Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Perceptual space of multiple complex tones

Stephan Töpken, Jesko Verhey and Reinhard Weber

Oldenburg University, Institute of Physics
Carl-von-Ossietzky Str. 9-11, 26129 Oldenburg, Germany

PACS: 43.66.Cb, 43.66.Jh, 43.66.Lj

ABSTRACT

Large fans often emit complex tones with a fundamental tone and a large number of harmonics. Two fans at different rotational speeds may generate complex tones with different fundamental frequencies and 'combination tones'. The perception of those large tone complexes has many aspects and varies e.g. according to the frequency ratio of the fundamentals, where only small ratio changes may lead to considerable changes in the pleasantness of the sounds. The aim of the study is to identify the perceptual space for these sets of tones and to investigate the impact of the fundamental's frequency ratio on the different perceptual dimensions. A semantic differential is used to determine denotative and connotative properties of fifteen different mixed complex tones. A factor analysis provides results that have specific relations to the three factors 'pleasant', 'powerful' and 'metallic'.

INTRODUCTION

Complex tones can occur as a part of machinery noise for all machines with rotating parts. In special assemblies of two fans, the resulting sound consists of two complex tones and combination tones. In a first approach to the perceptual aspects artificially generated sounds consisting of multiple superimposed sinusoids were investigated in two previous studies. It turned out that the perceptual characteristics of such sounds depend in large part on the frequencies of the fundamentals and on their frequency ratio. Using a semantic differential, it was found that a fundamental ratio equal to a ratio of small integers evokes a steady, hard and unpleasant sensation (Töpken et al. 2010b). Experiments with a paired comparison paradigm and a higher frequency resolution for the second fundamental frequency indicate that tone complexes with fundamental frequency ratios of small integers are judged louder and more unpleasant than those with fundamental frequency ratios of larger integers (Töpken et al. 2010a).

This study explores the perceptual space of the same sounds as in the paired comparison experiment of the second study (Töpken et al. 2010a) in order to identify the different perceptual dimensions in these sounds.

METHOD

Semantic Differential

The sounds are evaluated using a semantic differential of 16 adjective pairs, on individual 11-point categorial scales for each item pair (see table 1). Seven adjective scales are unipolar, 9 are bipolar. The adjectives have been composed on the basis of a pre-test (Töpken et al. 2010b) and to allow consistancy checks, the evaluations of two scales (pleasant - unpleasant and soft - hard) are repeated at the end of each session. According to the mother tongue of all participants the German originals of English translations of the adjectives in table 1 are used in the experiments. Regarding the order of evaluations, all sounds are assessed using one adjective pair. Next all sounds are assessed using the next adjective pair and so on. For this, the scales are arranged in 18 booklets (16 adjective scales plus the two repeated scales), with single scales on 15 pages, one page for each sound. The direction of the scales and the presentational order of the sounds are randomized differently for each scale, but uniformly for all participants.

Table 1: English translations of the adjective scales in the chronological order of the evaluation

No.	adjective scale
1	pleasant - unpleasant
2	smooth - rough
3	noisy - tonal
4	dull - sharp
5	not fluctuating - fluctuating
6	soft - hard
7	harmonic - discordant
8	not loud - loud
9	low - high
10	clean - dirty
11	not functional - functional
12	not intrusive - intrusive
13	vague - clear
14	not dominating - dominating
15	not hammering - hammering
16	not threatening - threatening
17	pleasant - unpleasant (rep.)
18	soft - hard (rep.)

Stimuli

The 15 stimuli are synthesized superpositions of two complex tones with fundamental frequencies f_{10}, f_{01} ($f_{10} < f_{01}$) and 29 higher harmonics as well as additional combination tones f_{ij}. In detail, the frequency components of the test stimuli are as follows:

$$
\begin{array}{ll}
f_{i0} = i \cdot f_{10} & i = 1,2,3,...,30 \\
f_{0j} = j \cdot f_{01} & j = 1,2,3,...,30 \\
f_{ij} = i \cdot f_{10} + j \cdot f_{01} & i,j = 1,2,3,...,20
\end{array}
$$

Thus each test sound consists of 460 partial tones. All partials have a random starting phase and decrease by 6 dB/octave in level. The combination tones are attenuated by 10 dB compared to the two complex tones. All stimuli have a duration of 5 seconds. The lower fundamental f_{10} is fixed (100 Hz). The higher fundamental f_{01} (and therewith the ratio of the fundamentals) is varied from 128.57 Hz to 137.5 Hz. Table 2 shows the details of the stimuli parameter. The ratios include 5 ratios of small integers as well as 10 different, detuned ratios, resulting in ratios of large integers. The names of the sound are made up of the fundamental frequencies and the corresponding small integers in case of an exact ratio. All stimuli are generated and stored in a computer with a sampling rate of 22050 Hz and a resolution of 16 bits. Figure 1 and 2 show the powerspectra and the time series of two exemplary sounds.

Table 2: Parameters of the 15 stimuli: fundamental frequencies f_{10}, f_{01}, ratio of the fundamental frequencies f_{01}/f_{10} , detuning Δf_{01} with respect to the corresponding exact integer ratio and name of the sound

f_{10} /Hz	f_{01} /Hz	ratio f_{01}/f_{10}	Δf_{01} /Hz	name
100	128,57...	9/7		100_128_7_9
100	129,3		-0,7	100_129.3
100	130	13/10		100_130_10_13
100	130,7		+0,7	100_130.7
100	132,33...		-1	100_132.33
100	132,66...		-0,66...	100_132.66
100	133,07...		-0,25	100_133.07
100	133,33...	4/3		100_133_3_4
100	133,58...		+0,25	100_133.58
100	134		+0,66...	100_134
100	134,33...		+1	100_134.33
100	135,86...		-0,5	100_135.86
100	136,36...	15/11		100_136_11_15
100	136,86...		+0,5	100_136.86
100	137,5	11/8		100_137.5_8_11

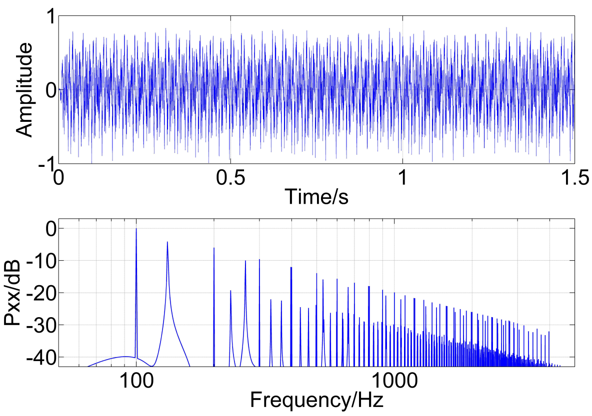

Figure 1: Normalized time series and spectrum of the sound *100_132.33* with the fundamentals' frequency ratio $f_{01}/f_{10} \approx 4/3$ (slightly mistuned)

Participants

37 paid volunteers (21 female, 16 male) with a median age of 21 (min=19a, max=62a) participated in the experiment. All participants reported no hearing problems.

Experimental setup

The experiments take place in a seminar room. The investigator and up to five participants sit around a hexagonal table. The sounds are played back by an audio software (Adobe Audition) with an external soundcard (M-Audio, Fast Track Pro) and an active loudspeaker (Mackie, HR 824). Whithin the software individual tracks for each of the 18 scales are arrangend with different randomizations of the 15 test sounds. Each sound is preceded by a recorded announcement with the corresponding number of the sound (e.g. 'Sound 1'), to avoid confusion in matching the correct scale, e.g. using the wrong page of the booklet. The sounds have a duration of 5 seconds with a pause before the next announcement to allow for making the assessment of the adjective scale. All stimuli are presented at the same constant level of 70 dB(A), with a sampling rate of 22050 Hz and a resolution of 16 bits.

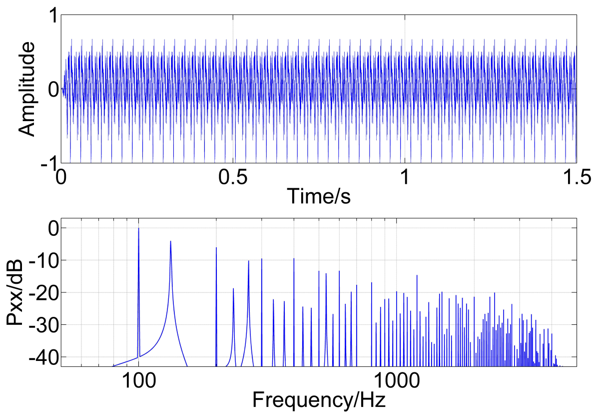

Figure 2: Normalized time series and spectrum of the sound *100_133_3_4* with the fundamentals' frequency ratio $f_{01}/f_{10} = 4/3$ (exact)

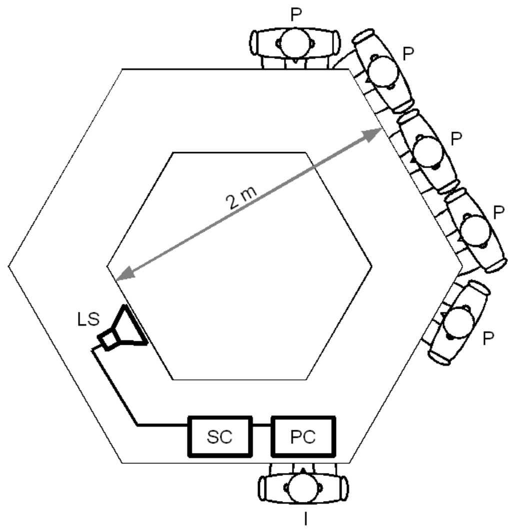

Figure 3: Schematic diagram of the experimental setup (top view): computer **PC**, soundcard **SC**, loudspeaker **LS**, participant **P**, investigator **I**

Experimental procedure

A session starts by handing out written instructions to all participants. After having clarified open questions, an orientation phase follows in playing back the 15 sounds twice in random order. Then the evaluation of the sounds with the semantic differential starts with the first scale booklet.

RESULTS

pleasant - unpleasant

Figure 4 shows the pleasantness rating of all 37 participants for the 15 sounds. The means of the ratings are in the upper unpleasant part of the figure and sound *100_133_3_4* is judged most unpleasant.

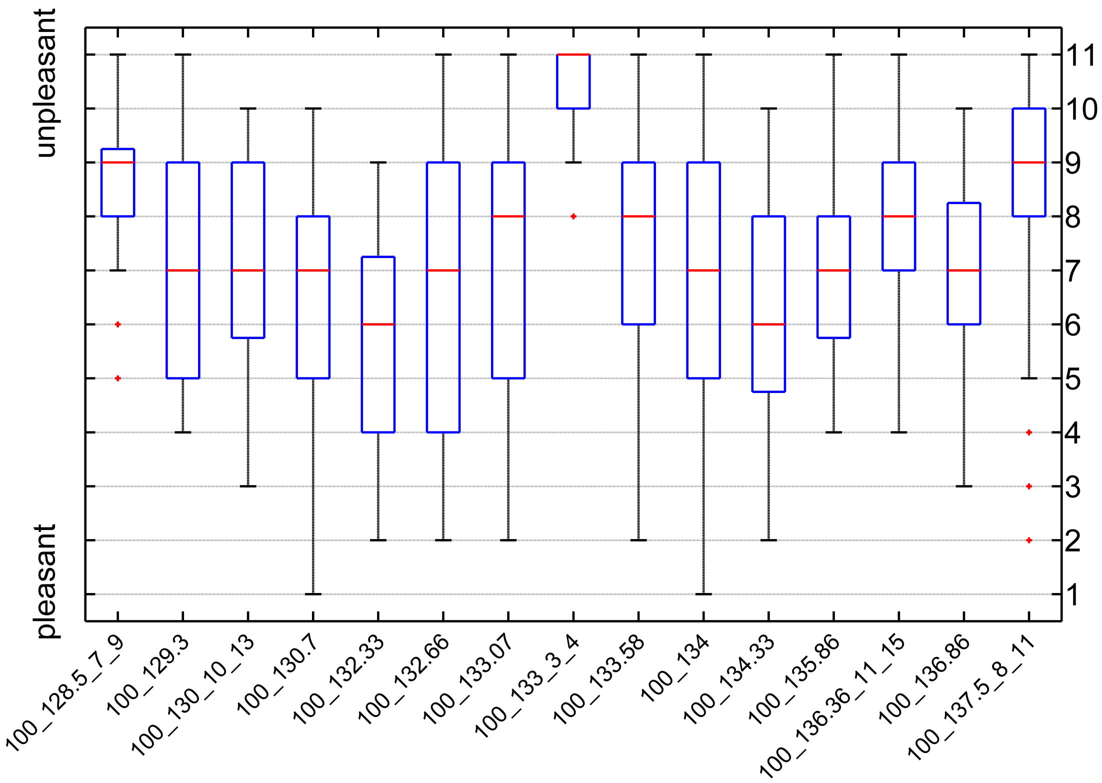

Figure 4: Boxplots of the judgements for the adjective scale *pleasant* - *unpleasant* plotted over the names of the 15 sounds

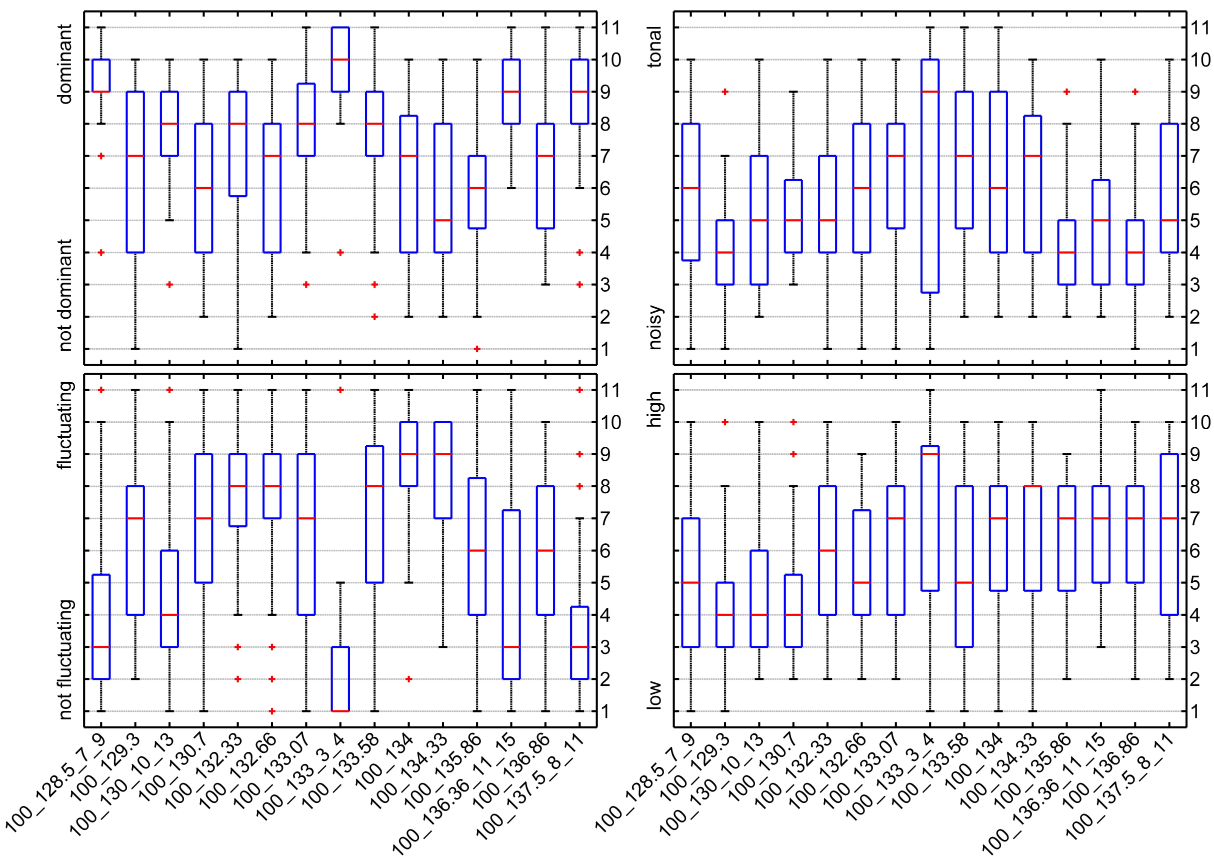

Figure 5: Boxplots of the judgements for the adjective scales *not dominant* - *dominant*, *noisy* - *tonal*, *not fluctuating* - *fluctuating* and *low* - *high* plotted over the names of the 15 sounds

Table 3: Perceptual factors - (varimax-rotated component matrix)

adjective scales	**component**				
	1	**2**	**3**	**4**	**5**
pleasant - unpleasant (rep.)	0,77	-0,01	0,06	-0,03	0,06
not intrusive - intrusive	0,72	0,07	0,01	0,04	0,05
harmonic - discordant	0,69	-0,03	0,30	0,07	-0,04
pleasant - unpleasant	0,66	-0,03	0,14	0,03	0,13
soft - hard (rep.)	0,65	0,13	-0,12	0,39	0,13
soft - hard	0,64	0,053	-0,017	0,396	0,168
not loud - loud	0,61	0,15	0,11	0,04	-0,07
dull - sharp	0,50	0,17	-0,23	0,07	0,33
vague - clear	0,10	0,95	-0,17	0,051	0,060
not dominant - dominant	0,12	0,95	-0,17	0,05	0,07
noisy - tonal	0,06	0,16	-0,71	-0,01	0,24
clean - dirty	0,17	-0,25	0,66	-0,11	0,17
smooth - rough	0,28	0,02	0,58	0,27	0,13
not fluctuating - fluctuating	-0,17	-0,05	0,03	-0,71	-0,13
not hammering - hammering	0,46	-0,03	0,11	0,53	0,13
not functional - functional	-0,38	0,07	0,13	0,52	-0,21
low - high	-0,02	0,15	0,12	-0,07	0,79
not threatening - threatening	0,17	-0,08	-0,10	0,28	0,59
explained variance	22%	11%	9%	8%	7%

The difference in pleasantness between sound *100_133_3_4* and all other sounds is statistically significant (Tamhane post-hoc-test, p<0,05). The sounds *100_132.33* and *100_134.33* are judged as the most pleasant sounds in comparison to the others.

Dimensions of the perceptual space

To explore the dimensions of the perceptual space a factor analysis with the adjective scales as variables has been carried out. As a prerequisite the results of the KMO test (KMO=0.823, 'meritorious') and the Bartlett test of sphericity ($p \leq 0,0001$) prove the adequacy of the data set for a factor analysis. In contrast to the Kaiser and Scree criterion (four factors) a five factor solustion is chosen with an explained total variance of about 57%. It delivers dimensions that can be interpreted well. Table 3 shows the resulting varimax rotated component matrix.

The first factor shows high loadings on the adjective scales *pleasant - unpleasant*, not *intrusive - intrusive*, *harmonic - discordant*, *soft - hard* and *not loud - loud*. It can be regarded as the *pleasant* factor.
The second factor shows high loadings on the adjective scales *vague - clear* and *not dominant - dominant*. It can be regarded as the *powerful* factor.
The third factor shows high loadings on the adjective scales *noisy - tonal*, *clean - dirty* and *smooth - rough*, describing the clearity or tone-to-noise-ratio of the sounds.
The fourth factor shows high loadings on the adjective scales *not fluctuating - fluctuating*, *not hammering - hammering* and *not functional - functional*, which describe the temporal structure of the sounds.
The fifth factor shows high loadings on the adjective scales *low - high* and *not threatening - threatening*, describing the spectral content.
These last three factors - three, four and five - together describe the sound character and can be regarded as the *metallic* dimensions. The results of representative adjective scales for each of the five factors are shown in figures 4 (*pleasant - unpleasant*) and 5 (*not dominant - dominant*, *noisy - tonal*, *not fluctuating - fluctuating* and *low - high*).

Grouping the sounds

In order to find similarities between the sounds a factor analysis is carried out with the sounds as variables. The results of the KMO test (KMO=0.817, 'meritorious') and the Bartlett test of sphericity ($p \leq 0,0001$) manifest the adequacy of the data set for a factor analysis. The Kaiser criterion as well as the scree criterion suggest an extraction of 3 factors. Table 4 shows the resulting varimax rotated component matrix. The first factor shows high loadings on the sounds which are detunings of the 10:13 ratio (*100_129.3* and *100_130.7*) or the 11:15 ratio (*100_135.86* and *100_136.86*). The second factor shows high loadings on the sounds with small integer fundamentals ratios' (e.g. *100_133_3_4*, *100_128_7_9*, *100_137.5_8_11*). The third factor shows high loadings of the 6 sounds which represent detunings of the 3:4 ratio (e.g. *100_132.66*, *100_134*, *100_132.33*).

Table 4: Factor analysis with sounds as variables (varimax-rotated component matrix)

sounds	**component**		
	1	**2**	**3**
100_129.3	0,73	0,15	0,06
100_135.86	0,71	0,09	0,14
100_136.86	0,67	0,19	0,10
100_130.7	0,61	0,11	0,11
100_133_3_4	-0,20	0,80	0,01
100_128_7_9	0,27	0,74	-0,14
100_137.5_8_11	0,30	0,69	-0,14
100_136_11_15	0,43	0,64	-0,04
100_130_10_13	0,45	0,60	-0,12
100_132.66	0,09	-0,17	0,74
100_134	0,04	-0,10	0,72
100_132.33	0,28	-0,12	0,71
100_134.33	0,26	-0,17	0,69
100_133.58	-0,34	0,45	0,63
100_133.07	-0,47	0,37	0,56
explained variance	24%	19%	18%

SUMMARY

The perception of multiple complex tones, consisting of two complex tones and combination tones is evaluated. The frequency ratio of the two fundamentals is the parameter of the test. It is varied in keeping the frequency of the lower fundamental constant and only changing the frequency of higher one. A total of 15 different sounds, composed as a superposition of sinusoids, are judged by 37 participants using a semantic differential of 16 different 11-point categorial adjective scales.

Fundamental frequency ratios corresponding to *ratios of small integers* are judged rather unpleasant. A detuning of ratios of small integers (resulting in a *ratio of large integers*) leads to a rather pleasant judgement. These results are well in line with pleasantness judgements from two former studies (Töpken et al. (2010b), Töpken et al. (2010a)).
A factor analysis on the adjective scales leads to five well interpretable factors. The obtained perceptual dimensions can be meaningfully compared to those timbre factors described by Namba et al. (1991), Namba et al. (1992) and Kuwano et al. (2006). In the actual study the first factor can be regarded as the *pleasant* factor and the second one as the *powerful* factor. The last three factors together can be identified with the *metallic* property, whereas each of these three factors describes a different aspect of the sound character. The third factor describes the clearity, the fourth factor describes the temporal structure and the fifth factor describes the spectral structure of the sounds. The variation of only one signal parameter - the frequeny ratio of the fundamentals - leads to a five dimensional perceptual space, hence showing the multiple perceptual properties of the the multiple complex tones investigated.

REFERENCES

S. Kuwano, H. Fastl, S. Namba, S. Nakamura, and H. Uchida. Quality of door sounds of passenger cars. *Acoust. Sci. Tech.*, 27:309–312, 2006.

S. Namba, S. Kuwano, T. Hashimoto, B. Berglund, Z. Da Rui, A. Schick, H.Hoege, and M. Florentine. Verbal expression of emotional impression of sound: A cross cultural study. *J. Acoust. Soc. Jap.*, 12(1):19–29, 1991.

S. Namba, S. Kuwano, K. Kinoshita, and K. Kurakata. Loudness and timbre of broad-band noise mixed with frequency-modulated sounds. *J. Acosut. Soc. Jap.*, 13(1):49–58, 1992.

S. Töpken, J. Verhey, and R.Weber. Psychoacoustic evaluation of mixed complex tones. In *Internoise 2010*, 2010a.

S. Töpken, J. Verhey, and R. Weber. Psychoakustische Bewertung von Tonkomplexen. In *Fortschritte der Akustik - DAGA 2010*, pages 589–590, 2010b.