6. General Conclusion

The influence of experimental parameters on the results in experiments concerning equal-loudness level contours is investigated. A model which is proposed by Gabriel [1996] (and fitted to data which are obtained under the same experimental conditions as used by Betke [1991] and Fastl et al. [1990]) is applied to data from the literature. In the model differences between results are attributed to the employed experimental setup. The model takes in particular the absolute position of the employed test-tone level range influences the resulting point of subjective equality (PSE) into account. Gabriel [1996] adjusted the model parameters to minimize the differences between data which are determined under comparable conditions (e.g. by Betke and Fastl for the 30 phone equal-loudness level contours ($ELLC$)). The model with this set of parameters fails to minimize differences between data from literature for other $ELLCs$. Even for the original data for the 30 phon contour the aim to minimize differences can only partly be achieved. Although for some frequencies the calculated curves seem to differ less, for others the difference is increased.

A modification of the determination of model parameters is investigated. Unreasonable data can be avoided by modifying the calculation of the REL value, which reflects the presented test-tone levels in the measurement procedure. Instead of calculating REL, as proposed by Gabriel [1996], from the maximum and minimum of the presented level with a weighting of 80% to 20%, a substitution by using a loudness calculation with the N_{35} percentile avoids unreasonable low values. But the 'correct' (i.e. 'absolute')

ELLC-values remains still unclear. Since only some differences could be explained by the model, the requirements and limitations of the parameters involved are investigated. Is is shown that for some frequencies it is impossible to minimize the differences between *ELLC* from different studies with this model.
Considering the findings of Gabriel and respecting the recommendations from ISO/ TC43/ WG1/N122 [1988], an adaptive procedure is proposed for the determination of *ELLC*. First results obtained with the method of constant stimuli and with the adaptive procedure are compared. The influence of the measurement setup on the results is smaller in comparison to the results obtained with the method of constant stimuli.
The parameters which have to be set by the investigator in the adaptive procedure are systematically varied. The influence on the resulting *PSE* is quantified and it is shown that a large initial step-size in the adaptive procedure minimizes the influence of the starting level. While a large initial step-size has the disadvantage that the test-tone can get too loud or below threshold of hearing after a few steps, a moderate initial step-size of 8 dB is proposed. It is shown that in a simple adaptive procedure the starting level can still influence the result. Even with two interleaved tracks the results are biased. Instead of presenting one adaptive track, at least 4 frequencies with different starting levels have to be randomly presented in an interleaved procedure in order to minimize the influence of the experimental parameters. The influence of the starting level on the resulting *PSE* can be reduced to the same order of individual reproducibility. Therefore, the proposed method is insensitive enough against the experimental setup and can be used for the determination of equal-loudness level contours.
Verhey [1999] reported that the influence of the starting level almost vanishes for shorter signals (10 ms). Maybe due to the fact that these short signals do not produce a hearing sensation of pitch, the task is equal difficult for all frequencies. Heldmann [1994] claimed that for the impression of pitch 35 ms are necessary (frequency > 100 Hz). In might be concluded that bias effect in the determination of *ELLC* can be reduced by taking

shorter signals. From Gabriel[1996] it is known that the bias is largest for the largest differences between test and reference-tone. This is because for subjects it is most difficult to compare tones which differ in pitch.
The reason for the bias remains unclear. At least from one experiment it is concluded that the process can be attributed to a more central mechanism in the processing of acoustic sensation. Although it is clear that it is a difficult task for the subject comparing two different stimuli, measurements are repeated with high reproducibility. But the variance between subjects is rather large. All the data from the literature show that the variance between individuals (intraindividual variance) is much smaller than differences between different subjects (interindividual variance). This variance can partly be minimized when taking into account the individual threshold of hearing. When subtracting the individual threshold of hearing from the individual PSE for each frequency, at least at low levels and low frequencies the standard deviation is reduced about a factor of two. The obtained equal-loudness level contours can be interpreted as *equal-loudness level contours sensation level level* $(ELLC - SL)$.

The proposed measurement setup for the determination of new $ELLCs$ should be restricted to the proposed schemes in order to avoid biasing the results by the employed procedures. The results of this investigation will help to determine comparable loudness level contours in different laboratories under free-field conditions in order to be used for a revision of the contours for a new standard.