Suprathreshold perception of tonal components in noise under conditions of masking release[*]

Jesko L. Verhey [†]

Department of Experimental Audiology
Otto von Guericke University Magdeburg
Leipziger Str. 44, 39120 Magdeburg, Germany

Stephan J. Heise

Institute of Physics
Carl von Ossietzky University Oldenburg
26111 Oldenburg, Germany

Abstract

Many technical sounds contain tonal components in a noise background. The present study investigates how such tonal components embedded in noise are perceived under conditions of enhanced detectability of the tone due to (i) amplitude modulation of the noise background or (ii) interaural disparities in the signal's phase. To quantify the sensation elicited by a tonal component in these masking-release conditions, the listeners are asked to adjust the level of the tone to have the same magnitude of tonal content (German "Tonhaltigkeit") or to be perceived as equally loud as a tone which is masked by an unmodulated noise and presented diotically (baseline condition). For both sensations the signal-to-noise ratio at equal magnitude of the sensation is lower for the tone in the masking-release conditions than for the baseline condition. This is mainly due to the lower masked threshold of the tone in masking-release conditions. The effect is most prominent at low levels and decreases towards higher levels. A high correlation is observed between the data for the magnitude of tonal content and the partial loudness of the tone, demonstrating the similarity between these two sensations.

PACS numbers: 43.50.Rq, 43.66.Cb, 43.66.Dc, 43.66.Mk, 43.66.Rq, 43.66.Ba

1 Introduction

In general, environmental sounds containing tonal components are considered to be more annoying than sounds without these components (e.g. DIN 45681, 2005). Many technical sounds contain tonal components originating from rotating parts, such as wind turbines, electric engines, aircrafts, ventilation, and tire/road noise. Thus tonal components are important when assessing noise pollution. On the other hand, the perception of tonal components is important for sound design. For example, the reduction of the random noise background in hybrid or electric cars compared to conventional combustion engines increases the number of audible tonal components in the car. The low noise level of hybrid and electric cars has also promoted a discussion of the necessity of artificial warning sounds. It is probable that these warning sounds will contain clearly audible tonal components. In several countries acoustical alarms with a tonal character are already common as warning signals for reversing vehi-

cles (Burgess and McCarthy, 2009). How we perceive audible tonal components in background noise is thus also an important question for the assessment of the perception of warnings sounds in urban environments.

Several standards include sections dedicated to the assessment of tonal components in sound (DIN 45681, 2005; IEC 64100-11, 1998; ANSI S1.13, 2005; ISO 1996-2, 2007). These standards have in common that they estimate the magnitude of the tonal components (tonal foreground) relative to the noise background. This magnitude is either derived from a comparison of the level in a critical band with the levels in adjacent critical bands (prominence ratio, ANSI S1.13, 2005) or by identifying tonal components within a critical band and comparing their magnitude with the intensity of the noise background in this critical band (e.g., tonality, IEC 64100-11, 1998, tone-to-noise ratio, ANSI S1.13, 2005).

The present study focusses on the German standard DIN 45681 (2005). The standard uses the term "Tonhaltigkeit" for audible tonal components in a noise background. An appropriate translation of this term is "magnitude of tonal content" (see Hansen *et al.* (2011), for a review). In the standard, the procedure used to calculate the magnitude of tonal content is similar to

---

[†]Jesko.Verhey@uni-oldenburg.de

the one used for tonality or tone-to-noise ratio. The level of the tonal components L_T and that of the noise background L_G are estimated within each critical band. To calculate the magnitude of tonal content at the frequency f, masking index a_V are subtracted from the level difference between tone and noise.

$$\Delta L = L_T - L_G - a_V \quad \text{in dB} \tag{1}$$

$$\text{with } a_V = -2 - \log\left(1 + \left(\frac{f}{502}\right)^{2.5}\right) \quad \text{in dB} \tag{2}$$

In a second step, the higher annoyance of sounds with tonal components is addressed by making a tone adjustment (German "Tonzuschlag") which depends on the magnitude of tonal content (see table 1 in DIN 45681, 2005)[1]. Up to a ΔL of 6 dB the tone adjustment increases by 1 dB every 2 dB. From 6 to 12 dB an increase of the tone adjustment of 1 dB only occurs every 3 dB. Thus the relation between tone adjustment (due to annoyance) and magnitude of tonal content is compressive. In the framework of the standards the perception of tonal components is restricted to the annoyance of the tones. How to explicitly measure the sensation associated with magnitude of tonal content is not addressed in the standards.

The aim of the present study is twofold. On the one hand it is investigated whether the sensation of the magnitude of tonal content can be assessed by asking listeners to rate the loudness of the tonal portion of the sound. This would facilitate future experimental studies on the magnitude of tonal content since most listeners are more familiar with the term loudness than with the term magnitude of tonal content. In addition, it would have implications for the standards of the perception of tonal components in noise. The current standards have in common that they are based on the analysis of intensities. If the magnitude of tonal content is equal to the partial loudness of the tonal components in noise, a better description of its perception may be achieved if the nonlinear relation between intensity and loudness is taken into account, i.e., if the standards for the magnitude of tonal content are linked to the standards for the sensation loudness. The second aim is to investigate the change in the magnitude of tonal content when the masking of the tonal component is reduced due to the presence of additional cues. In a first experiment a masking release is introduced by modulating the noise background. Since the masking release is particularly large for bandwidths considerably larger than the critical bandwidth centred at the tonal signal, it was initially assumed that the masking release results from a comparison of the temporal level fluctuations in different critical bands (Hall *et al.*, 1984, however see Verhey *et al.*, 1999; Verhey and Ernst, 2009). For a modulated masker these level fluctuations are coherent, i.e., they are comodulated. Therefore this type of

masking release is commonly referred to as comodulation masking release (CMR, Hall *et al.*, 1984; Verhey *et al.*, 2003 for a review). In a second experiment a release from masking was introduced by changing from a diotic presentation of tonal target signal and noise masker to a dichotic presentation of the target signal. This masking release is commonly referred to as binaural masking level difference (BMLD, Jeffress *et al.*, 1956).

In the present study artificial stimuli are used to generate a large magnitude of masking release and to facilitate the comparison with previous psychoacoustical studies. It is likely that comparable masking-release conditions occur in our everyday acoustical environment, for example in a car, when electrical equipment such as an electrically operated window regulator generates a localized tonal sound which is embedded in diffuse broadband wind and road noise.

2 Experiment 1: Perception of tones in modulated noise

2.1 Methods

2.1.1 Listeners

Nine listeners (2 female, 7 male) participated in the experiment. None of them reported any hearing difficulties. Their age ranged from 23 to 32 years (26.8 years, on average).

2.1.2 Apparatus and stimuli

Stimuli were generated digitally at a sampling rate of 44.1 kHz. A standard personal computer controlled stimulus generation and presentation and recorded results using a software package developed at the University of Oldenburg. Stimuli were D/A converted (RME ADI-8 PRO), amplified (Tucker-Davis HB7), and presented via Sennheiser HD 650 headphones. Listeners were seated in a double-walled, sound-insulated booth.

The target signal was a 986-Hz pure tone with a duration of 600 ms including 50-ms raised-cosine ramps at signal onset and offset. The target signal was temporally centred in a masking noise. The duration of the masker was 700 ms including 50-ms raised-cosine ramps at masker onset and offset. The masker level was 65 dB sound pressure level (SPL). The band-limited noise had a lower cut-off frequency of 250 Hz and an upper cut-off frequency of 4000 Hz. The noise was generated in the frequency domain by transforming a 700-ms white noise time signal into the frequency domain via a Fast Fourier Transform (FFT) and setting all Fourier components outside the desired passband to zero. A subsequent inverse FFT on the complex buffer pair yielded the desired noise waveform.

The masker was either an unmodulated or an irregularly rectangularly modulated noise with a mean modulation frequency of 40 Hz (as used in Verhey and Ernst, 2009 and Ernst *et al.*, 2010). A schematic plot of the generation of the irregular square-wave modulator and
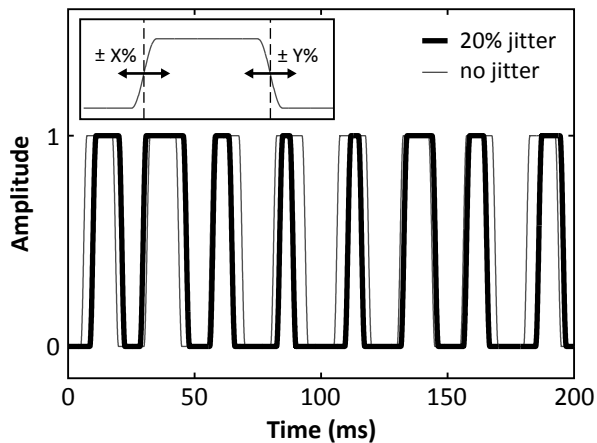
---

[1]Such a tone adjustment due to the presence of audible tonal components is also used in other standards. For example, the Australian environmental noise guidelines for wind farms add 5 dB(A) to the noise level if tonality is a characteristic of the noise (EPA, 2009).

Figure 1: Masker modulation used in the first experiment. An example of an irregular modulator with a 20 % jitter is shown with a thick black line. For comparison, a regular modulator (no jitter) is shown with a thin grey line. The inset schematically shows for one period how an irregular square wave modulator is generated. Starting with a regular square wave modulator, onset and offset were jittered independently. The magnitudes of the jitter (X and Y) had to be smaller than or equal to 20 %.

an example of a masker modulator are shown in Figure 1. The modulation was a unipolar square wave $(0, 1)$, i.e., the carrier was switched on and off by the modulator. It has been shown that square-wave modulators give rise to larger CMR than low-pass noise modulators, the latter being a modulator type that was commonly used in CMR experiment (e.g., Carlyon *et al.*, 1989; Verhey *et al.*, 1999). The average duty cycle of the regular square wave was 50 %, i.e., for each 25-ms period, the signal was switched on for half of the time. In order to generate the irregular square waves, onset and offset times were slightly jittered (see Figure 2 in Verhey and Ernst, 2009). Jittering the modulation rate was introduced to avoid a periodicity pitch in the masker. The magnitude of the jitter was 20 % of a period, which introduced random fluctuations of the duty cycle in the range from 10 to 90 %. On- and offset times were jittered independently. The square-wave modulator was convolved with a 5-ms raised-cosine window in order to avoid onset and offset clicks. The modulated masker was generated by multiplying the modulator with the band-limited noise. For each presentation of the masker a new noise sample and, in the case of the modulated masker, a new modulator sample were generated.

2.1.3 Procedure

The experiment was performed in three steps. In the first step, the masked threshold of the tone was measured for both masking conditions. A three-alternative, forced-choice (3-AFC) procedure with adaptive signal-level adjustment (1-up 2-down) was used to determine the thresholds (Levitt, 1971). Intervals in a trial were separated by 500-ms silence intervals. The tonal target signal was added to one of these intervals. This signal

interval was randomly selected for each trial. Listeners had to indicate which of the intervals contained the signal. Visual feedback was provided after each response. The signal level was initially adjusted in steps of 8 dB. After the first two upper reversals, the step size was set to 2 dB and after the third upper reversal the step size was set to the minimum step size of 1 dB. The run terminated after a further six reversals. The average of these six last reversals was taken as an estimate of the threshold. The procedure was repeated twice for each condition and the average of the two estimates was taken as the final estimate of the threshold.

In the next step, the perception of tonal components was measured at suprathreshold levels. The tone level at equal magnitude of tonal content was measured using an adaptive two-alternative, forced-choice procedure with a 1-up 1-down rule. Each trial consisted of one interval with a tone embedded in unmodulated noise and one interval with a tone embedded in modulated noise. Within a run only one tone level was varied, either the level of the tone in unmodulated noise or the level of the tone in modulated noise, while the other interval (reference) stayed fixed in level. In order to prevent listeners from focussing on the overall loudness of the stimuli in a run, the total level of signal plus masker was varied randomly from trial to trial in a range of 3 dB around the level which would have been used without level roving. The order of the intervals was randomized for each interval. After the presentation of the two intervals of the trial, listeners were asked which of the intervals had the higher magnitude of tonal content (German: "Welcher Stimulus klang tonhaltiger?"). Listeners had to indicate the interval with the higher magnitude of tonal content by pressing the corresponding button on the keyboard. For each adaptive track, the starting level of the tone was randomly chosen in the range from 8 dB below to 8 dB above the individual masked threshold. The tone level was initially adjusted in steps of 8 dB. The step size was halved at each upper reversal until the minimum step size of 2 dB was reached. The run continued for four reversals with the minimum step size. The mean of these four last reversals was taken as the level of equal magnitude of tonal content. This level was estimated for reference stimuli of both masker types (modulated and unmodulated noise) and at five levels above the respective individual threshold of the tone: 3, 6, 9, 12, and 18 dB. To reduce potential bias effects, all ten adaptive tracks were interleaved (e.g. Buus *et al.*, 1997; Verhey *et al.*, 1999). This procedure was repeated three times. In general, the average over the levels obtained in each run was taken as the final estimate of the level at equal magnitude of tonal content. If the level in a run was below the masked threshold it was not included in the average.

In the last step of the experiment, the level at equal loudness of the tone was estimated using essentially the same procedure as described for the estimation of the level at equal magnitude of tonal content. In contrast the second part of the experiment, listeners were asked to indicate the interval containing the louder

tonal component (question in German: "Bei welchem Stimulus klang der Ton lauter").

2.2 Results

Figure 2 shows individual data for the nine listeners participating in the experiment. For all listeners the cross indicating the masked thresholds is below the diagonal, i.e., as expected, all listeners have a lower threshold for the signal embedded in the modulated noise than for the baseline condition with unmodulated noise. There are individual differences in level at threshold and in the magnitude of masking release. Thresholds for the masking-release condition range from 28 dB SPL (listener SJH) to 34 dB SPL (listener AH) and are, on average, 31 dB SPL. For the baseline condition listener AH has the lowest threshold of 46 dB SPL and listener BL the highest threshold of about 50 dB SPL. The majority of the remaining listeners (AO, CHU, JH, JS, SJH, and SK) has a threshold of 47 dB SPL, which is also the average threshold for all listeners. The masking release ranges from 12 dB (listeners AH and JH) to 19 dB (e.g., listeners BL, SJH, and JV).

In general, for suprathreshold tonal components, the data points for the magnitude of tonal component (circles) are below the diagonal, i.e., the sound pressure level of the tone in the masking release condition is always below that of the tone in the baseline condition. The level at equal magnitude of tonal content is about the same for both masking conditions at the highest level. There are individual differences in the shape of the curve of equal magnitude of tonal content. Some listeners show an approximately linear increase with a slope higher than one for the whole range of tone levels used (e.g., listener AO). Other listeners show a linear increase parallel to the diagonal for low levels above threshold followed by steeper increase at higher levels (e.g., listener SK). Standard deviations are only shown if they are larger than the size of the symbols (i.e., larger than 1.5 dB). In general standard deviations are small (2.3 dB, on average). However, for listener JS and JV, standard deviations are up to 10.3 dB, indicating that these listeners had difficulties with the task.

The data for the partial loudness of the tone (squares) are very similar to those for the magnitude of tonal content. The correlation between the two data sets is shown in the top left corner of each panel; it ranges from $\rho = 0.88$ to $\rho = 0.99$. The average standard deviations are slightly smaller than for the data of equal magnitude of tonal content (1.7 dB). The listeners JS and JV showing difficulties in the estimation of levels of equal magnitude of tonal content do not show a comparable uncertainty when adjusting the level to elicit the same partial loudness of the tonal component.

Figure 3 shows mean data averaged across the nine listeners. Data representation in the bottom panel of the figure is the same as in Figure 2, i.e., levels are expressed in dB SPL. The top panel shows the same data but now expressed in dB relative to the masked threshold. For this representation, the data are close
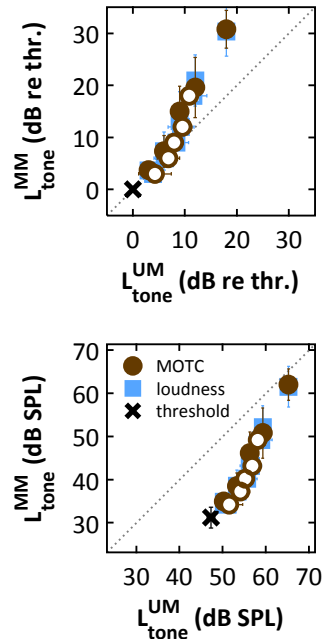

Figure 3: (colour online): Mean data of the first experiment for the magnitude of tonal content and the partial loudness of the tone. Averages were taken for each reference condition across the individual data of the nine listeners. Symbols are the same as used in Figure 2. The upper panel shows the data in dB relative to the masked threshold, while the bottom panel shows the data in dB SPL (as in Figure 2).

to the diagonal for levels up to about 9 dB. At higher levels the level of the tone above threshold in the modulated masking condition is higher than that of the tone in the baseline condition. For the data representation with the levels expressed in dB SPL (bottom panel) the levels at equal magnitude of tonal content asymptotically reach the diagonal at very high levels. A similar behaviour is found for the loudness data, which is expected since the correlation between loudness data and data for the magnitude of tonal content is already high on an individual basis. The correlation between the average data for the magnitude of tonal content and that of the partial loudness of the tone is higher than 0.99.

3 Experiment 2: Perception of dichotic tones in diotic noise

3.1 Methods

Ten listeners (3 female, 7 male) participated in the experiment. None of them reported any hearing difficulties. Their age ranged from 20 to 28 years (24.8 years, on average). Three of them had already participated in the first experiment.

The tonal component of the stimuli in the experiment was a 500-Hz pure tone which was either presented diotically or with an interaural phase difference of π. The masker was a Gaussian noise with a
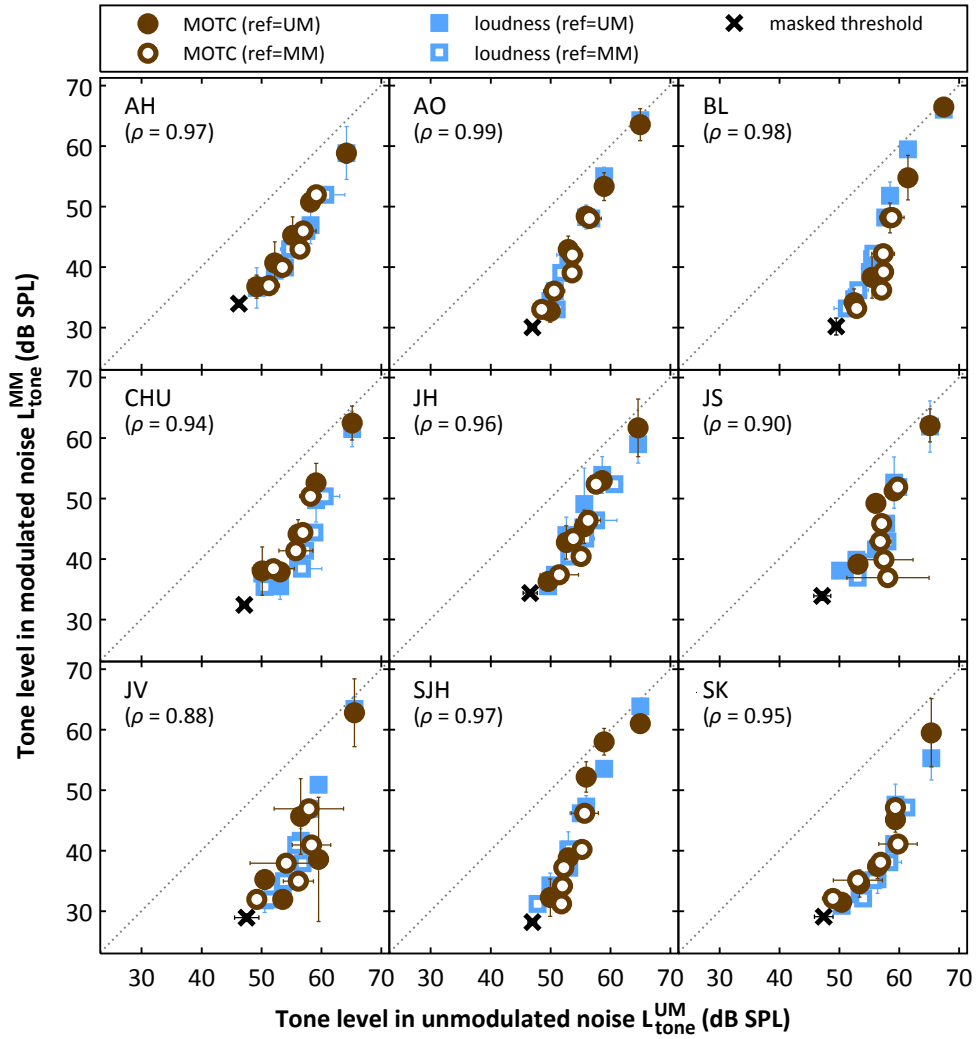
Figure 2: (colour online): Individual data for the magnitude of tonal content (MOTC, circles) and partial loudness of the tone in noise (squares) for the first experiment. In addition masked thresholds are indicated by a cross. Each panel shows data for one listener. The level of the tone embedded in the modulated masker (MM) is plotted versus the level of the tone embedded in the unmodulated masker (UM). Open symbols indicate the data where the level of the tone masked by the modulated masker was fixed and the level of the tone embedded in the unmodulated noise was adjusted to elicit the same sensation of magnitude of tonal content or loudness. Filled symbols indicate the symbols where the tone masked by the modulated noise was varied while the level of the tone masked by the unmodulated noise was fixed. Levels are expressed in dB SPL. Error bars indicate individual standard deviations.

white spectrum in the range from 200 and 4000 Hz. The band-limited noise was generated in the frequency domain as described for the first experiment. The masker duration was 700 ms including 50-ms raised-cosine ramps at masker onset and offset. The masker level was 65 dB SPL. The 600-ms tone was temporally centred in the masker. The signal was gated on and off with 50-ms raised-cosine ramps.

In general, the experimental procedure was the same as used in the first experiment, i.e., in a first step, masked thresholds were measured for the diotic and the antiphasic tone in the presence of the band-limited noise masker. In the following steps the levels at equal magnitude of tonal content and equal partial loudness of the two tonal signals were measured using the same procedure as in the first experiment. The levels of the fixed-level tones were again 3, 6, 9, 12, and 18 dB above individual threshold. In contrast to the first experiment, the term "magnitude of tonal content" was introduced prior to the experiment. Listeners were presented five sounds, two of which were taken from Vormann *et al.* (1998) (their sounds 2 and 3)[2], another one contained a tonal component in noise that was reduced in level over time, and the last two sounds were a combination of a tone and a noise with two different signal-to-noise ratios.

3.2   Results

Individual differences in masked thresholds and suprathreshold perception are similar to the ones shown in Figure 2 for the first experiment (not shown). Figure 4 shows average data for the second experiment. The threshold for the tonal component is, on average, 30 dB SPL for the dichotic presentation and about 42 dB SPL for the diotically presented tone, i.e., the average BMLD for this experiment is 12 dB (see bottom panel of Figure 4). Thus the masking release is slightly smaller than the one of the first experiment. The general characteristics of the data on the perception of the suprathreshold tones is very similar to those found in the first experiment. For levels of up to about 9 dB above masked threshold, the magnitude of tonal content in the diotic condition is about the same at the same level above threshold in the dichotic condition (top panel of Figure 4). At higher levels the level above masked threshold is higher in the masking-release condition than in the baseline condition. At the highest level of the diotic tone used in this study, the same magnitude of tonal content is obtained at about the same sound pressure level of the tone in the two conditions. The data for the partial loudness of the tone is highly correlated with the data for the magnitude of
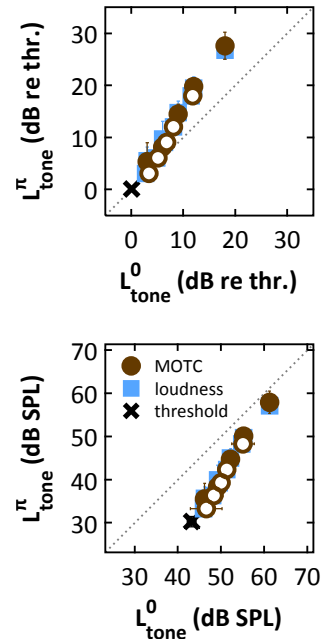
[2] The reference sounds in Vormann *et al.* (1998) were 30 s long and consisted of a 700-Hz tone embedded in pink noise. For the second sound of Vormann *et al.* (1998), the level of the tone was rising for the first half of the signal and falling for the second half of the signal while the level of the noise was constant. For the third sound of Vormann *et al.* (1998), the level of the noise was falling for the first half of the signal and rising for the second half of the stimulus while the level of the tone was constant for the whole duration of the stimulus.

Figure 4: (colour online): Mean data of the second experiment for the magnitude of tonal content and the partial loudness of the tone. Averages were taken for each reference condition across the individual data of the ten listeners. The level of the tone in the dichotic masker is plotted versus the level of the tone in the diotic masker. Data representation is the same as in Figure 3.

tonal content (correlation $\rho > 0.99$).

4   Discussion

4.1   Comparison with previous studies

The masking releases found in the present study are comparable to those found in the literature. Several previous studies have shown a masking release for broadband noise maskers that were modulated with a square-wave modulator (Carlyon *et al.*, 1989; Hall *et al.*, 1996; Verhey *et al.*, 1999; Verhey and Ernst, 2009; Ernst *et al.*, 2010). The CMR can be as large as 40 dB for a modulation rate of 10 Hz (Hall *et al.*, 1996). For the modulation rate of 40 Hz used in the present study, the masking release ranged from 15 dB (Ernst *et al.*, 2010) to about 20 dB (Carlyon *et al.*, 1989; Hall *et al.*, 1996), the latter using a regular square-wave modulator. For the same amount of jitter used in the present study (20 %), Verhey and Ernst (2009) obtained a masking release of 17 dB. Thus the masking release of 16 dB of the first experiment is in agreement with the results of previous studies.

A change in the interaural phase of the signal from 0 to π (experiment 2) resulted in a BMLD of 12 dB. The same BMLD was found, e.g., in Hall *et al.* (1983) at a comparable masker spectrum level (30 dB SPL). The BMLD in the present study is slightly smaller than the BMLD of 15 dB for a comparable condition in van de Par and Kohlrausch (1999) and Nitschmann *et al.* (2009).

Only very few studies have investigated suprathreshold perception in conditions of masking release. There are three studies comparable to the experiments of the present study. Townsend and Goldstein (1972) used a loudness balance procedure to estimate the level difference at equal loudness between a masked diotic tone and a masked tone with interaural phase difference of π embedded in a broadband noise. For a frequency of 500 Hz, they found that the benefit due to the binaural masking release deteriorates at a sound pressure level that corresponds to a sensation level of about 20 dB for the diotic signal. This result is comparable to the present study: At a sensation level of 18 dB for the diotic signal, almost the same sound pressure level at equal loudness for the diotic and dichotic signals was found. A similar result was found in Soderquist and Shilling (1990), who measured equal loudness level contours for tones in the presence of a noise masker. The 100-Hz standard tone was always presented diotically. The frequency of the comparision tone was either 200, 300, 400, 500, 750, or 1000 Hz and the interaural phase was either 0 or π. Levels were considerably lower for the dichotic signals than for the equally-loud diotic signals at low levels of the diotic standard. However, for levels of 20 dB sensation level for the diotic standard, the equal loudness levels were the same for diotic tones and dichotic tones with an interaural phase of π.

For a 250-Hz tone, Fastl and Zwicker (2007) showed levels of a monaurally presented tone that is equally loud as a binaurally presented tone embedded in noise. The interaural phase difference was either 0 or π. At low levels, they found differences in the level at equal loudness which they attributed to the BMLD. This difference decreased as the level increased. A residual effect was found for sensation levels as high as 40 dB. In the present study comparable results are found at low sensation levels. The difference between the present results and those in Fastl and Zwicker is presumably mainly due to differences in the signal frequency. Townsend and Goldstein (1972) showed that for 250 Hz the effect of masking release is stronger than for 500 Hz.

4.2 Magnitude of tonal content and partial loudness of the tone

Vormann *et al.* (1999a) showed that the magnitude of tonal content ("Tonhaltigkeit") can be considered a sensation. Since the term is not commonly used, it was discussed whether an orientation phase should precede experiments on the magnitude of tonal content (Vormann *et al.*, 1998). However, even with an orientation phase listeners report more difficulties in rating the magnitude of tonal content than in rating the partial loudness of the tonal component (Hansen and Weber, 2011).

Using a similar adaptive experimental paradigm as used in the present study, Vormann *et al.* (1999b, 2000) showed that the magnitude of tonal content increases as the number of equally audible tones, each falling into a different critical band, increases. This increase is sim-

ilar to the increase in the partial loudness of the tone (Vormann, 2011). In a parametric study, Hansen and Weber (2011) measured the magnitude of tonal content and partial loudness of a two-tone complex embedded in noise with different levels of the two tones. The data for the two sensations were very similar, indicating that the magnitude of tonal content may be measured by asking the listeners to rate the partial loudness of the tone. In agreement with their data, the present data show a high correlation between the data for the magnitude of tonal content and the results for the partial loudness of the tone. Thus, for an experimental assessment of the magnitude of tonal content of a stimulus, listeners may be asked to rate the loudness of the tonal portion of the stimulus rather than the magnitude of tonal content. Such a procedure may overcome the difficulties in rating the magnitude of tonal content in conditions where the noise is barely audible. For the extreme condition of a pure tone without noise, Vormann *et al.* (1999a) found that some listeners rated a pure tone to have a rather low magnitude of tonal content since they could not contrast the level of the tonal foreground to that of a noise background.

Whereas only a few studies compared the loudness of tones in different masking conditions, there are several studies where a masked tone was matched in loudness to a tone presented in quiet (Steinberg and Gardner, 1937; Zwicker, 1963; Gleiss and Zwicker, 1963; Hellman and Zwislocki, 1964; Stevens and Guirao, 1967). The characteristics of their results are similar to those of the present study: The effect of the elevated threshold in the masking condition is most prominent close to the masked threshold and decreases towards higher levels. At very high levels, the level of the masked tone and the equally loud tone in quiet are similar and an increase of 1 dB of the masked tone results in an increase of 1 dB of the equally loud tone in quiet. At low levels, the level of the tone in quiet has to be increased by substantially more than 1 dB to match a 1-dB increase of the masked tone. The transition depends on the masker spectrum, the masker level, and the signal frequency. For a signal frequency and masker level comparable to those used in the present study, and a broad masker spectrum, the transition occurs at about 20 dB sensation level of the masked tone. The similarity of the data indicate that, for tones above about 20 dB sensation level of the masked tone in the diotic condition, the masker has a negligible effect on the loudness of the tone, i.e., its loudness is essentially the same as for a tone in quiet.

4.3 Modelling partial loudness of a tone in noise

The high correlation between magnitude of the tonal content and partial loudness of a tone indicates that predictions of loudness models may be used to estimate the magnitude of tonal content. To study the behaviour of a loudness model in a masking-release condition of the present study, the dynamic loudness model (DLM) of Chalupper and Fastl (2002) was used to simulate the results of the first experiment. The model

consists of (i) a 50-Hz high-pass filter, to simulate the limit in the audible range, (ii) a bank of overlapping auditory filters (channels) to simulate the frequency-place transformation at the level of the cochlea, (iii) a correction factor to account for the transmission through outer and middle ear, (iv) a nonlinear transformation of intensity to instantaneous loudness, (v) a stage to account for postmasking, (vi) a summation across channels, and, finally, an 8-Hz low-pass filter, to simulate temporal integration. Masked thresholds were simulated using the model as an artificial observer in the same procedure used for the experiment. For each interval the maximum of the short-term loudness of the stimulus in the critical band centred at the frequency of the tone, i.e., the specific loudness in this critical band, was calculated. In a second step, the specific loudness in this critical band was calculated for an independent realisation of the noise and, in the last step, subtracted from the specific loudness calculated in the first step. For the interval containing the tone, this procedure results in an estimate of the partial loudness of the tone. In the other intervals, the difference will be close to zero, since the loudness of different realisations of noise is similar. In each trial the model chose the interval with the highest partial loudness, i.e., the tone was considered to be detected if the partial loudness of the signal interval was higher than the maximum of the partial loudnesses of the other intervals. The same adaptive procedure was used as in the experiment, i.e., the mean of the last six reversals was taken as an estimate of the masked threshold. The predicted masked threshold was calculated as the mean of ten of those estimates.

As for the masked thresholds, the same procedure as in the psychoacoustical experiment was used to estimate the level at equal loudness of the tone. Partial loudness was calculated for each interval of the trial using the same algorithm as used for the prediction of the masked threshold and the level of the tone in the next trial was chosen by means of the same adaptive procedure as in the experiment. Again the final estimate of the level at equal partial loudness was averaged across the results of ten runs.

Figure 5 shows an example of the signals (top panels) and their internal representations (bottom panels) for the noise only (grey curves) and for the noise with an added tone at signal-to-noise ratio of −10 dB (black curves) for the baseline condition (left panels) and the masking-release condition (right panels). Circles indicate the maximum of the short-term loudness in the critical band centred at the tone which was taken as an estimate of the specific loudness in this band. Specific loudness for the combination of tone and noise was higher than that for the noise only. The difference is very similar for the two noise types, thus the model does not predict a higher audibility of the tone in modulated noise. This is primarily due to the temporal resolution of the model, which is limited by the cut-off frequency (8 Hz) of its final low-pass filter stage. This limited temporal resolution prevents th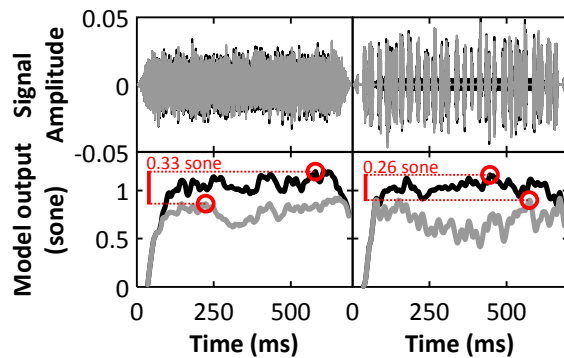e model from using the higher signal-to-noise ratio at temporal positions where the noise amplitude is low. In addition, the model does not include an across-channel process which enhances the signal representation when the noise masker is comodulated.

Figure 5: (colour online): Stimuli (top panels) and short-term loudness (bottom panels) for the two masking conditions of the first experiment: Baseline condition with unmodulated noise (left panels) and masking-release condition with modulated noise (right panels). Grey and black lines indicate the time waveforms and short term loudnesses of the noise only and of the combination of signal and noise, respectively. Circles indicate the maximum of the short-term loudness which is taken as an estimate of the overall loudness of the stimulus.

Figure 6 shows the predictions of the model for the partial loudness of the tone (diamonds) together with the measured data (squares) using the same data representation as in Figure 3. As expected from the internal representations shown in Figure 5, predicted levels at equal partial loudness of the tone are close to the diagonal. The results of the predicted masked threshold are about the same for the two masking conditions: 47 dB SPL for the masking-release condition and 45 dB SPL for the baseline condition. As a result of the inability of the model to predict a masking release, the correlation between measured data and predictions of the levels at equal loudness of the tone is only 0.56.

A mechanism eliciting the masking release in the modulated condition was simulated by reducing the noise level of the modulated noise by 22 dB to 43 dB SPL. Simultaneously, a detection criterion of $N_c = 0.05$ sone was introduced: The interval containing the tone was detected if its partial loudness was higher than the maximum of the partial loudnesses of the other intervals by this critical amount N_c. These settings were found to yield the best correspondence between predicted and average experimental masked threshold. The data for the simulation with this altered model approach are shown with black triangles in Figure 6. The model now predicts the effect of masking on suprathreshold perception, i.e., a gradual decrease of the effect of masking release towards higher levels of the tone. The correlation between data and predictions for the levels at equal loudness of the tone is 0.98. Thus, main aspects of the data can be predicted if the reduced masking in the masking-release condition is taken into account by reducing the masker level
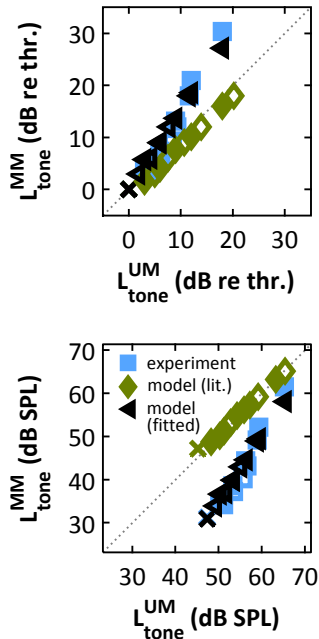
Figure 6: (colour online) Predicted partial loudness (diamonds and triangles) and measured data (squares) for experiment 1. The top panel shows predictions together with the measured data in dB relative to the masked threshold while the bottom panel shows predictions and measured data in dB SPL. The crosses indicate masked thresholds. Open and closed symbols are used as in Figure 2 to indicate which of the tones was varied. The predictions of the model for the original stimuli are shown in diamonds. The black triangles indicate predictions where a release of masking is simulated by reducing the level of the modulated masker.

accordingly. Since the loudness model was not built to account for binaural presentation, it will not predict the data of the second experiment, unless the masking release is introduced artificially as in the simulations of the first experiment. Such simulations were not performed for the second experiment, but it is likely that the correlation between predictions and data is equally high, since the data for the two experiments are very similar.

Similar predictions are expected, for instance, from the model for time-varying loudness (Glasberg and Moore, 2002), since the differences between the models are only visible in a very few experiments (Rennies *et al.*, 2010). The present study focused on dynamic loudness models for the simulations since stationary models are not sensitive to amplitude modulations. The model by Chalupper and Fastl was used since it is computational efficient and is implemented in the time domain, i.e., it uses a filterbank to simulate frequency selectivity of the auditory system.

Chalupper and Fastl (2002) did not define how the partial loudness of a tone in noise should be calculated. The current algorithm is based on the simple assumption that the partial loudness is the difference between the loudnesses of the combination of tone and noise and of noise only. The model of partial loudness by Moore *et al.* (1997) uses a more elaborate algorithm. This

may result in a slightly different partial loudness, especially at low levels of the tone. However it is unlikely that the model predicts a release of masking, since the model works on the long-term spectrum of the stimulus which is the same for the masking-release condition and the baseline condition.

4.4 Implications for the standard DIN 45681

The German standard DIN 45681 (2005) assumes that the threshold of a tonal component is determined by the noise level in the critical band centred at the frequency of the tone. This threshold estimate is based on the assumption of the power spectrum model and thus cannot account for lower thresholds due to (co-)modulation of the masker or binaural cues. Since masking release also affects suprathreshold perception the standard cannot be used in its present form when such cues are present, at least not at levels close to masked threshold. For levels that correspond to sensation levels of the tone in the baseline condition larger than about 20 dB, the present data indicate that the level at equal magnitude of tonal content is likely to be the same for the baseline condition and the masking-release conditions and thus the standard is applicable in its present form. At low levels, up to about about 9 dB above masked threshold, the same magnitude of tonal content is obtained at about the same sensation level. Thus, for these levels the standard can account for the data when a_V is changed to the reduced threshold in the masking-release condition. For intermediate sensation levels a more elaborate measure is needed to account for the data. The similarity of the data for partial loudness of the tone and the results for the magnitude of tonal content indicate that loudness models may be used to estimate the magnitude of tonal content. The simulations of the present study indicate that loudness models may be used to estimate the partial loudness of the tone in a masking-release condition if the masking release is simulated by reducing the masker level in those conditions. A simulation of the partial loudness in the masking-release conditions with the original stimuli is not yet possible, because current loudness models do not include mechanisms to predict a masking release due to binaural cues or due to (co-)modulation of the noise background. Thus for complex stimuli with multiple cues a measurement of the magnitude of tonal content may still be necessary.

5 Summary and Conclusions

The perception of audible tonal components in a noisy background is an important question for the assessment of the perception of technical sounds, and several standards have been developed to address this point. The data of the present study indicate that the perceived magnitude of these tonal components is closely related to their partial loudness. On the one hand this finding has implications for future experiments on the

perception of tonal components, since the magnitude of tonal content may be assessed by asking the listeners to rate the partial loudness of the tonal component. On the other hand it has implications for algorithms that calculate the magnitude of tonal content. Algorithms based on loudness values may provide a better correspondence to the perception of tonal components in noise than level-based algorithms which are currently used in the standards. Finally, the present data on the partial loudness of a tone embedded in noise under conditions of masking release may help to develop loudness models that include stages which are sensitive to interaural disparities and common envelope fluctuations in different frequency regions. It has been shown that current loudness models can only account for the data if the release from masking is simulated by reducing the level of the masker.

Acknowledgments

References

ANSI S1.13 (**2005**). "Measurement of Sound Pressure Levels in Air", ANSI S1.13-2005.

Burgess, M. and McCarthy, M. (**2009**). "Effectiveness of non-tonal audible movement warning alarms for construction sites", in *Proceedings of Acoustics 2009*, 1–4 (Adelaide, Australia).

Buus, S., Florentine, M., and Poulsen, T. (**1997**). "Temporal integration of loudness, loudness discrimination, and the form of the loudness function", J. Acoust. Soc. Am. **101**, 669–680.

Carlyon, R. P., Buus, S., and Florentine, M. (**1989**). "Comodulation masking release for three types of modulator as a function of modulation rate", Hear. Res. **42**, 37–45.

Chalupper, J. and Fastl, H. (**2002**). "Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners", Acta Acustica united with Acustica **88**, 378–386.

DIN 45681 (**2005**). "Acoustics – Detection of tonal components of noise and determination of tone adjustment for the assessment of noise immissions", DIN 45681.

EPA (**2009**). "Wind farms environmental noise guidelines", URL `http://www.epa.sa.gov.au/xstd_files/Noise/Guideline/windfarms.pdf`, Environmental Protection Agency.

Ernst, S. M. A., Uppenkamp, S., and Verhey, J. L. (**2010**). "Cortical representation of release from auditory masking", NeuroImage **49**, 835–842.

Fastl, H. and Zwicker, E. (**2007**). *Psychoacoustics: facts and models*, Springer series in information sciences (Springer).

Glasberg, B. R. and Moore, B. C. J. (**2002**). "A Model of Loudness Applicable to Time-Varying Sounds", J. Audio Eng. Soc **50**, 331–342.

Gleiss, N. and Zwicker, E. (**1963**). "Loudness Function in the Presence of Masking noise", J. Acoust. Soc. Am. **36**, 393–394.

Hall, J. W., Grose, J. H., and Hatch, D. R. (**1996**). "Effects of masker gating for signal detection in unmodulated and modulated bandlimited noise", J. Acoust. Soc. Am. **100**, 2365–2372.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectro-temporal pattern analysis", J. Acoust. Soc. Am. **76**, 50–56.

Hall, J. W., Tyler, R. S., and Fernandes, M. A. (**1983**). "Monaural and binaural auditory frequency resolution measured using bandlimited noise and notched-noise masking", J. Acoust. Soc. Am. **73**, 894–898.

Hansen, H., Verhey, J. L., and Weber, R. (**2011**). "The Magnitude of Tonal Content. A Review", Acta Acustica united with Acustica **97**, 355–363.

Hansen, H. and Weber, R. (**2011**). "Partial loudness as a measure of the magnitude of tonal content", Acoust. Sci. & Tech. **32**, 111–114.

Hellman, R. P. and Zwislocki, J. (**1964**). "Loudness functions of a 1000-cps tone in the presence of a masking noise", J. Acoust. Soc. Am. **36**, 1618–1627.

IEC 64100-11 (**1998**). "Wind turbine generator systems – Part 11: Acoustic noise measurement techniques.", International Electrotechnical Commission.

ISO 1996-2 (**2007**). "Acoustics – Description, measurement and assessment of environmental noise – Part 2: Determination of environmental noise level.", International organization of Standardization.

Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Wood, C. L. (**1956**). "Masking of Tonal Signals", J. Acoust. Soc. Am. **28**, 416–426.

Levitt, H. (**1971**). "Transformed Up-Down Methods in Psychoacoustics", J. Acoust. Soc. Am. **49**, 467–477.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness", J. Audio Eng. Soc **45**, 224–240.

Nitschmann, M., Verhey, J. L., and Kollmeier, B. (**2009**). "The role of across-frequency processes in dichotic listening conditions", J. Acoust. Soc. Am. **126**, 3188–3198.

Rennies, J., Verhey, J. L., and Fastl, H. (**2010**). "Comparison of loudness models for time-varying sounds", Acta Acustica united with Acustica **96**, 383–396.

Soderquist, D. R. and Shilling, R. D. (**1990**). "Loudness and the Binaural Masking Level Difference", Bull. Psychonomic. Soc. **28**, 553–555.

Steinberg, J. C. and Gardner, M. B. (**1937**). "The Dependence of Hearing Impairment on Sound Intensity", J. Acoust. Soc. Am. **9**, 11–23.

Stevens, S. and Guirao, M. (**1967**). "Loudness Functions under Inhibition", Perception Psychophys. **2**, 459–465.

Townsend, T. H. and Goldstein, D. P. (**1972**). "Suprathreshold Binaural Unmasking", J. Acoust. Soc. Am. **51**, 621–624.

van de Par, S. and Kohlrausch, A. (**1999**). "Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters", J. Acoust. Soc. Am. **106**, 1940–1947.

Verhey, J. L., Dau, T., and Kollmeier, B. (**1999**). "Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model", J. Acoust. Soc. Am. **106**, 2733–2745.

Verhey, J. L. and Ernst, S. M. A. (**2009**). "Comodulation masking release for regular and irregular modulators", Hear. Res. **253**, 97–106.

Verhey, J. L. and Heise, S. (**2010**). "Einfluss der Zeitstruktur des Hintergrundes auf die Tonhaltigkeit und Lautheit des tonalen Vordergrundes [Influence of the temporal structure of the background on the magnitude of tonal content of the tonal foreground]", in *Fortschritte der Akustik – DAGA 2010*, 595–596 (DEGA e.V., Berlin).

Verhey, J. L., Pressnitzer, D., and Winter, I. M. (**2003**). "The psychophysics and physiology of comodulation masking release", Exp. Brain Res. **153**, 405–417.

Verhey, J. L. and Stefanowicz, S. (**2011**). "Binaurale Tonhaltigkeit [Binaural magnitude of tonal content]", in *Fortschritte der Akustik – DAGA 2011* (DEGA e.V., Berlin).

Verhey, J. L., Stefanowicz, S., and Heise, S. J. (**2011**). "Magnitude of tonal content in conditions of masking release", in *British Society of Audiology Conference 2011*.

Vormann, M. (**2011**). "Untersuchungen zu psychoakustischen Mess- und Berechnungsverfahren der Tonhaltigkeit [Investigations into psychoacoustic methods for measuring and calculating the magnitude of tonal content]", Ph.D. thesis, University of Oldenburg.

Vormann, M., Meis, M., Mellert, V., and Schick, A. (**1999**a). *Psychophysics, Physiology and Models of Hearing*, chapter A New Approach for the Evaluation of Tonal Noise (Tonality), 109–112 (World Scientific, Singapore).

Vormann, M., Verhey, J. L., Meis, M., Klatte, M., and Mellert, V. (**1998**). "Tonhaltigkeit aus psychophysikalischer Sichtweise: Untersuchung zur Versuchspersoneninstruktion bei der Bewertung der Tonhaltigkeit [Magnitude of tonal content from a psychoacoustical perspective: Experiments on the subjects' instructions when assessing the magnitude of tonal content]", in *Fortschritte der Akusitk – DAGA 1998* (DEGA e.V., Berlin).

Vormann, M., Verhey, J. L., Mellert, V., and Schick, A. (**1999**b). "Factors influencing the subjective rating of tonal components in noise", in *Joint Meeting "Berlin 99" integrated 25th German Acoustics DAGA Conference, March 1999*, 4 (DEGA e.V.).

Vormann, M., Verhey, J. L., Mellert, V., and Schick, A. (**2000**). "Ein adaptives Verfahren zur Bestimmung der subjektiven Tonhaltigkeit [An adaptive procedure for determining the subjective magnitude of tonal content]", in *Fortschritte der Akustik – DAGA 2000*, 304–305 (DEGA e.V., Oldenburg).

Zwicker, E. (**1963**). "Über die Lautheit von gedrosselten und ungedrosselten Schallen", Acustica **13**, 194–211.