Applied experimental psychology and psychophysics

SEMANTIC DIFFERENTIAL

- INTRO TO CENSOR PANEL SOFTWARE -

November 30th. 2017

Group members:

Andreas Kornmaaler Hansen

Emil Bonnerup

Juliane Nilsson

Lucca Julie Nellemann

Sara Nielsen

Group 782
Electronic Systems, Aalborg University
17gr782@es.aau.dk

Introduction

This paper aims to analyse data obtained from Christiansen et al. 2011 where they tested the influence of a loudspeaker's appearance and the perceived sound quality on 20 subjects with 16 bipolar word-pairs on open ended Visual Analogue Scales (VAS). Half of the dataset has been replicated so that more data-points are available. The objective is to analyse the semantic difference between the answers using the censor panel software *PanelCheck* and gain experience working with the software. Multiple plots will be presented and analysed.

Method

The subjects were shown pictures of each set of speakers as they appear in Figure 0.1. They were then asked to rate their attributes on a VAS with 16 bipolar word-pairs.

Figure 0.1. Overview of the speakers used in the data set. They are all from Bang & Olufsen in order to avoid any bias relating to specific brands.

The data is analyzed using the software *PanelCheck*. The analysis is quite explorative, since it is still unknown what we are looking for.

PanelCheck Software

PanelCheck is a free software which is made to be easy-to-use. It is mainly used for visualisation of sensory data which helps to gain insight in a given assessor and panel performance. The software is shown on Figure 0.2.

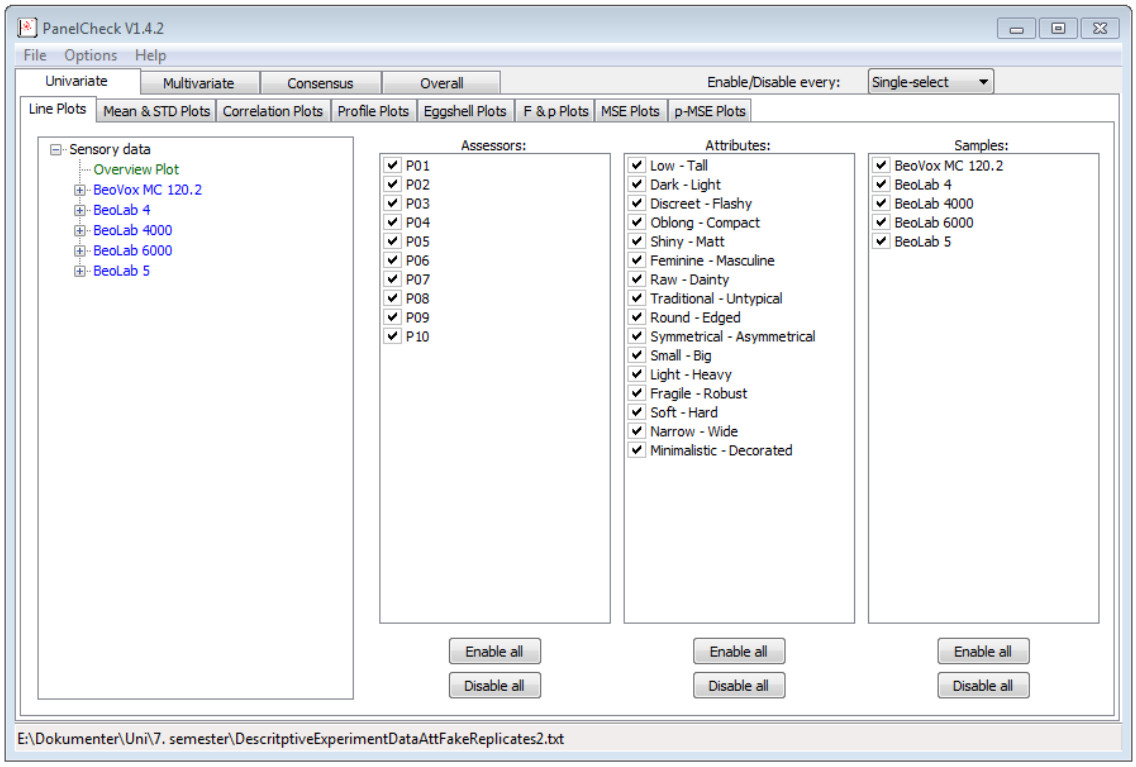

Figure 0.2. The start interface when *PanelCheck* is started and a dataset has been imported. The 16 different attributes are shown along with the five different speakers (Samples). The ten assessors are named from P01-P10. There are four main tabs: Univariate, Multivariate, Consensus, and Overall.

The initial interface contains an overview of the subjects (*assessors*), the word-pairs (*attributes*), and the different speakers (*samples*). At the top it is possible to choose between different tabs: *Univariate*, *Multivariate*, *Consensus*, or *Overall*. *Univariate* and *Multivariate* relates to how many variables are being compared, and *Consensus* relates to the underlying dimensions in the data and therefore contains PCA-related options. As the name implies, *Overall* can be used to gain an overall impression of the data e.g. by showing multiple ANOVA-plots in the same calculation. Within each of the four tabs there exists varying sub-tabs which contains statistical plots relating to the chosen tab. In this example, within *Univariate* it is possible to choose *Line Plots*, *Mean & STD Plots*, *Correlation Plots*, etc. It is also possible to exclude *assessors*, *attributes*, and *samples*. One just have to undo the check marks in the unwanted category as shown in Figure 0.2.

Results and Analysis

Using a 2-way ANOVA, a significant assessor effect ($p<0.01$) has been found on the "Dark - Light" attribute, which means that the assessors disagree about how light/dark the speakers are. See Figure 0.3. This could mean that people do not have the same understanding of what it means for something to be light or dark.

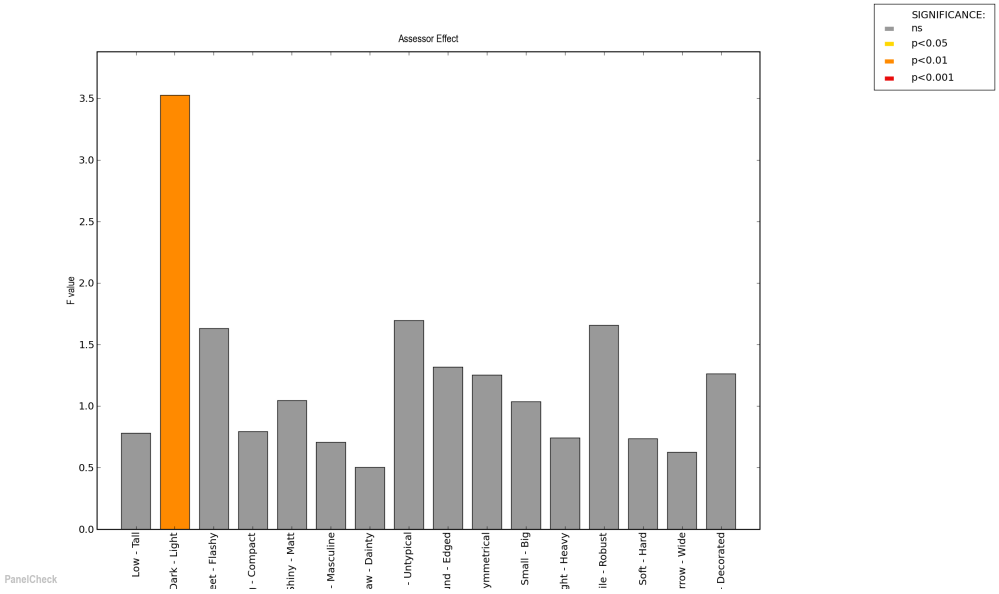

Figure 0.3. A plot showing the assessor effect. The variance was only significantly different in the results relating to the "Dark - Light" attribute, indicated by the orange bar. This means that the assessors did not agree upon which of the speakers were the lightest or darkest.

Regarding the product effect on ratings, all attributes are rated significantly different at at least $p<0.01$, except for "Symmetrical - Asymmetrical". That means that the speaker model has an influence on the ratings of all the attributes, except for symmetry, which is unaffected of which speaker model is rated.

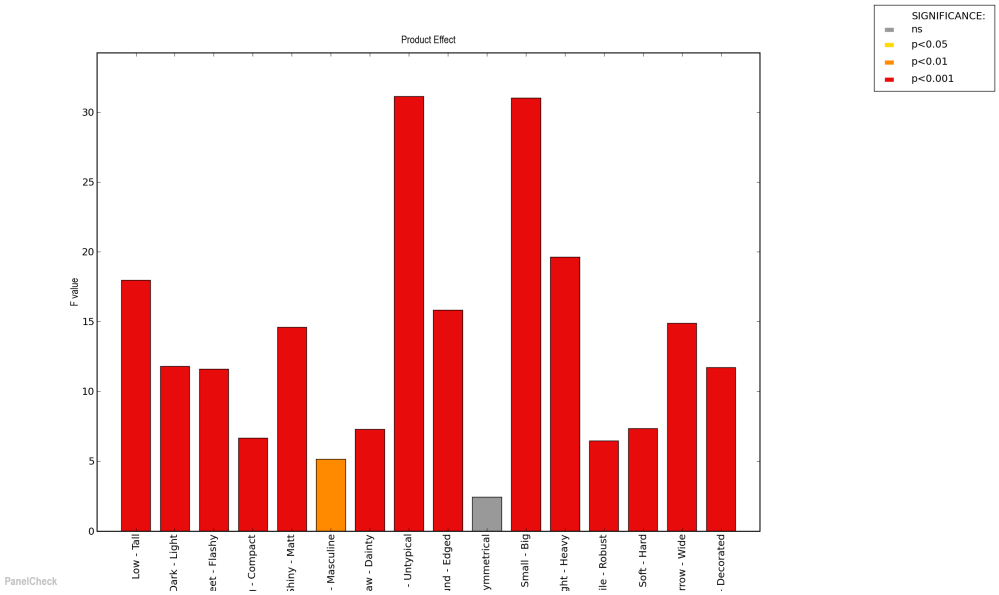

Figure 0.4. A plot showing the product effect. The only attribute which is unaffected by speaker model is the one regarding symmetry, indicated by the grey bar. All other attributes are affected by speaker model, indicated by the red bars for $p<0.001$ and orange bar for $p<0.01$.

This can be confirmed by looking at the means and standard deviations for the different attributes. For example, in Figure 0.5, the symmetry attribute has a very little standard deviation, which means that it is fairly unaffected by speaker model, compared to some of the other attributes. This means that the attribute is unimportant for the rating of the speakers.

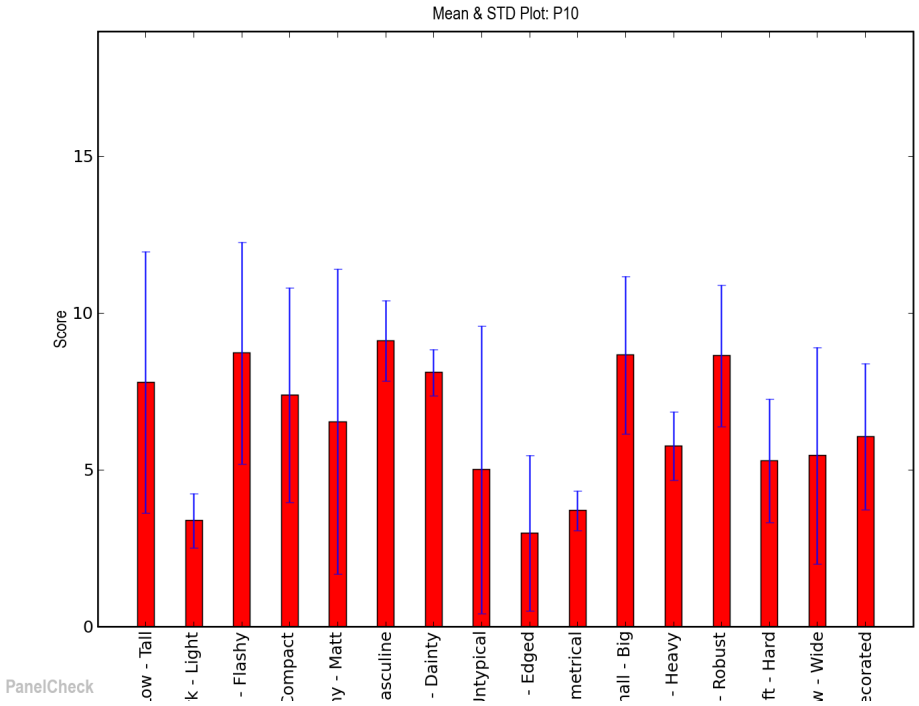

Figure 0.5. A plot showing attribute importance according to their individual mean and standard deviation.

Looking at the product effect alone for all of the five speakers does not tell us anything about which of the speakers where rated different than the others, but only that some of them were. By doing a 2-way ANOVA on only two of the speakers, it is possible to see which attributes separates the two. For example, when comparing the Beolab 4000 with the Beolab 6000 (see Figure 0.1), the main differentiators are: *Low - Tall*, *Light - Dark*, and *Wide - Narrow*, as shown in Figure 0.6. This approach can be repeated across the different combinations of speakers, to assess how each differs from the others.

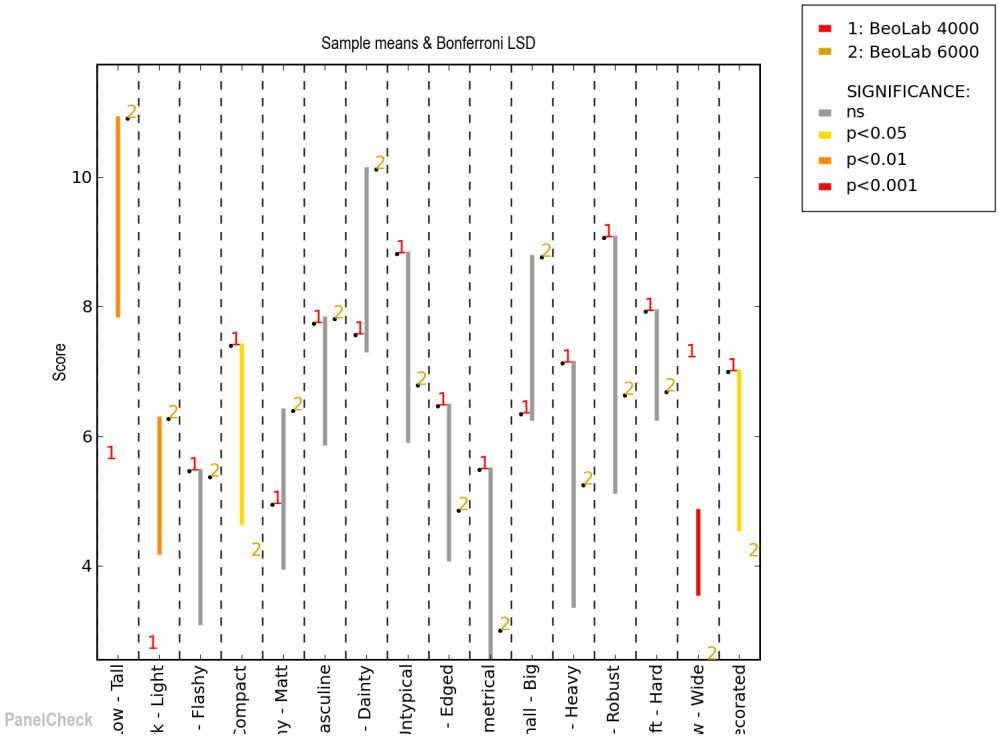

Figure 0.6. Comparison of the attribute ratings for the Beolab 4000 and the Beolab 6000. It is possible to see which of the attributes are significantly different from each other. In this example it is: "Low - Tall", "Light - Dark", and "Wide - Narrow".

It is possible to look at a spider plot in order to get a sense of how the speakers were rated overall, as illustrated on Figure 0.7. Here the left word is the inner point and the right word is the outer point. For example the BeoLab 5 overall scores the highest in the attributes: *Flashy*, *Decorated*, *Wide*, and *Untypical*. This seems reasonable when presented with the speakers next to each other in Figure 0.1. In general this paints a fitting picture regarding which features stand out the most and it creates a unique attribute profile for each speaker when shown individually. Though when shown all together as in Figure 0.7 some points risk ending up on top of each other and some insights may be lost due to this. That being said, a spider plot is generally considered quite useful in displaying multivariate data visually.

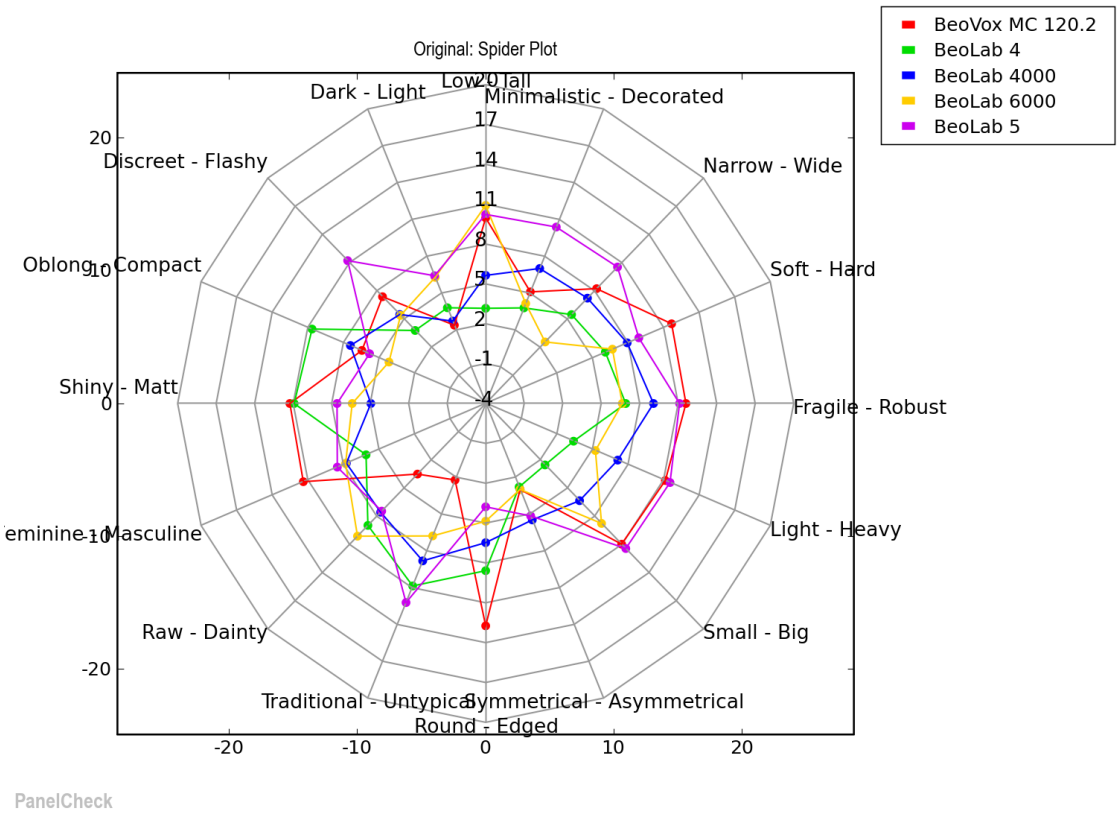

Figure 0.7. A spider plot of the subjects' preference of the five different speakers according to each attribute.

Possible Correlation

The spider plot in Figure 0.7 also paints an overall picture of which attributes might correlate though they are not arranged according to correlation. The plot can also get a bit confusing when multiple points are arranged close or on top of each other as previously stated. Luckily *PanelCheck* contains a *PCA Correlation loadings plot* which presumably arranges attributes with higher correlation close to one another. The plot is shown on Figure 0.8.

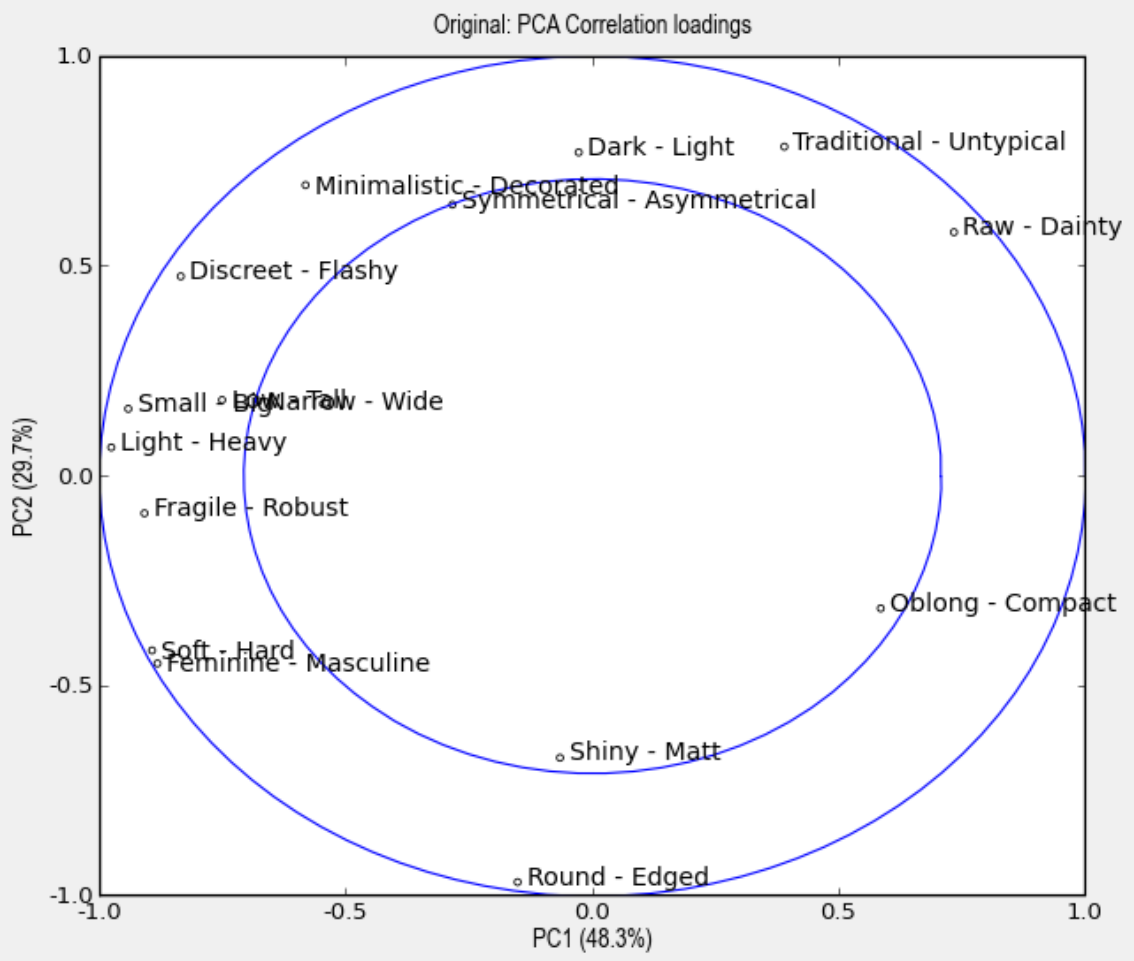

Figure 0.8. A plot containing the PCA Correlation loadings. It becomes clearer which attributes correlates e.g. "Small - Big" and "Light - Heavy".

In this PCA Correlation loadings plot a group appears to the left containing: *Light - Heavy*, *Small - Big*, *Low - Tall*, *Narrow - Wide*, and *Fragile - Robust*. This means that there seems to exist a correlation with these attributes. The group makes sense from an objective point of view where all attributes relates to size except *Fragile - Robust*. However small, narrow, light, etc. might often be experienced to be fragile. The same goes for robust and being big etc. Another clear group is *Soft - Hard* and *Feminine - Masculine*. It could be investigated further if these correlations are strong enough to stand alone meaning, perhaps they mean the same to the assessors and the number of word pairs might be reduced without losing much resolution in the data.

Discussion

Overall it seems like the subjects were able to assign fitting scores for each attribute. The software *PanelCheck* allow easy access to a vast variety of statistical methods and visualisation. However it does take away some of the freedom seen in other software such as *R* or *Matlab*. It is not easy to change the appearance e.g. if another colour scheme is preferable or the labels are badly placed. Often the labels are placed on top of each other making it impossible to read. *PanelCheck* is an easy to use tool but it could use some more documentation on the different plots. It is not hard to imagine using an easy generated plot only to misinterpret it due to this.

Conclusion

PanelCheck is an easy to use software and statistical tool but could do with some more documentation and customisation in the form of editing labels, colours, etc.. With the help of *PanelCheck* it was found that the assessors disagreed on the attribute: *Dark - Light*. The attribute *Symmetrical - Asymmetrical* had a very low standard deviation and the assessors seemed to agree that all speakers shown were symmetrical. All the assessors' answers were significantly different from each other in one or more attribute ratings. By using *PanelCheck* it is possible to compare the attribute ratings for two different speakers to figure out where the difference lies. An example was shown and found that the speakers BeoLab 4000 and the BeoLab 6000 were rated significantly different from one another on the following attributes: *Low - Tall*, *Light - Dark*, and *Wide - Narrow*. A spider plot was used to identify each speakers attribute profile and a plot containing PCA Correlation loadings showed that there most likely exists correlation between the attributes: *Light - Heavy*, *Small-Big*, *Low - Tall*, *Narrow - Wide*, and *Fragile - Robust* and again between *Soft - Hard* and *Feminine - Masculine* respectively where the latter group showed the strongest correlation.

Bibliography

Christiansen, A. S. G. et al. (2011): "Investigation of Loudspeakers' Visual Appearance and its Influence on the Evaluation of Sound Quality". In: