Audio Engineering Society

Convention Paper 7785

Presented at the 126th Convention
2009 May 7–10 Munich, Germany

The influence of sound processing on listeners' program choice in radio broadcasting

Hans-Joachim Maempel[1], Fabian Gawlik[1]

[1]Dep. of Audio Communication, Technische Universität Berlin
Sekr. EN-08, Einsteinufer 17c, 10587 Berlin, Germany

hans-joachim.maempel@tu-berlin.de, fgawlik@googlemail.com

ABSTRACT

Many opinions on broadcast sound processing are founded on tacit assumptions about certain effects on listeners. However, those have lacked support by internally and ecologically valid empirical data so far. Thus, under largely realistic conditions it has been experimentally investigated to what extent broadcast sound processing influences listeners' program choice. Technical features of stimuli, socio-demographic data of the test persons, and data of listening conditions have been additionally collected. In the main experiment, subjects were asked to choose one out of six audio stimuli varied in content and sound processing. The varied sound processing caused marginal and statistically not significant differences in frequencies of program choice. By contrast, a subsequent experiment enabling a direct comparison of different sound processings of the same audio content yielded distinct preferences for certain sound processings.

1. SUBJECT

Today, radio stations differ not only in broadcast content but also in sound appearance. Audio processing is performed in the playout centers of the broadcasting corporations in order to conform to legal requirements [1] and to encourage the attachment of a certain audience. Aesthetic concepts seem to cover the field between a relative pureness and a considerable colouration of the sound, as well as between the preservation of dynamics and loudness maximisation. In particular radio stations normally try to impress a unique sound, positively affecting the listeners. The development of the respective sound profiles and the fixing of their specific parameters imply the basic assumption that broadcast sound processing does affect listener reception appreciably at all. Till now, research has focused on a few specific characteristics of broadcast sound. For instance technical measures have been taken in order to determine loudness, and listening tests have been performed so as to assess overall sound quality [2]. However, there exist further relevant perceptual criteria potentially affected by sound, e.g. aspects of aesthetic impression, communication of brand values, intelligibility, recognition, and listening convenience in different situations and at varying degrees of receptiveness, respectively. Theoretically, in combination with the individual preferences for a specific content these aspects influence the listener behaviour, and in the end listeners make a short-term or a long-term choice of a radio program. Data on short-term program choice can be reliably collected under well-defined conditions, which has been taken advantage of at the investigation at hand. The hypothesis tested proposes that broadcast sound processing can influence listeners' spontaneous program choice.

2. METHOD

Subjects were asked to choose one of six playable audio programs on a graphical user interface. The listening test comprised two stages: In the first experiment the stimuli varied in content and sound processing as on a real radio. In the subsequent experiment only the sound processing was varied, enabling a direct comparison of sound, and subjects were once again asked to make a choice according to their preference. Since the first experiment provides a certain ecological validity, it can be generalised for the reality of radio listening to some extent, whereas the second experiment rather afforded a more sensitive measurement.

2.1. Stimuli

2.1.1. Selection of audio contents

The audio content used in the experiments consisted of five musical pieces and a report. In order to cover the spectrum of typical styles of music in radio broadcasting one piece of classical music and four pieces of popular music were chosen from playlists not more than six months old. The report was an unremarkable text spoken by a male professional voice. Since the report was copied from a backup recording, the voice had already passed through a studio voice processor. Table 1 lists the used audio content.

Category	Creator / Interpreter	Title	Year
Classical	Carl Nielsen	Aladdin Suite – oriental festive march	1991
Pop	Madonna	Ray of light	1998
Rock	Herbert Grönemeyer	Chaos	1994
Jazz	Till Brönner	Out of nowhere	2001
Acoustical	Norah Jones	What am I to you	2004
Speech	Inforadio RBB	Economic reportage	2007

Table 1: Audio content of the listening test

2.1.2. Audio processing

The typical path for broadcast signal processing resembles other sum processing chains for audio signals followed by a multiplex limiter and comprises processings for stereo, dynamics and level control (figure 1). First, a phase rotator [3] changes the relative phase of the frequency components in order to raise the polar symmetry of the time domain signal enabling a higher possible loudness [4]. Then a stereo enhancer homogenises the stereo correlation by dynamically adjusting the ratio of the mid and side signal. The following automatic gain control is a long-term leveller reducing high and raising low amplitudes. Thereby, a dynamic expander prevents the increase of noise in silent passages. At the next stage the signal passes a frequency dependent and frequency selective dynamic compression enabled by a multiband structure with typically 2 or 5 bands. This process homogenises the distribution of energy in the spectrum of the signal and allows for a higher loudness [5]. In order to conform to the limitation of frequency modulation for radio broadcast given by [6], finally a multiplex limiter controls the level of the audio signal depending on the anticipatory power of the MPX signal [7]. The several components are often integrated in one broadcast processor. Its output signal is fed to the multiplex encoder providing the modulated signal with a band frequency range of 10 Hz to 57 kHz as a mid-side signal (MPX-signal).

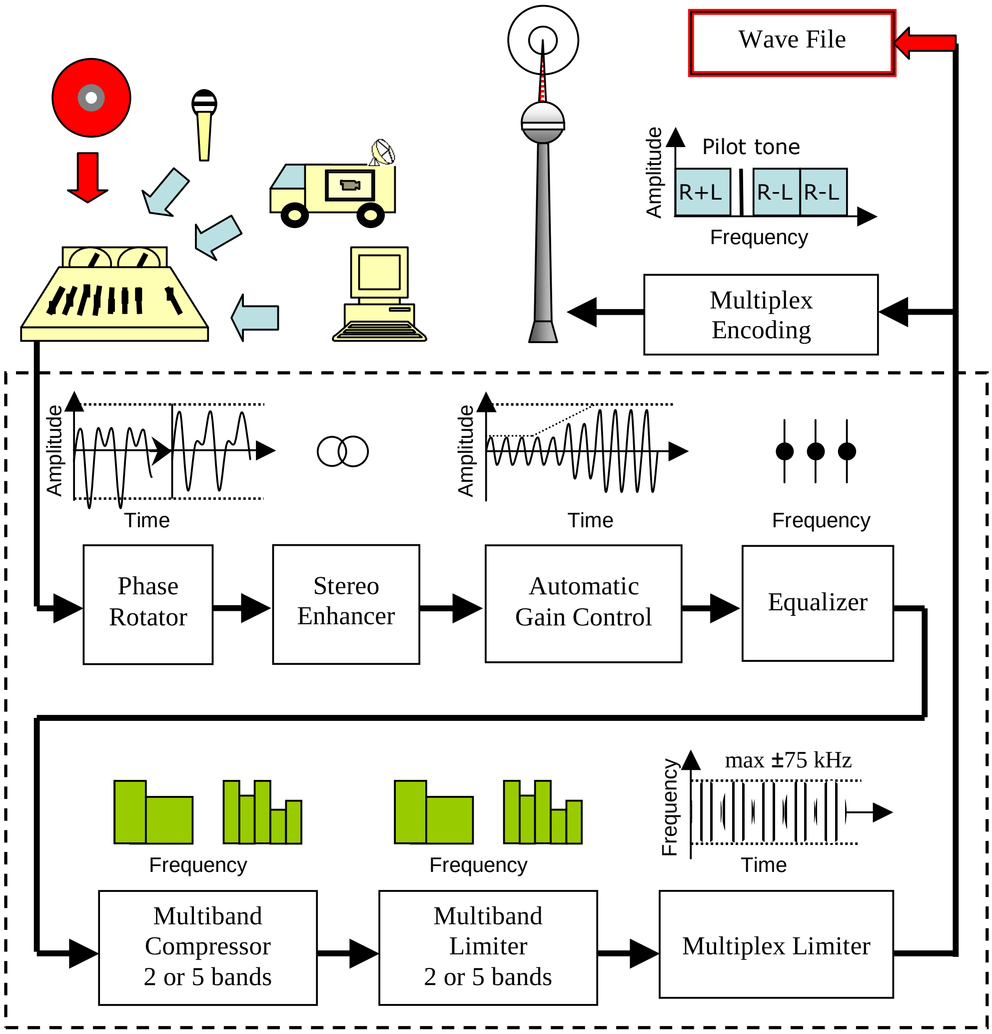

Figure 1 : Broadcast processing chain

The overall time- and frequency-dependent regularisation can cause audible effects as 'pumping', 'time smearing' and the 'commercial sound' as well as a high loudness and intelligibility. Since broadcasting corporations are generally reluctant to provide information on their sound processing, it was performed by a leading manufacturer of integrated broadcast processors. Therefore, presets of different German radio stations were applied. By this way, as such, the values of the technical parameters are not transparent, but the audio contents could be processed as they would have been by German radio stations during broadcast. Five differently processed versions were produced from each content. Additionally, unprocessed versions were prepared by normalising each audio content to -9 dBFS according to the ARD norm level [8]. With respect to the predominant terrestrial analog transmission, the five processing chains also contain the MPX limiter. Although a MPX encoding, preemphasis and transmission path can further change the audio quality [9], they were not emulated, so as to allow for a certain generalisation over DAB and other types of digital transmission. Table 2 provides an overview of the generated stimuli.

Code	Content A	Content B	Content C	Content D	Content E	Content F
Processing 1	A1	B1	C1	D1	E1	F1
Processing 2	A2	B2	C2	D2	E2	F2
Processing 3	A3	B3	C3	D3	E3	F3
Processing 4	A4	B4	C4	D4	E4	F4
Processing 5	A5	B5	C5	D5	E5	F5
Unprocessed	A6	B6	C6	D6	E6	F6

Table 2 : Stimuli of the listening test

2.1.3. Characteristics of the stimuli

In order to allow for a description of the effect of the processings, besides music and speech some test signals (i.e. impulse, white noise and sine waves) were processed. Figure 2 shows the peak levels of the processed and unprocessed test signals. Through a comparison of the test signal levels, curves of compressor characteristics can be identified. A comparison of the impulse and noise signals of the same level shows the effect of different time constants. Finally, the difference in amplitudes of the two sine signals differing in frequency indicate a multiband compression and/or equalising [10].
Figure 3 shows the amplitude spectra of white noise (-9 dBFS) before and after the processing as smoothed RMS levels. The observable spectral changes are put down to equalisation and multiband compression, thus differing for other input levels. The largest differences can be observed in the bass region and between 2 and 7 kHz. Processing 3 stands out with the highest level in the bass region. The steep low-pass filter at 15 kHz keeps the upper frequency region free with regard to the 19 kHz pilot signal.
Figure 4 shows the crest factors, loudness levels according to [11] and peak levels of the processed and the unprocessed versions for all audio contents and white noise. The lowering of the crest factors points to a dynamic compression. As expected, processings producing lower crest factors stand out by a higher loudness. Overall, the different values of the unprocessed signals were homogenised by the processings. However, complex sound qualities can only be poorly predicted by technical measures.

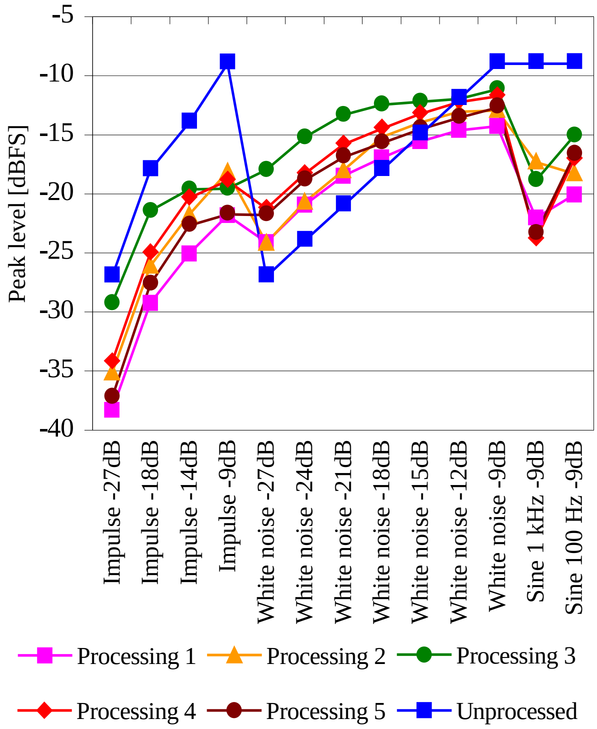

Figure 2 : Peak levels of the processed and unprocessed test signals

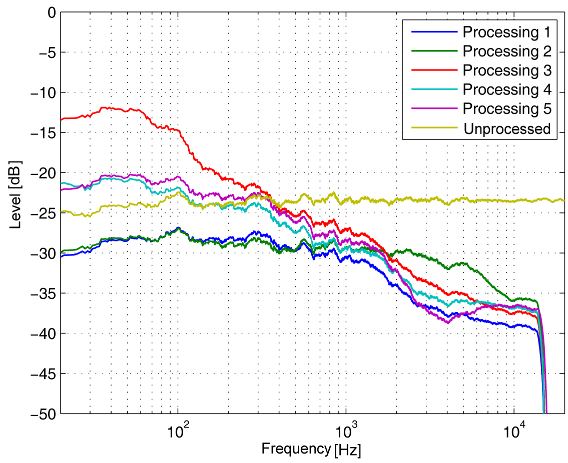

Figure 3 : Amplitude spectra of white noise before and after processing

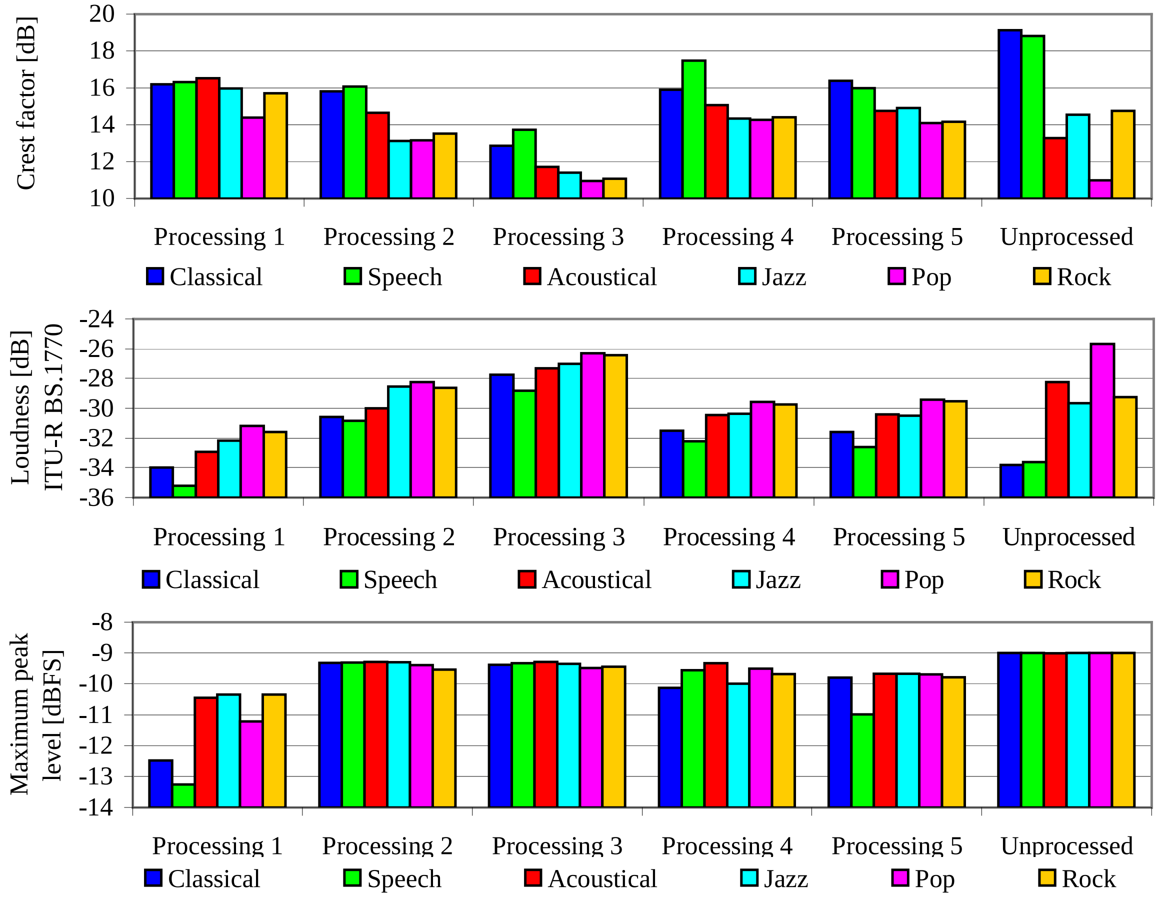

Figure 4 : Crest factors, loudness levels [11] and peak levels of the processed and the unprocessed versions

2.2. Test design

The listening test consisted of two experiments, which each subject ran through subsequently. Experiment 1 is in line with a largely realistic variation of conditions, allowing for results that bear ecological validity. However, experiment 2 achieves a high internal validity and reliability, thus following criteria applied in listening tests predominantly.

2.2.1. 1st Experiment

In order to provide options and thereby be in accordance with a natural program panel on a radio, sound and content must be confounded (systematically co-varied). Thus, an independent and complete variation of conditions (full factorial design) is inapplicable. A confounding of sound and content is given by coinstantaneous offering of six programs: In each decision task (trial), one set of stimuli arranged diagonally in table 2 (same-colour) was presented to the subject. The order was randomized so as to avoid position effects. On the one hand, due to the confounding the audio content has to be regarded as an interfering variable degrading the reliability of the measurement, on the other hand it is a factor providing a certain ecological validity to the experiment. In order to be able to measure the effect of the factor of sound processing independently, the confounding itself is required to be variated and thereby balanced. The different confoundings were realised by shifting the order of processings relative to the order of audio contents, and were presented subsequently in time: Each subject ran through six trials (respectively through all of the colours in table 2). The order of presentation was randomized.

2.2.2. 2nd Experiment

With the objective of measuring accuracy, the confounding was released in the second experiment. In this way, the subject was able to directly compare different sound processings of the same audio content. Thereby, the audio content chosen by the subject in experiment 1 was used. So in each trial the respective set of stimuli arranged vertically in table 2 was presented to a subject. The positions of the stimuli on the control panel were randomised in each trial.

2.3. Measure

As a measure of preference the frequency with which subjects chose a given program (experiment 1) and a given sound profile (experiment 2) was counted. The measuring instrument consisted of the sample of subjects and a graphical user interface for the presentation of the stimuli and the registration of the program choice. Additionally the sound pressure level was measured (A weighted) by using the Brüel&Kjaer sound level meter 2205 with the adjustment on slow ($t_{integration}$=1s).

2.3.1. Experimental setup

The tests were carried out individually. Figure 5 shows the technical setup. For monitoring Genelec 1029A were used (active monitors). With regard to the ecological validity the setup was designed to be portable and was applied in typical radio listening situations [12] – frequently in surroundings familiar to the subjects ones. The interface was as simple as possible in order to allow an easy and intuitive handling even by technically unexperienced subjects. The volume could be easily adjusted by a hardware-knob. Since the studio monitors show a relatively linear frequency response ranging from 55 Hz to 18 kHz, the sound quality exceeds that of typical radios.

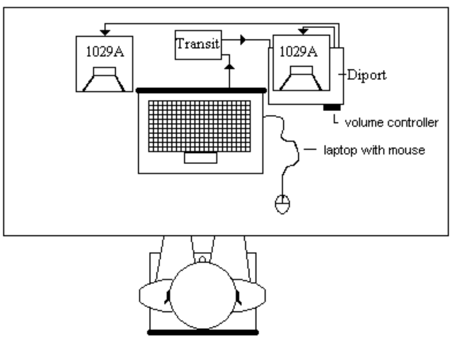

Figure 5 : Technical setup of the listening test

The optical representation of a simple radio receiver panel, the interaction on the subject by staying on a selected stimuli, and the registration of the program choice were carried out by a graphical user interface on a notebook (Figure 6) in connection with a controlling software driven by Matlab.

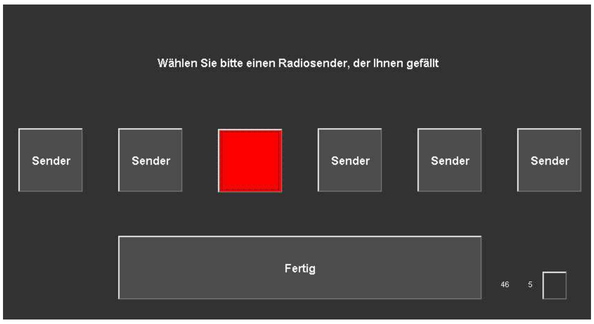

Figure 6 : Graphical user interface

2.3.2. Sample

The sample consisted of 60 non-expert subjects, in view of high population validity. According to media analysis data [13] the proportion of men to women was 48% to 52%. This pool, however, represents a relatively better educated and a younger part of the population. The age distribution ranges from 16 to 61 years (median 29 years) with a main field of 82% between 20 and 39 years. 47% have graduated high school or college and 45% have completed secondary school. The total sample size is given by the quantity of six trials (if valid) per subject and is N=358 (first experiment) and N=357 (second experiment) respectively. It allows for the statistical coverage of a medium effect size [14].

2.4. Statistical evaluation

Due to nominal scale measurement, the evaluation of the frequencies for a statistical inference is based on χ^2-methods. On the one hand the frequencies were tested for independence of content and processing in a two dimensional way, on the other hand the one-dimensional goodness-of-fit test of equal distribution was used for the separate contents as well as for their totals. The expected frequencies were defined as the average of the observed frequencies. If the considered part of the sample is nearly 60 or less, the Yates' correction was used. Based on [14] the test for independence is computed by:

$$\chi^2 = \sum_{j=1}^{m}\sum_{k=1}^{r}\frac{\left(f_{jk} - f_{jk}^*\right)^2}{f_{jk}^*} \qquad (1)$$

with the expected frequencies

$$f_{jk}^* = \frac{\sum_{a=1}^{m} f_{ak} \cdot \sum_{b=1}^{r} f_{jb}}{\sum_{a=1}^{m}\sum_{b=1}^{r} f_{ab}} \qquad (2).$$

An mutual influence of sound processing and content at the listeners program choice can be assumed by a statistically significant dependence, otherwise it would be a stochastic association. An influence of sound processing or content on the listeners program choice is shown by an inequal distribution. The level of significance was set at α=0.05.

3. RESULTS

3.1. Main results

The collected data were plotted in figure 7 and 8 and tested for independence. The analysis of the first experiment shows no significant dependence of content and processing (table 3). Considering only the musical content, the χ^2 value drops from 32.35 (p=0.15) to 16.62 (p=0.68), thereby indicating the special influence of the speech, which has exceptional characteristics. However, in the second experiment the analysis corroborates a significant dependence of content and processing.
The marginal totals for the variation of content are shown in figure 9 and for the variation of the sound processing in figure 10 (first experiment) and figure 11 (second experiment).
The goodness-of-fit test for equal distribution was performed on each content group as well as on the marginal totals (table 4). The tests show significantly inequal distributions of preferences due to the variation of content with χ^2 = 53.32 (p=0.00) and to the variation of sound processing with fixed content (table 4, second experiment). An obvious preference of the loud and warm sound processing 3 can be found in four of the six contents. By contrast, the test indicates an equal distribution of preferences due to the systematic co-variation of content and sound processing (table 4, first experiment). Exceptions are given by the speech stimuli, that show an opposite statistical result and acoustical music with a statistically assumable equality of distribution in both experiments.

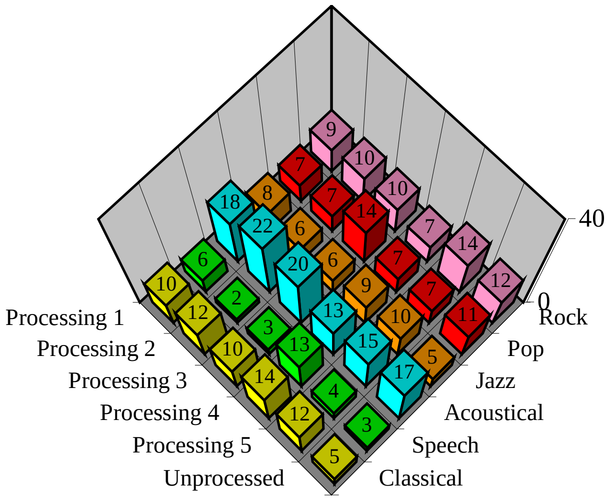

Figure 7: Observed frequencies in the 1st experiment

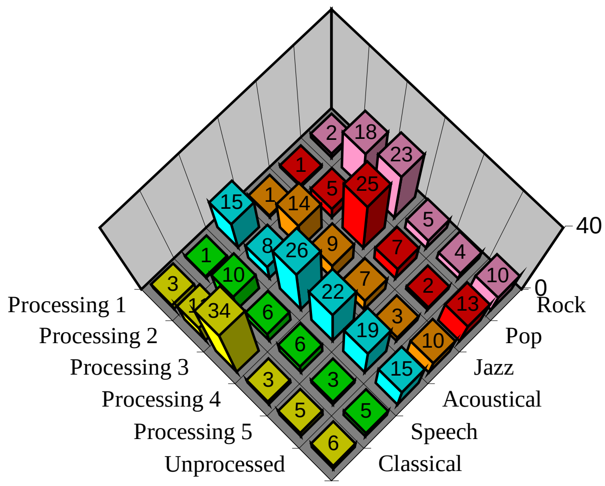

Figure 8: Observed frequencies in the 2nd experiment

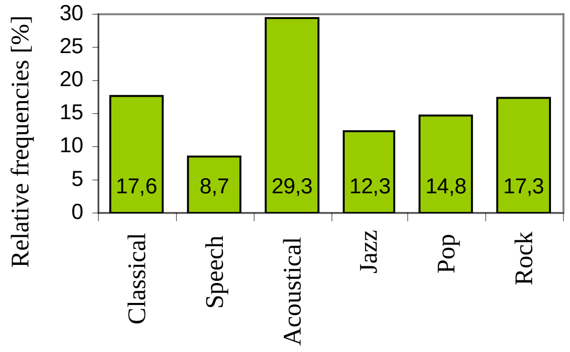

Figure 9 : Marginal totals of content (χ^2 = 53.32, p=0.00)

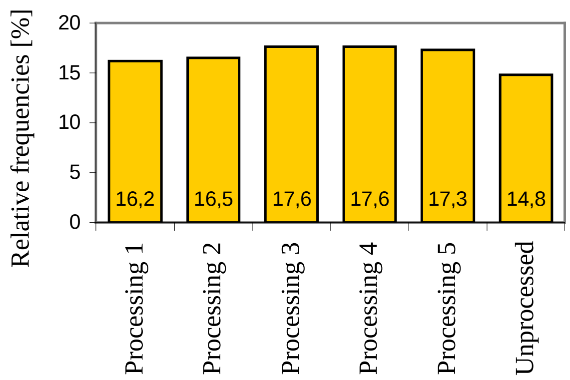

Figure 10: Marginal totals of processings (1st experiment)

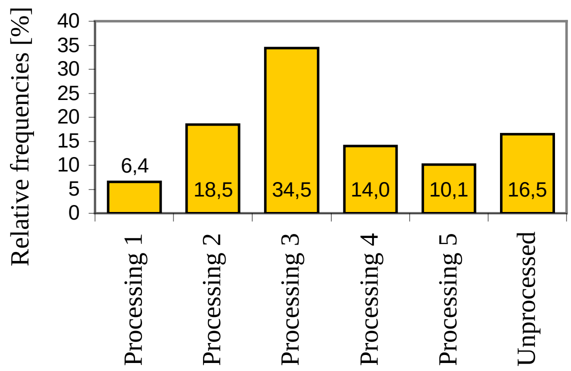

Figure 11: Marginal totals of processings (2nd experiment)

		1st experiment		**2nd experiment**	
	df	χ^2	p	χ^2	p
Total	25	32.35	0.15	77.53 **	0.00
Only music	20	16.62	0.68	70.91 **	0.00

** highly significant (α=0.01)

Table 3: Test of stochastic independence of content & processing

	1st experiment			**2nd experiment**		
df =5	χ^2	χ^2 (YK)	p	χ^2	χ^2 (YK)	p
Classical	-	3.43	0.634	-	64.66 **	0.000
Speech	-	12.97 *	0.024	-	6.38	0.233
Acoustical	-	2.29	0.808	-	9.17	0.084
Jazz	-	1.48	0.915	-	12.66 *	0.027
Pop	-	3.58	0.611	-	41.39 **	0.000
Rock	-	1.95	0.856	-	30.79 **	0.000
Total	1.26	-	0.939	101.67 **	-	0.000

YC: Yates' correction; * significant (α=0.05), ** highly significant (α=0.01)

Table 4: Test for equal distribution of the processings

3.2. Additional results

Additional results of this exploration are the average listening sound level with 58 dB(A), the average remaining time on the programs of 4.6 s (first experiment) and 5,3 s (second experiment), and an average switching frequency of 6.5 (first experiment) and 8.6 (second experiment).
Furthermore, the relative frequencies of the preferentially used types of radio players named by the subjects (figure 12) show that the kitchen radio is most commonly used (35%), followed by the car radio (22%)

and the small hifi radio (20%). In conclusion it can be assumed that normally listening conditions and acoustical environment are of low quality and attentive listening is not in focus. For further analyses see [15].

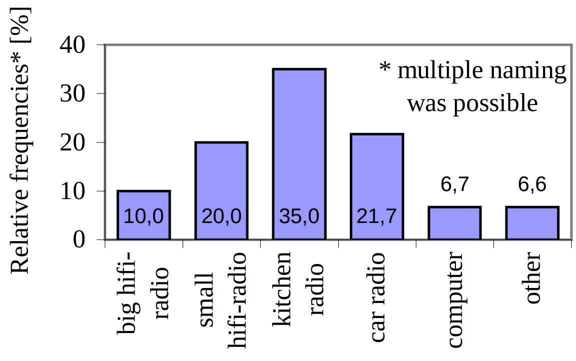

Figure 12 : Mainly used radio players

4. DISCUSSION

By comparing both experiments, the presented exploration clearly emphasizes the significance and the difficulty of achieving ecological and external validity. In the first experiment the content and sound processing are confounded factors and largely comply with a realistic situation. Except in the case of speech, the variation of sound processing caused marginal and statistically insignificant differences in frequencies of program choice, which was rather strongly determined by the offered content. In addition no significant stochastic dependence could be found between content and sound processing. The descriptives do not show any appreciable preferences or tendencies caused by the processing, except a lower total frequency of the unprocessed versions. Therefore, the hypothesis could not be corroborated with respect to a medium effect size. On the other hand, the second experiment, allowing for a direct sound comparison, which cannot be found in broadcast reality, yielded distinct preferences for specific sound processings, corroborated by statistical significance, except for speech and acoustical content. The inseparability of musical structure and sound, the principle of inherence [16], empirically appears in the second experiment, which bought its higher sensitivity by a loss of external validity. The observed preferences for processing number 3 affirms the assumption of a positive effect from high loudness but also from a high bass found in four of six cases of content (figure 8). So in fact the particular spectrum, if selectable by the listener, can also cause a high influence on the preferences as well.
One explanation for the divergence of the results is that the two experiments drew subjects' attention to different entities and changed the demand of a respective tradeoff between them. However, this methodical side effect shows that in general the criterion of external validity should be more often considered in the design of listening tests. Furthermore a reactivity of the listeners by the unavoidable transparent variation of conditions and question is probable. A replication and refinement is advisable if small preferences with regard to a high amount of listeners should be relevant. In order to cover a small effect size the hypothesis can be inspected as in the first experiment, but with a much higher sample and with a larger sample of test programs. A further analysis of the action of choosing a program in the sense of formulating theory, a mutual dependence of the variables of loudness, loudness adjustments, sound quality and time of staying on a program should be considered.

5. CONCLUSION

By all means under realistic conditions the specific sound appearance hardly seems to play a role in listeners' spontaneous program choice, as long as the audio content passes through a broadcast processor at all. Content is the primary factor. The ability to compare sound profiles directly, thereby accepting an unrealistic variation of stimuli, induces to a widespread opinion, that loud signals are preferred. However, due to the specific processings used in the experiments, this is connected with a high bass amplitude.
In any case it should be pondered that such effects of loudness generally don't last very long. They disappear with the first spontaneous loudness correction by the listener, in contrast to the persistent loss of sound quality by the compression and the following risk of medium-term annoyance [17].

6. ACKNOWLEDGEMENTS

The authors are deeply grateful to Peter van Beusekom from Orban Europe for processing the audio content.

7. REFERENCES

[1] Bundesnetzagentur, Referat 222, Vorläufige Verwaltungsvorschrift für Frequenzzuteilungen für den Rundfunkdienst VVRdfk, (2007.03)

[2] Spikofski, Gerhard; Klar, Siegfried, Digitalradio and CD-Qualität – Realität oder Traum?, 22. Tonmeistertagung, Hannover, (2002)

[3] Wolters, Martin, State of the Art Speech Processing for Broadcasting, Omnia.com, White paper, (2009)

[4] Werrbach, Donn, The Aphex Model 230’s Phase Rotator, Aphex.com, White paper, (2005)

[5] Foti, Frank; Orban, Robert, What Happens to My Recording When it's Played on the Radio?, AES Convention 111, Paper Nr.5469, (2001)

[6] International Telecommunication Union, Planning standards for terrestrial FM sound broadcasting at VHF, ITU-R BS, 412, (1998)

[7] International Telecommunication Union, Limiters for high-quality sound-programme signals, Rec. ITU-R BS.642-1, (1990)

[8] Institut für Rundfunktechnik, Empfehlung 15 der ARD-Hörfunkbetriebsleiterkonferenz: Headroom bei digitalen Tonsignalen, IRT, München, (1994)

[9] Foti, Frank, Audio Processing For Digital Broadcast Mediums, 116th AES Convention 2004, Convention Paper 6040, (2004)

[10] Orban, Robert; Foti, Frank, What Happens to My Recording When it’s Played on the Radio?, 111th AES Convention, Convention Paper 5469, (2001)

[11] International Telecommunication Union, Algorithms to measure audio programme loudness and true-peak audio level. Rec., ITU-R BS.1770, (2006)

[12] Frisch, Anne-Linda, Senderbindung im Hörfunk - Eine Sekundärauswertung der Media-Analyse 2000 zur Hörfunknutzung in Deutschland, Exam, Univ., Jena (2000)

[13] Arbeitsgemeinschaft Media-Analyse, ma 2007 Radio II, „“http://reichweiten.bik-gmbh.de/ma/online/2007radioU/rw“, (2007.12)

[14] Bortz, Jürgen, Statistik für Human- und Sozialwissenschaftler, Heidelberg: Springer Medizin Verlag, (2005)

[15] Gawlik, Fabian, Der Einfluss des Klangprocessings auf die Senderwahl im Hörfunk, Exam, Techn. Univ., Berlin, (2008)

[16] Maempel, Hans-Joachim, Klanggestaltung und Popmusik. Eine experimentelle Untersuchung, Labor Synchron; 1, Heidelberg: Synchron, (2001)

[17] Wagner, K., Zur Lautheit von Rundfunkprogrammen, In: Radio, Fernsehen, Elektronik: die Zeitschrift der Unterhaltungselektronik 46 (3), S.42-46, (1997)