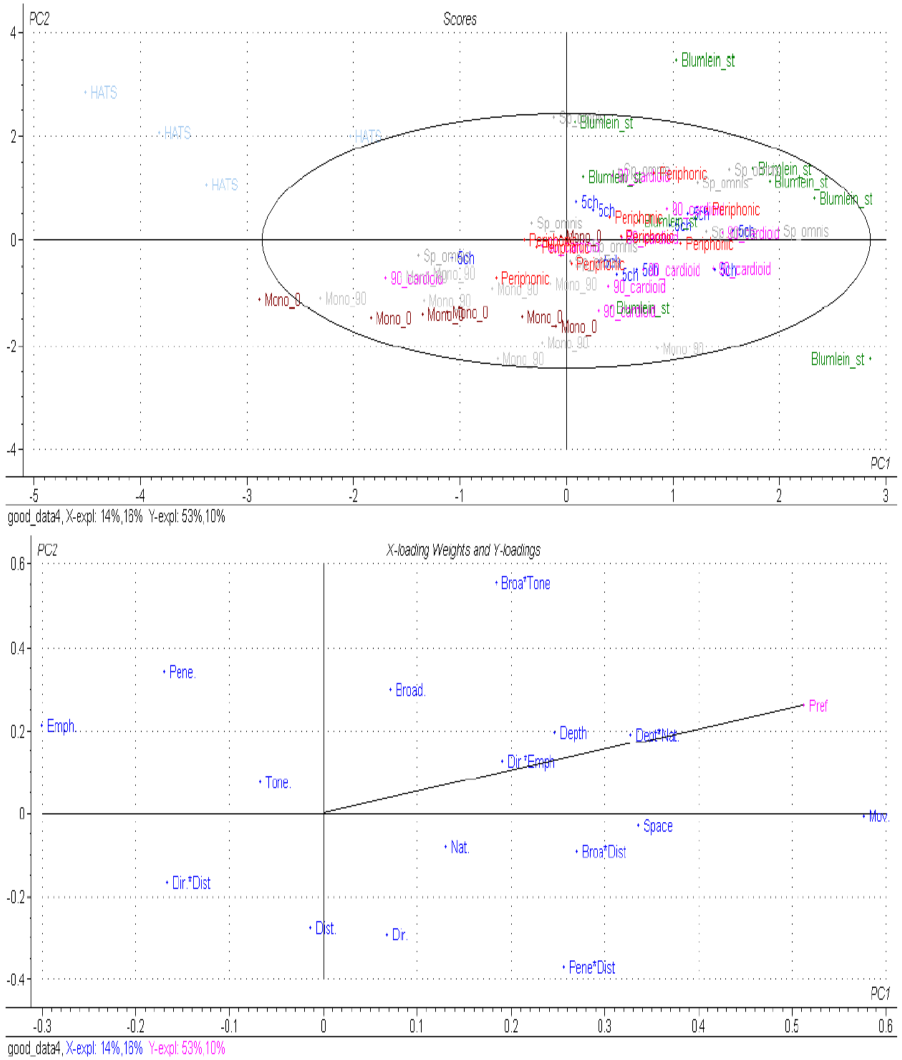
- Descriptive analysis and external preference mapping of spatial sound reproduction systems
    - 104 stimuli
  - Traditional attribute development
    - 12 attributes
      - 8 spatial
      - 4 timbral

Negative end-word	**Spatial attribute**	**Positive end-word**
Huonosti välittyvä *Ill-defined*	Suunnan tuntu *Sense of direction*	Hyvin välittyvä *Well defined*
Huonosti välittyvä *Ill-defined*	Syvyyden tuntu *Sense of depth*	Hyvin välittyvä *Well defined*
Huonosti välittyvä *Ill-defined*	Tilantuntu *Sense of space*	Hyvin välittyvä *Well defined*
Huonosti välittyvä *Ill-defined*	Liikkuvuuden tuntu *Sense of movement*	Hyvin välittyvä *Well defined*
Olematon *Non-penetrating*	Pistävyys *penetration*	Runsas *penetrating*
Lähellä *Close*	Tapahtumien etäisyys *Distance to events*	Kaukana *Distant*
Suppea *Narrow*	Laajuus *Broadness*	Laaja *Broad*
Epäluonnollinen *Unnatural*	Luonnollisuus *Naturalness*	Luonnollinen *Natural*

Negative end-word	**Timbral attribute**	**Positive end-word**
Vähäinen *Thin*	Täyteläisyys *Richness*	Täyteläinen *Rich*
Pehmeä *Soft*	Kovuus *Hardness*	Kova *Hard*
Neutraali *Neutral*	Korostuneisuus *Emphasis*	Korostunut *Emphasized*
Tumma *Dark*	Tummuus *Tone colour*	Kirkas *Bright*

Zacharov (2001) – (3)

- Preference mapping performed using PLS-R1 model
  - Eliptical model found best suited
  - Explained variance: 71% of preference data
  - Subjects are not in complete agreement – several "outliers"
  - Four principle components required
    - Contributions: 53%, 10%, 6%, 2%



Mattila (2001)

- Large scale sensory evaluation of digital mobile phone sound quality
  - Large number of methods applied
  - Interal / external preference mapping
  - Large stimulus set
  - Large data sets

Nr.	Processing	Symb.	Nr.	Processing	Symb.
1	Land line phone to mobile phone transmission, GSM-EFR	LME	23	GSM-EFR and transmission channel errors (*C/I* = 7 dB)	ES2
2	Land line phone to mobile phone transmission, GSM-HR	LMH	24	GSM-EFR and transmission channel errors (*C/I* = 4 dB)	ES3
3	Mobile phone to mobile phone transmission, GSM-EFR	MME	25	GSM-EFR and transmission channel errors (*C/I* = 3-15 dB)	DH3
4	Mobile phone to mobile phone transmission, GSM-HR	MMH	26	GSM-HR and transmission channel errors (*C/I* = 7 dB)	HS2
5	Land line phone to mobile phone transmission, GSM-EFR, HATS	VTE	27	GSM-EFR, muting of lost frames	EMT
6	Land line phone to mobile phone transmission, GSM-HR, HATS	VTH	28	Transducer	TRD
7	PCM, ITU G.711	PCM	29	Wideband additive noise	GNS
8	ADPCM, ITU G.726	ADP	30	Narrow band additive noise	NBN
9	LD-CELP, ITU G.728	LDC	31	Addition of echo	ECH
10	GSM-FR, ETSI GSM 06.10	FR	32	Clipping	CLP
11	GSM-HR, ETSI GSM 06.20	HR	33	Center clipping	CCL
12	GSM-EFR, ETSI GSM 06.60	EFR	34	Highpass filtering	HP
13	EVRC, TIA IS-127	EVR	35	Low pass filtering	LP
14	TETRA, TETRA 06.20	TRA	36	Bandpass filtering	BP
15	PDC-HR, RCR	PDC	37	Angular pole distortion	APD
16	U-ALWE (noise suppression)	UAW	38	Radial pole distortion	RPD
17	UDRC (dynamic range control)	DRC	39	Pole distortion	POD
18	U-ALWE and GSM-EFR	EAW	40	Narrow band frequency distortion	NBD
19	U-ALWE and GSM-HR	HAW	41	No Processing	ORG
20	GSM-EFR and GSM-EFR (tandem connection)	ETD	42	No Processing (SNR = 5 dB)	OR5
21	GSM-HR and GSM-HR (tandem connection)	HTD	43	GSM-HR, ETSI GSM 06.20 (SNR = 5 dB)	EF5
22	GSM-EFR with DTX	EDX	44	GSM-EFR, ETSI GSM 06.60 (SNR = 5 dB)	HR5

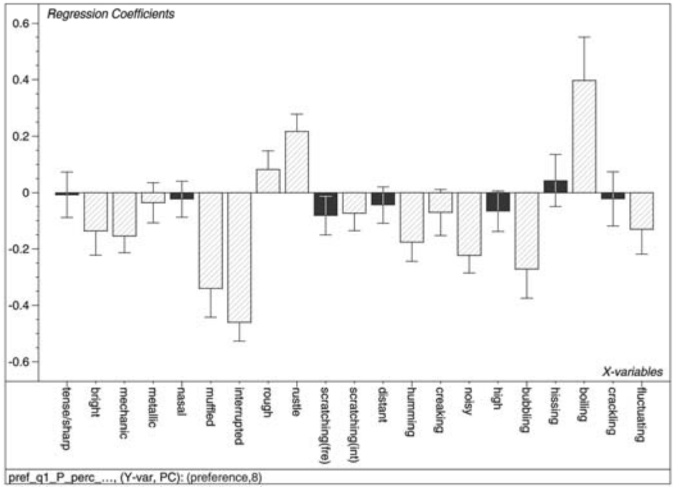
Mattila (2001) - (2)

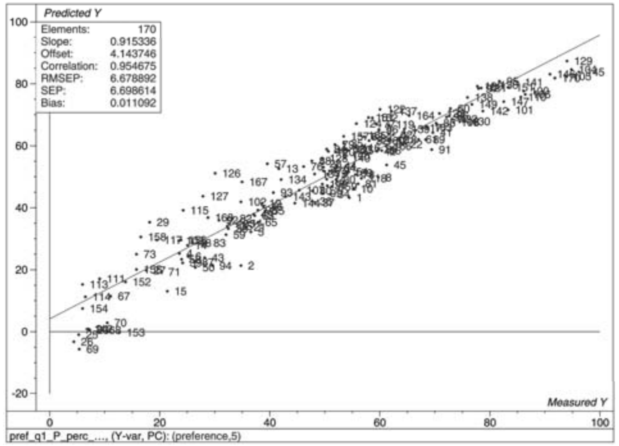
- 14 subjects
- Language develop required ~60 hrs
- 14000 words elicited
- 21 attributes and scaled defined
  - •Speech quality is one group
  - •Background noise forms a second group

	Nr.	Attribute in Finnish	Attribute in English
Speech	1	kireä / pingottunut	tense / sharp
	2	tumma - kirkas	dark - bright
	3	konemainen / mekaaninen	mechanic
	4	metallinen	metallic
	5	nasaalinen / inisevä	nasal / w hining
	6	selkeä - puuroutuva	clear - muffled
	7	ehjä - rikkonainen / katkonainen	smooth - interrupted
	8	käheä - karhea - rosoinen	rough
	9	särisevä	rustling
	10	särähtelevä / rasahteleva (frekvenssi)	scratching (frequency)
	11	särähtelevä / rasahteleva (voimakkuus)	scratching (intensity)
	12	avoin - etäinen / tukahtunut	open - distant
Background noise	13	humiseva	humming
	14	kitisevä	creaking
	15	kohiseva	noisy
	16	matala - korkea (taajuus)	low - high (frequency content)
	17	kupliva (frekvenssi)	bubbling (frequency)
	18	suhiseva	hissing
	19	poriseva	boiling
	20	ritisevä	crackling
	21	tasainen - vaihteleva - ryöpsähtelevä	steady - fluctuating

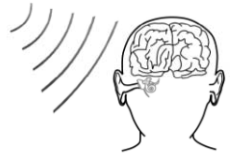
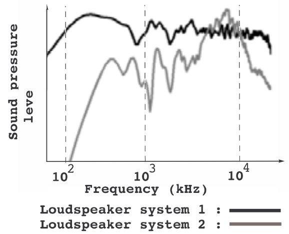
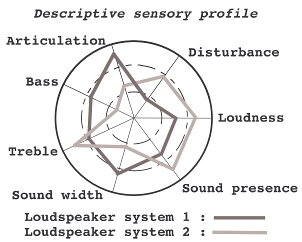
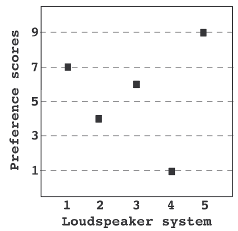
Mattila (2001) – (3)

- Preference mapping performed using PLS-R1 model
  - Vector model found best suited
  - Explained variance: 96% of preference data
  - 92% correlation between measured and modelled data
  - Subjects tended to be broadly grouped





Lorho (2005-**2010**)

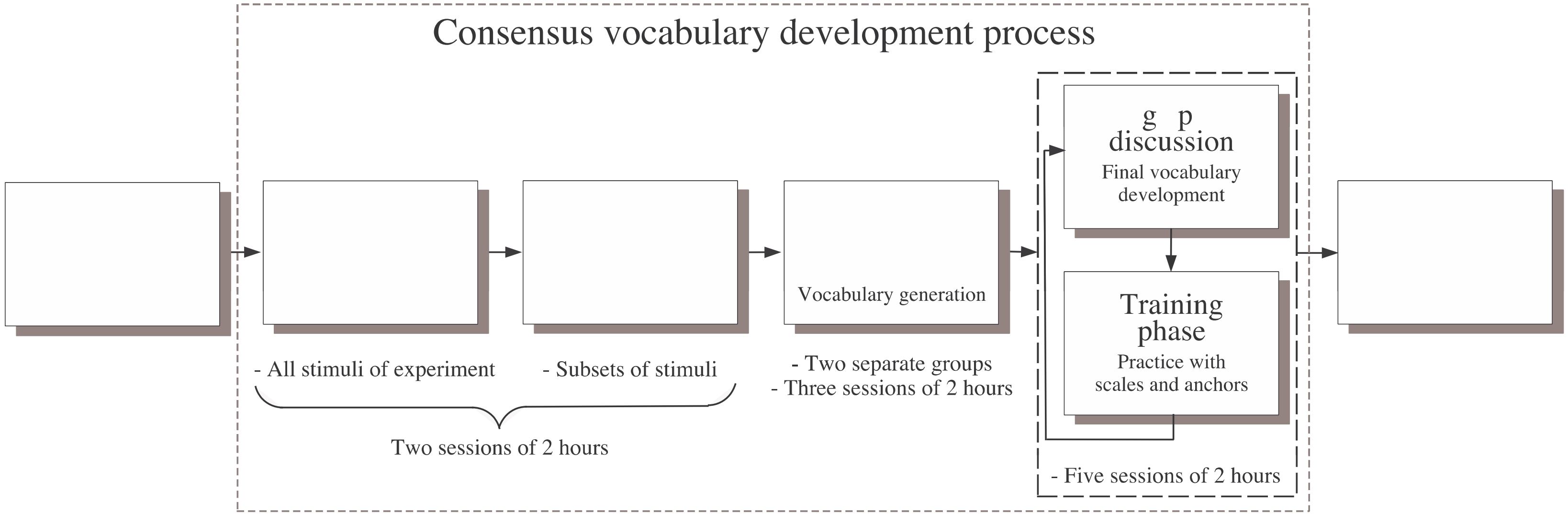
<table>
<tr><td colspan="2">Signal
(stimulus)</td><td colspan="3">
noise, speech or music</td></tr>
<tr><td colspan="2">Measurement
device</td><td>
microphone</td><td colspan="2">
auditory system</td></tr>
<tr><td rowspan="3">Stimulus
characterization</td><td>Domain</td><td>*Physical*</td><td>*Sensory*</td><td>*Affective*</td></tr>
<tr><td>Example</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Level of
Objectivity</td><td>*High*</td><td>*Medium*</td><td>*Low*</td></tr>
</table>

Lorho (2005-**2010**)

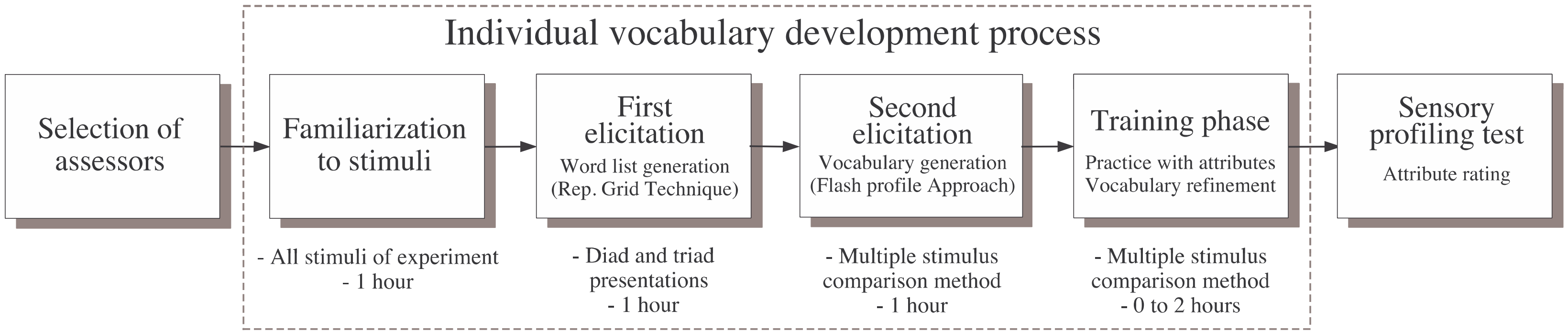
- Focused on headphone sound reproduction
- Applied and adapted a large number of techniques
- Consensus vocabulary development (CVP)
  - Extension of ADAM
- Individual vocabulary development (IVP)
  - Continuation of Flash profiling, etc.

- Indepth comparison on the two methods and their performance
- Deep understanding of headphone sound quality

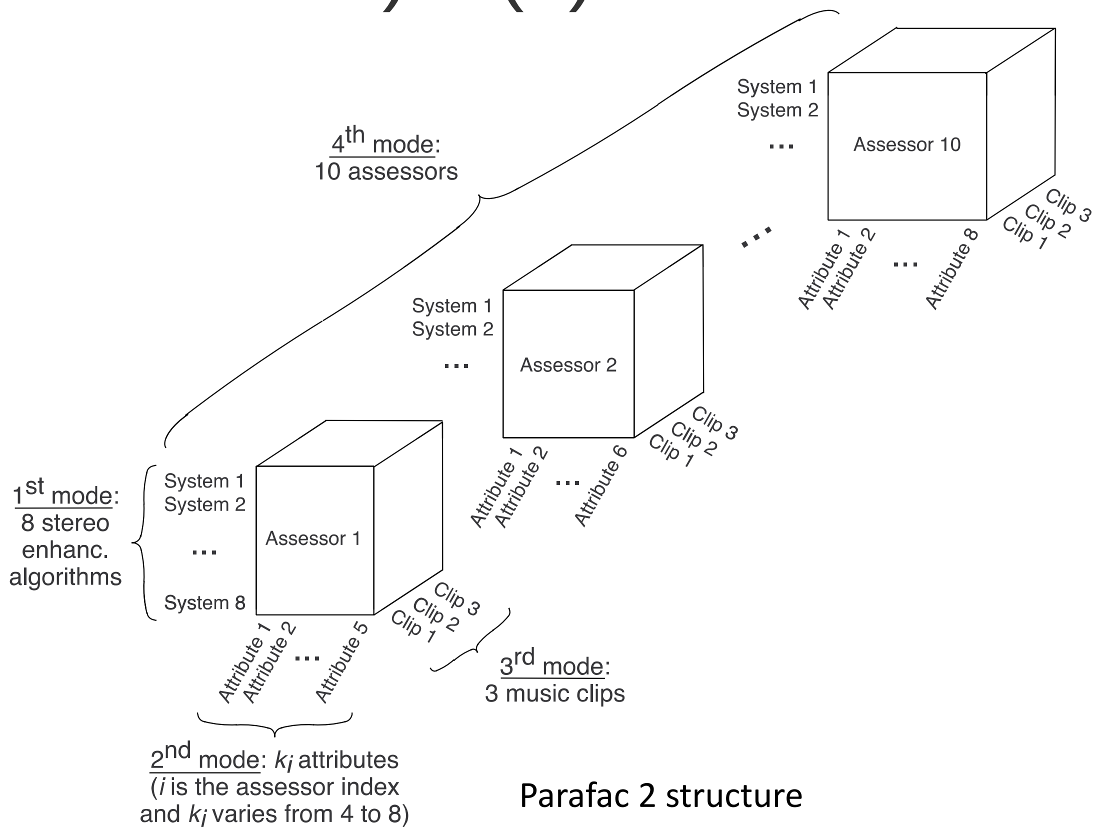
Lorho (2005-**2010**) – (2)



Lorho (2005-**2010**) – (3)

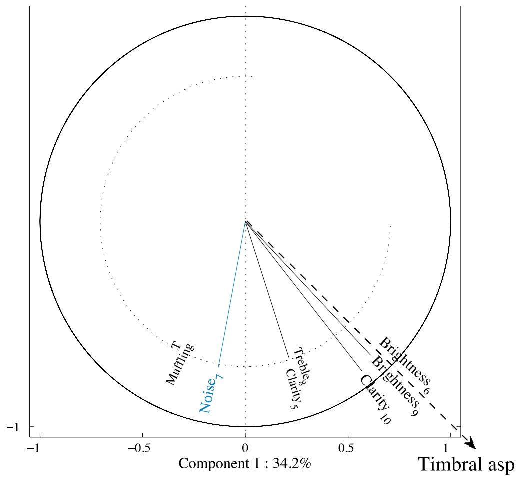


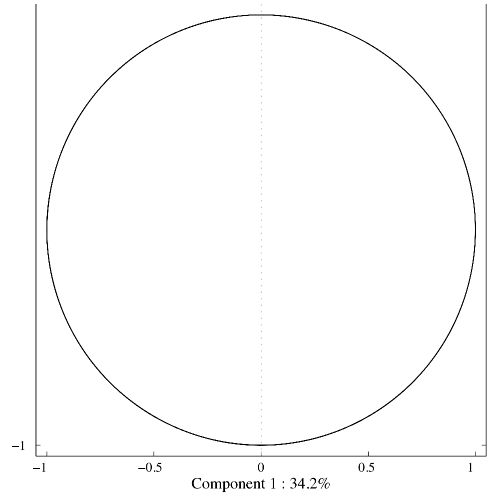
Lorho (2005-**2010**) – (4)

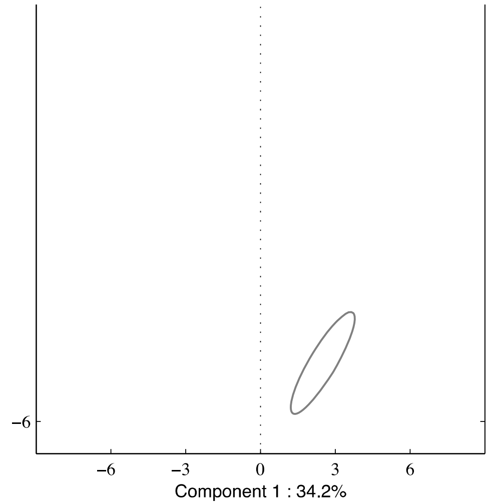


Parafac 2 structure

Lorho (2005-**2010**) – (4)



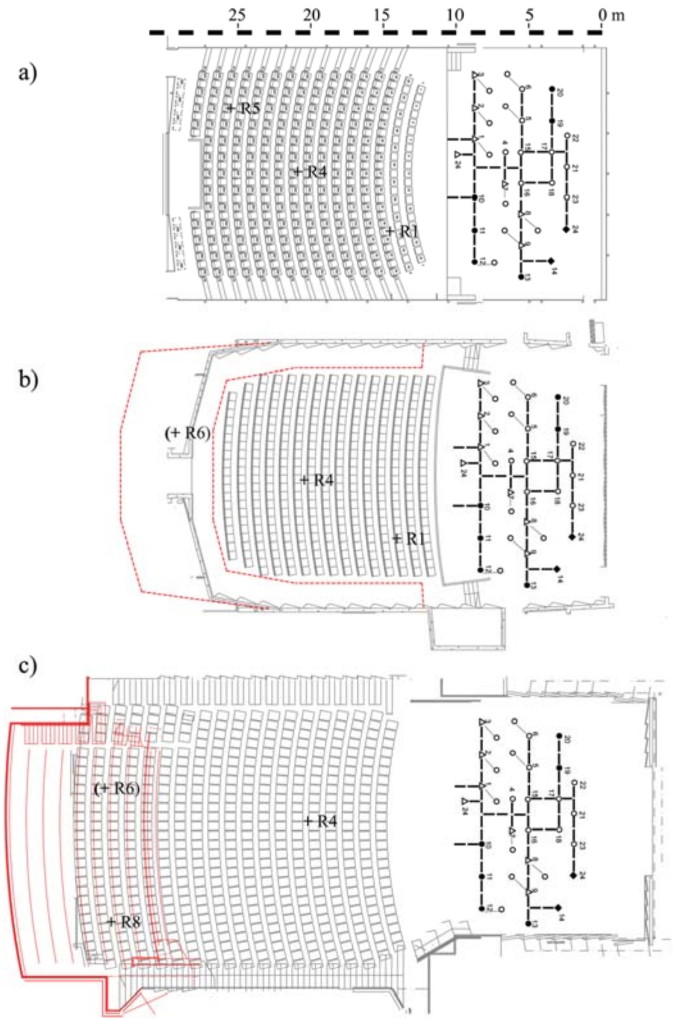




PCA analysis of the GPA average configuration

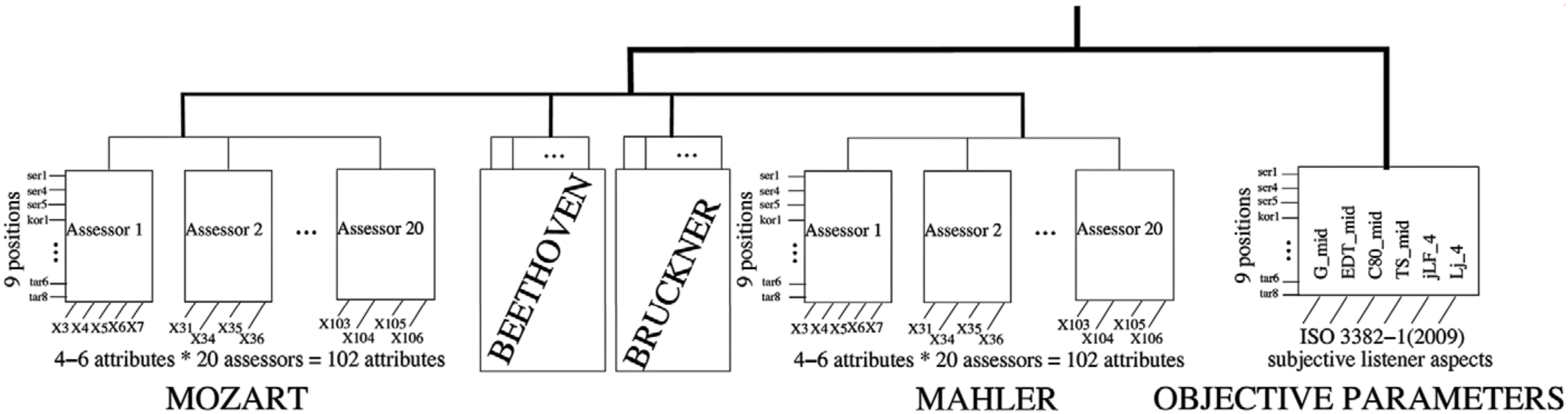
Lokki (2011)

- Latest and most definitive concert hall acoustic study
  - Fairly soon the conert hall will be understood
- Comparison of multiple halls at multiple positions
  - Using a virtual orchestra
  - IVP approach
  - (H)MFA analysis
    - Also for metric objective mapping

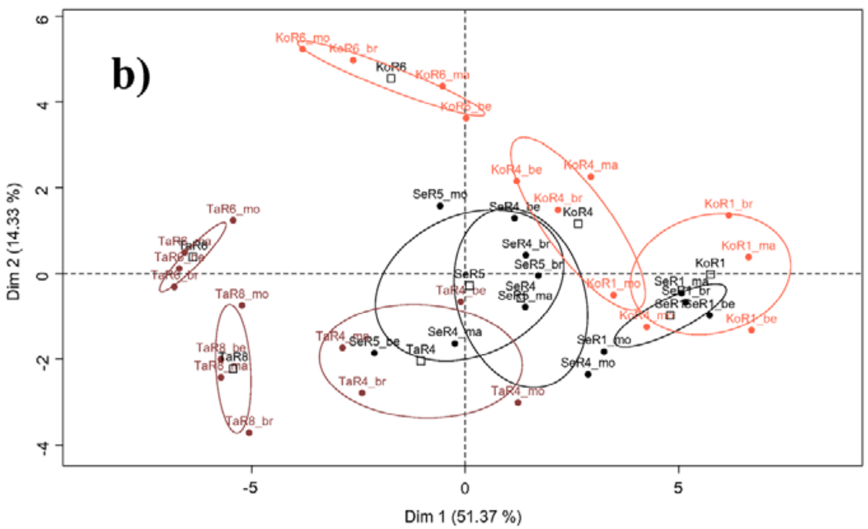


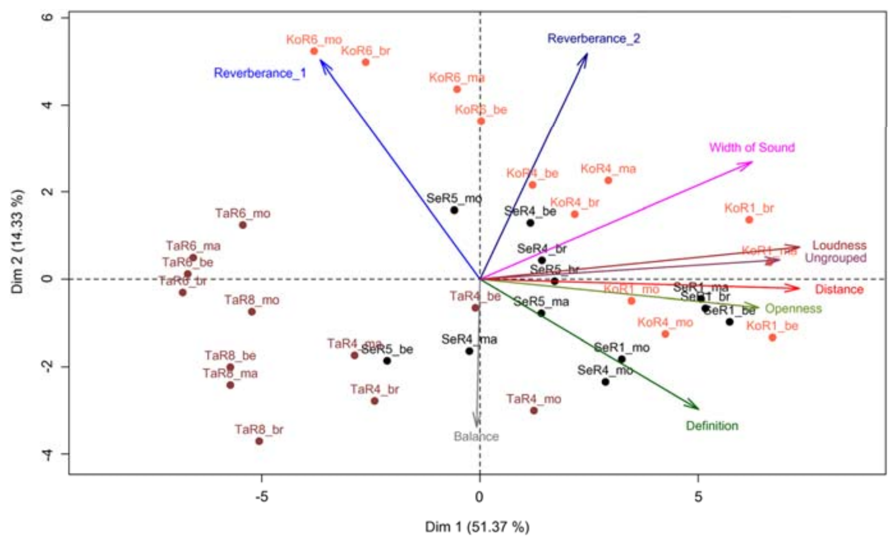
Lokki (2011) – (2)

- The basic data structure for HMFA analysis



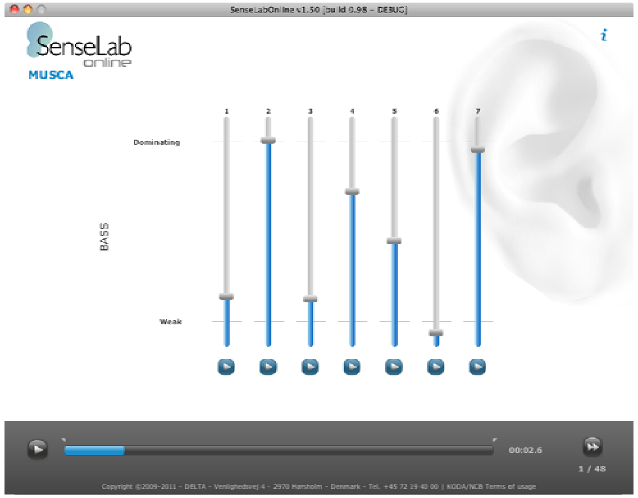
Lokki (2011) – (3)

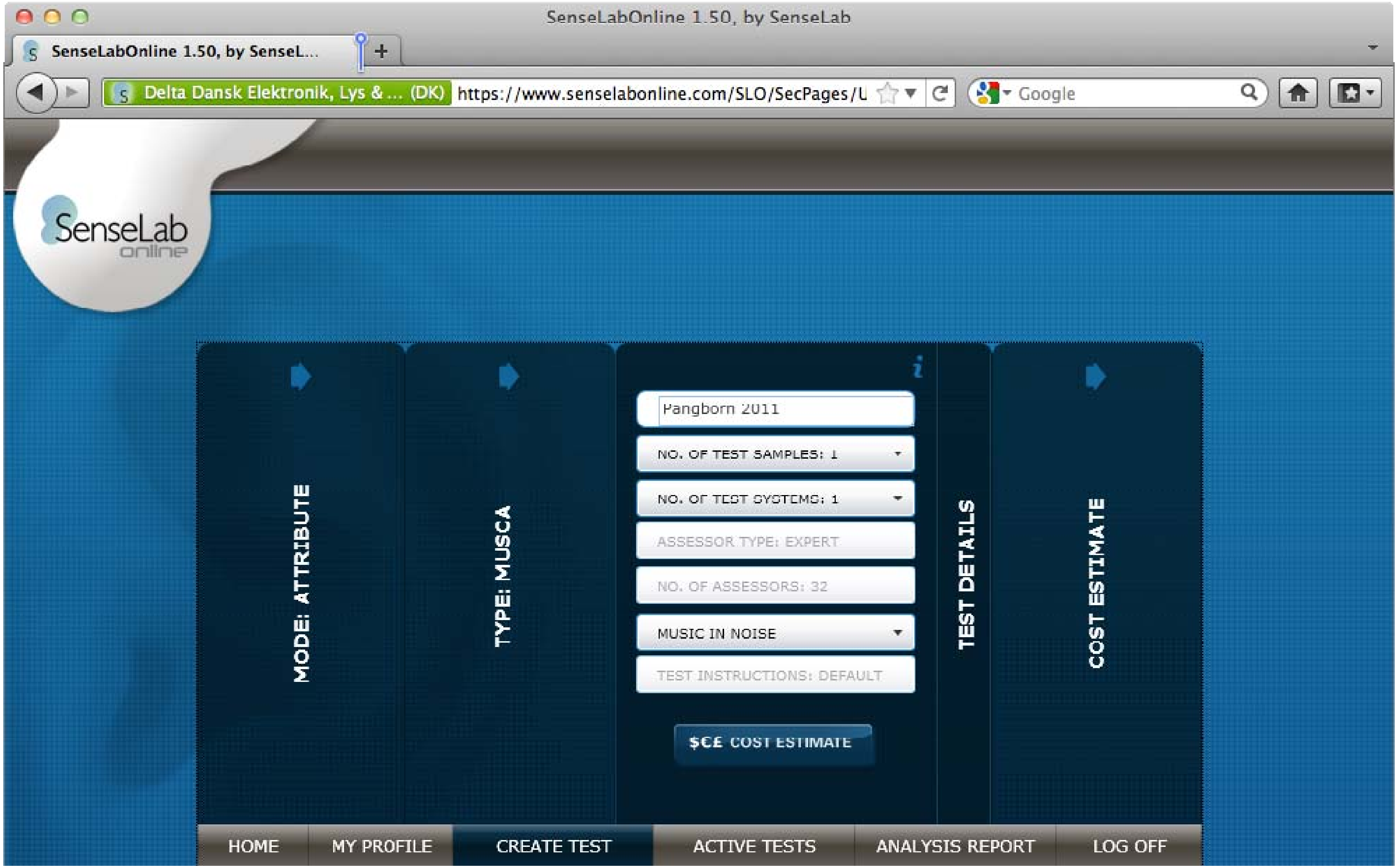

b)
Dim 1 (51.37 %)
Dim 2 (14.33 %)


Reverberance_1
Reverberance_2
Width of Sound
Loudness
Ungrouped
Distance
Openness
Definition
Balance
Dim 1 (51.37 %)
Dim 2 (14.33 %)

Ramsgaard (2011)

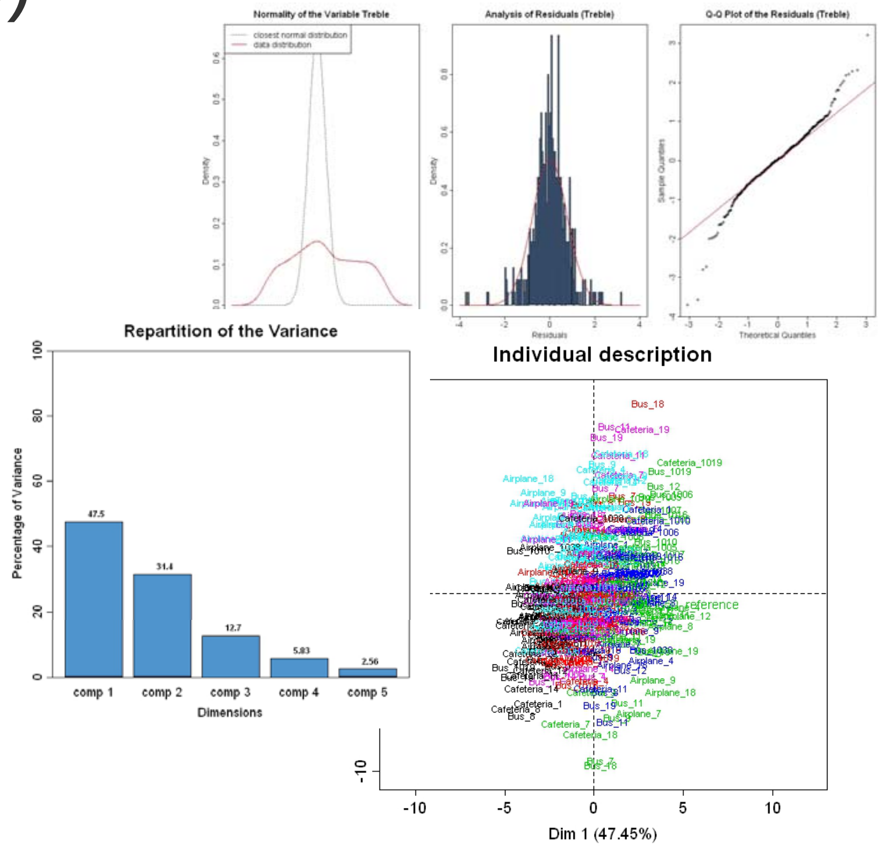
- SenseLabOnline
    - Easy creation of experiments (defining conditions, uploading files etc.)
    - Stimulus presentation and data gathering
    - Automatic statistical analysis
    - Automate everything where possible




SenseLabOnline 1.50, by SenseLab
SenseLabOnline 1.50, by SenseL...
Delta Dansk Elektronik, Lys & ... (DK)
https://www.senselabonline.com/SLO/SecPages/U
Google
SenseLab online
MODE: ATTRIBUTE
TYPE: MUSCA
Pangborn 2011
NO. OF TEST SAMPLES: 1
NO. OF TEST SYSTEMS: 1
ASSESSOR TYPE: EXPERT
NO. OF ASSESSORS: 32
MUSIC IN NOISE
TEST INSTRUCTIONS: DEFAULT
$€£ COST ESTIMATE
TEST DETAILS
COST ESTIMATE
HOME
MY PROFILE
CREATE TEST
ACTIVE TESTS
ANALYSIS REPORT
LOG OFF

Ramsgaard (2011) – (3)

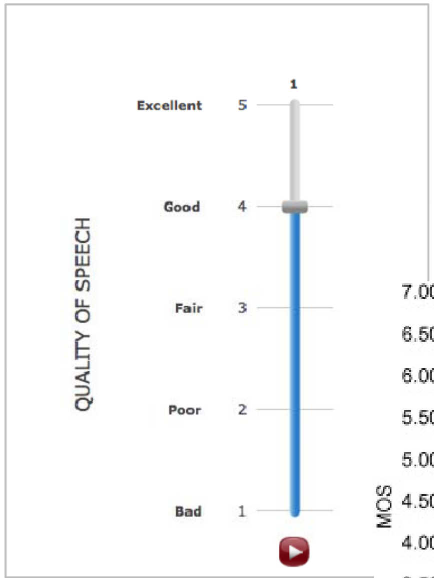
- Fully automated data analysis
  - And why not?

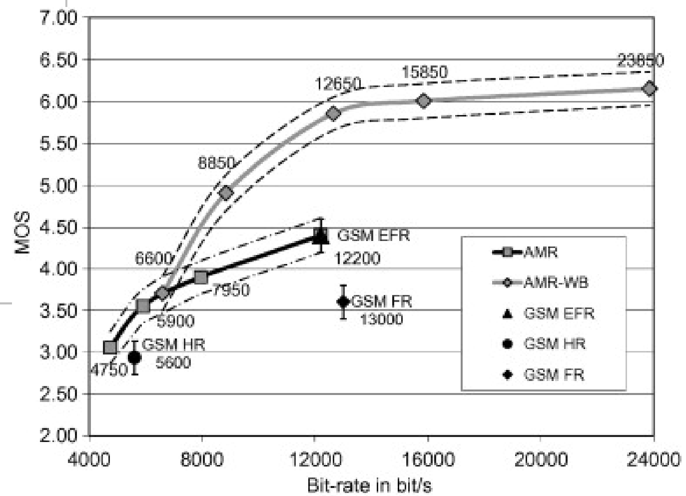


So what have we learnt?

What's been successful ?

- The Mean Opinion Score keeps a stronghold
  - Dominates the standards!
  - In the speech and audio codec world
  - Why is this so?
    - Fast and easy
    - Provides the needed level of information
    - The globally preferred codec
  - Only now are there moved towards attributes in speech
    - ITU-R P.835





Observations about CVP

- CVP works
  - But it requires **consensus**
    - This is not easy to get
    - **Check for consensus** needed
  - CVP takes a LOT of time
  - Also requires **expert assessors** (more time)
- All of this takes a significant amount of time, effort and expertise
  - Thus often its not done properly
    - Risk data quality
    - By default this IS the case ☹

Why is IVP gaining interest?

- Speed and ease !
  - Expert or naïve assessors can be used
    - Means to test assessor performance are basic, but functional
  - Latent attributes are derived from the IVP
  - Consensus is not needed
    - Its derived later (e.g. via clustering, etc.)
  - Its fast(er)
    - Therefore its cheaper

- More advanced stats are needed
  - This requires superior data analysts ☺
  - Still a PhD methodology ☹

- Its somewhat more consumer oriented (i.e. less expert)
  - The corporations value this

What do we know about analysis tools

- There is a time and a place for many tools
  - PLS – x
  - PCA
  - GPA
  - (H)MFA

  - Need to know your data
  - Need to understand the limits of the methods

- The power of the tools is excellent
  - But we would like more ways

- Once again this is NOT easy
  - It could be easier ;-)

What's not made headway (yet)?

- Advanced methods!
  - Preference mapping
  - Baysian methods?
  - PLS-Pathway
  - SEM

  - These methods are very potential
    - The key to many kingdoms (perhaps)
      - BUT….

- Why not?
  - Early days
  - The nature of the data sets
    - Large DOF experiments
      - Usually N-way
    - Smaller data sets (low number of consumers)
    - Massive experiments are
    - Some failed experiences
      - You need more than a PhD to apply some of these 😐

Making this easier

- Significant proportion of researchers cited are PhDs
  - Why are we making it so difficult to access?
    - Expertise is not about being clever
    - Its more about how to make this accessible
      - Hide the complexity?
    - e.g. **PanelCheck**
    - e.g. **SenseLabOnline**

Future direction and topics

- More multimodal
- More consumer oriented
    - Corporations don’t always *understand* expert assessors
- Even more ways
    - Speakers, environment, level, sample, assessor, attribute
- Would need to really benefit from the advanced methods
    - Preference mapping, PLS-pathway, Baysian methods, etc.
        - We want to key to the kingdoms

Conclusion

- Sensory science has had a big impact in the domain of sound
- Still much to be done
    - Mean Opinion Score is still dominating some fields
        - At the research level many advanced techniques are beneficially applied
            - But we need to make the methods more accessible (KISS)
    - Need methods that are better suited to our experimental designs
        - Large N-way, smaller sampling
    - Individual (IVP) approaches are gaining
        - Fast, more consumer oriented, avoid consensus challenges

Thank you!

- Questions?

s o u n d   p a s s i o n

References (1)

1. W. Sabine, “Reverberation: Introduction,” Am. Arch. (1900) [Reprinted in W. C. Sabine, Collected Papers on Acoustics (Harvard University Press, London, 1922), pp. 3–68.]
2. Beranek, L. L. (1962). Music, acoustics and architecture, John Wiley, New York.
3. M. Schroeder, G. Gottlob, and K. Siebrasse, “Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustics parameters” J. Acoust. Soc. Am. 56, 1195–1201 (1974).
4. H. Staffeldt, “Correlation Between Subjective and Objective Data for Quality Loudspeakers,” Audio Engineering Society Journal, Vol. 22, No. 6, 1974
5. H Wilkens, “Mehrdimensionale Beschreibung Subjektiver Beurteilungen der Akustik von Konzertsalen”, PhD Dissertation, TU Berlin, 1975
6. W. D. Voiers, “Diagnostic Acceptability Measure for Speech Communication,” Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP’77) , pp. 204–207, US-Washington (1977)
7. A Gabrielsson and H Sjogren, “Perceived Sound Quality of Sound-Reproducing Systems,” J. Acoust. Soc. Am. , 65(4): 1019–1033 (1979)
8. Gabrielsson, A. Statistical treatment of data for listening tests on sound reproduction systems. Tech. Rep. TA 92, Department of Technical Audiology, Karolinska Inst., Sweden, 1979.
9. F Toole, Listening Tests-Turning Opinion into Fact, J. Audio Eng. Soc, 30(6), 1982
10. Voiers, W. D., (1983), Evaluating Processed Speech using the Diagnostic Rhyme Test. Speech Technology pp. 30-39.
11. F Toole, Subjective Measurements of Loudspeaker Sound Quality and Listener Performance, JAES Volume 33 Issue 1/2 pp. 2-32; February 1985
12. F Toole, Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 1, JAES Volume 34 Issue 4 pp. 227-235; April 1986
13. F Toole, Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 2, JAES Volume 34 Issue 5 pp. 323-348; May 1986
14. Quackenbush, S. R., Barnwell, T. P., III and Clements, M. A., Objective Measures of Speech Quality, Prentice-Hall, New York, NY, 1988.
15. T. Letowski, “Sound-Quality Assessment: Concepts and Criteria”, Preprint 2825, Proc. 87th Conv. Audio Engr. Soc. (1989)
16. O. Warusfel, C. Lavandier, and J. Jullien, “Perception of coloration and spatial effects in room acoustics,” in Proceedings of the 13th International Congress on Acoustics (ICA’89), 173–176 (Belgrade, Yugoslavia) (1989).
17. C. Lavandier, “Validation perceptive d’un modèle objectif de caracteéisation de la qualité acoustique des salles (Perceptual validation of an objective model for characterizing the acoustical quality of rooms),” Ph.D. thesis, Université du Maine, Le Mans (1989).
18. Letowski, L. Timbre, tone colour, and sound quality: concepts and definitions Archives of Acoustics, 17 , pp.17-30, 1992.
19. Lundberg G, Ovegård A, Hagerman B, Gabrielsson A, Brändström U. Perceived sound quality in a hearing aid with vented and closed earmould equalized in frequency response. *Scan Audiol*. 1992; 21(2):87-92
20. S Bech, Perception of Timbre of Reproduced Sound in Small Rooms: Influence of Room and Loudspeaker Position, JAES Volume 42 Issue 12 pp. 999-1007; December 1994

References (2)

21. J. Blauert & U. Jekosch, “ Sound-Quality Evaluation – A Multi-Layered Problem”, ACUSTICA united with acta acustica 83, 747-753, (1996)
22. S. Bech, “Methods for subjective evaluation of spatial characteristics of sound”, 16th Conference of the Audio Engineering Society, 10-12 April 1999, Rovaniemi, Finland
23. V.-V. Mattila, “Perceptual analysis of speech quality in mobile communications”, PhD thesis, Tampere University of Technology, Tampere, Finland, October 2001.
24. N. Zacharov and K. Koivuniemi, “Perceptual audio profiling and mapping of spatial sound displays,” in Proceedings of the Int. Conf. on Auditory Display, p. 95-104, ICAD, 2001.
25. J. Blauert and U. Jekosch, “Concepts behind Sound Quality: Some Basic Considerations”, proceedings of Internoise, 2003
26. G. Lorho. Individual vocabulary profiling of spatial enhancement system for stereo headphone reproduction. In the 119th Audio Engineering Society Convention , New York, NY, USA, October 7-10 2005.
27. T. H. Pedersen, “The Semantic Space of Sounds, Lexicon of Sound-describing Words”, DELTA Tech. note, March 2005
28. J. Berg and F. Rumsey, “Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique,” J. Audio Eng. Soc., vol. 54, pp. 365–379 (2006 May).
29. Bech, S. and Zacharov, N. (2006). Perceptual Audio Evaluation – Theory, Method and Application, John Wiley & Sons, Chichester, England
30. S, Choisel and F. Wickelmaier, “Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference”, Journal of the Acoustical Society of America, 121, 2007
31. Ramsgaard, J., Le Ray, G., & Zacharov, N. (2009). Application of napping® to the evaluation of digital camera zoom lens function. Pangborn 2009 Symposium (Florence, Italy)
32. M. Wältermann, A. Raake, and S. Möller, “Quality Dimensions of Narrowband and Wideband Speech Transmission,” Acta Acustica united with Acustica , 96(6), pp. 1090–1103 (2010).
33. Lorho, G., Le Ray, G,. Zacharov, N. (2010). eGauge - A Measure of Assessor Expertise in Audio Quality Evaluations. AES 38th Conference (Piteå, Sweden).
34. Lorho, G, Perceived Quality Evaluation: An Application to Sound Reproduction over Headphones. PhD thesis, Helsinki University of Technology, Helsinki, June 2010.
35. Simonsen, C. S., Legarth, S. V. (2010). A Procedure for Sound Quality Evaluation of Hearing Aids. Hearing Review, Dec. 2010.
36. Zacharov, N., Ramsgaard, J., Le Ray, G., & Jørgensen, C. V. (2010). The multidimensional characterization of active noise cancellation headphone perception. Quality of Multimedia Experience (QOMEX) (Trondheim, Norway).
37. T. Lokki, J Pätynen and N Zacharov, “Concert hall acoustics assessment with sensory evaluation – tools and practices”, 8th International Conference on Auditorium Acoustics, Institute of Acoustics, Dublin, Ireland, 20-22 May 2011
38. Lokki, T., Pätynen, J., Kuusinen, A., Vertanen, H., and Tervo, S., "Concert hall acoustics assessment with individually elicited attributes," *Journal of the Acoustical Society of America*, vol. 130, no. 2, pp. 835-849, August 2011
39. Wankling, M, Fazenda, BM and Davies, WJ, 'The assessment of low-frequency room acoustic parameters using descriptive analysis' , Journal of the Audio Engineering Society, 60 (5), pp. 325-337, 2012
40. Blauert, J.; Jekosch, U., A Layer Model of Sound Quality, JAES Volume 60 Issue 1/2 pp. 4-12; January 2012