



## Research Paper

# Spatial and temporal disparity in signals and maskers affects signal detection in non-human primates



Francesca Rocchi, Margit E. Dylla, Peter A. Bohlen, Ramnarayan Ramachandran\*

Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN 37212, USA

## ARTICLE INFO

## Article history:

Received 22 February 2016  
 Received in revised form  
 10 October 2016  
 Accepted 15 October 2016  
 Available online 19 October 2016

## Keywords:

Detection threshold  
 Auditory scene analysis  
 Temporal asynchrony  
 Spatial release from masking  
 Auditory masking

## ABSTRACT

Detection thresholds for auditory stimuli (signals) increase in the presence of maskers. Natural environments contain maskers/distractors that can have a wide range of spatiotemporal properties relative to the signal. While these parameters have been well explored psychophysically in humans, they have not been well explored in animal models, and their neuronal underpinnings are not well understood. As a precursor to the neuronal measurements, we report the effects of systematically varying the spatial and temporal relationship between signals and noise in macaque monkeys (*Macaca mulatta* and *Macaca radiata*). Macaques detected tones masked by noise in a Go/No-Go task in which the spatiotemporal relationships between the tone and noise were systematically varied. Masked thresholds were higher when the masker was continuous or gated on and off simultaneously with the signal, and lower when the continuous masker was turned off during the signal. A burst of noise caused higher masked thresholds if it completely temporally overlapped with the signal, whereas partial overlap resulted in lower thresholds. Noise durations needed to be at least 100 ms before significant masking could be observed. Thresholds for short duration tones were significantly higher when the onsets of signal and masker coincided compared to when the signal was presented during the steady state portion of the noise (overshoot). When signal and masker were separated in space, masked signal detection thresholds decreased relative to when the masker and signal were co-located (spatial release from masking). Masking release was larger for azimuthal separations than for elevation separations. These results in macaques are similar to those observed in humans, suggesting that the specific spatiotemporal relationship between signal and masker determine threshold in natural environments for macaques in a manner similar to humans. These results form the basis for future investigations of neuronal correlates and mechanisms of masking.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Background sounds often result in an increase in threshold levels for detecting signals. In the real world, maskers may occur in complex spatiotemporal configurations relative to the signal to be detected. The auditory system relies on temporal, spectral and spatial cues to detect signals in noise. Low-level features, such as the signal onset and the spatial location of sound energy, represent important cues for signal detection and formation of auditory scenes (Bregman, 1990). Here we report the results of studies

investigating how the spatio-temporal properties of background noise affect masking.

Behavioral and neuronal responses to a target sound are sensitive to other sounds that occur around the time of occurrence of the target sound. Not surprisingly, masked detection is sensitive to the temporal structure of the target signal and the masker. Behaviorally, these have been best exemplified by studies of simultaneous masking, and studies of forward and backward masking (non-simultaneous masking), which involve manipulations of stimulus onset asynchrony of maskers and signals (e.g. de Maré, 1940; Lüscher and Zwislocki, 1947; Munson and Gardner, 1950; Zwislocki et al., 1959; Plomp, 1964; Elliott, 1971; Widin and Viemeister, 1979; Jesteadt et al., 1982; Moore and Glasberg, 1983; Nelson, 1991; Plack and Oxenham, 1998). The behavioral consequences of such temporal relationships have not been well studied

\* Corresponding author. Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, 111 21st Ave S, WH 065, Nashville, TN 37212, USA  
 E-mail address: [ramnarayan.ramachandran@vanderbilt.edu](mailto:ramnarayan.ramachandran@vanderbilt.edu) (R. Ramachandran).

in animal models that will directly allow exploration of neuronal mechanisms. Inhibition/suppression, neuronal adaptation, inhibition or forward suppression, and offset inhibition are all mechanisms that have been proposed to contribute to simultaneous and non-simultaneous masking (e.g. Duifhuis, 1973; Nelson and Swain, 1996; Oxenham and Moore, 1995, 1997; Harris and Dallos, 1979; Brosch and Schreiner, 1997; Gai, 2016).

It is well known that the relative locations of signals and maskers/distractors strongly influence audibility of the target signal (e.g., Santon, 1987). Separation of target signal and masker/distractor in azimuth and/or elevation can cause detection thresholds to be lower by 10–12 dB (spatial release from masking, Saberi et al., 1991). While much of this work has been done in humans, to the best of our knowledge there are no studies in nonhuman animal models that address spatial effects of maskers and distractors on masked detection (however see Sollini et al., 2016). Most of the work done in humans using free field sounds has mainly focused on speech intelligibility (e.g. Plomp, 1976; Plomp and Mimpen, 1981). Only a few studies (e.g. Santon, 1987; Gatehouse, 1986, 1987; Kidd et al., 1998; Saberi et al., 1991) of masking using tones and noise have been conducted in a free field. This study extends the spatial release in masking literature to animal models; here, the benefit of spatial separation of signal and noise on detection threshold (audibility) is directly addressed.

In this study, we explored the effect of varying spatial and temporal relationships between a signal and a masker in free field on masked signal detection. Although spatial and temporal release from masking have been extensively studied in humans, they have not been well explored in animal models that permit direct exploration of the neuronal mechanisms underlying these phenomena, especially in macaques. The use of non-human primates as behavioral model in hearing research has increased over the last decade (e.g., O'Connor et al., 2011; Osmanski et al., 2013; Dylla et al., 2013; Christison-Lagay and Cohen, 2014). The similarity in behavioral performance across human and nonhuman primate species observed in many of the above studies suggests that nonhuman primates may be a good model for human hearing; phylogenetic similarity suggests that the neuronal mechanisms underlying behavioral performance may be similar across these two species. While macaques are being increasingly used in behavioral and physiological studies of the auditory system, there are very few studies on masking. We used macaques to obtain behavioral results for factors influencing masked detection; these provide a baseline for planned investigations of neuronal mechanisms underlying hearing in more naturalistic environments, in which signals and maskers may have spatial and temporal disparities. These also form a baseline for studies of the consequences of hearing impairment.

## 2. Methods

### 2.1. Subjects

Five male macaque monkeys (four *Macaca mulatta* and one *Macaca radiata*) were used as subjects. Monkeys A, B, C and D (*Macaca mulatta*) were 10, 9, 7 and 6 years old, respectively at the start of the study, while Monkey G (*Macaca radiata*) was 7 years old. All monkeys showed audiometrically normal hearing thresholds measured (in dichotic conditions) over the range of frequencies spanning 0.125–40 kHz (Dylla et al., 2013; Bohlen et al., 2014). All the procedures were approved by the Animal Care and Use Committee of the Vanderbilt University Medical Center and were strictly consistent with the guidelines for animal research established by the National Institutes of Health.

### 2.2. Surgical procedure

Monkeys were prepared for chronic experiments using standard techniques employed in previous studies (e.g. Ramachandran and Lisberger, 2005; Dylla et al., 2013). The surgical procedure was performed prior to the behavioral tasks. A metal head holder (Crist Instruments, Hagerstown, MD) was placed on the skull of the monkeys to maintain the position of their head at a constant location with respect to the loudspeakers. Bone cement and eight-mm-long stainless steel screws (Synthes Inc., and Veterinary Orthopedic Instruments) were used to secure the head post to the skull. Analgesics, and if necessary, antibiotics were administered to the monkeys under veterinary oversight. Further details about the surgical procedures are given in Dylla et al. (2013).

### 2.3. Apparatus and stimuli

Experiments were conducted in a sound treated booth (model 1200, Industrial Acoustics Corp., NY or Acoustic Systems) where the monkeys were seated in an acrylic primate chair (audio chair, Crist Instrument Co., Hagerstown, MD). The holder implanted on the monkeys' skull was fastened to the chair. The distance between the loudspeakers and the head of the monkeys varied depending on the task used (for details, see *behavioral task* section). Sounds between 0.05 and 40 kHz (SA1 loudspeaker, Madisound, WI) could be delivered by the loudspeakers, whose on-axis output varied less by than 3 dB between 100 Hz and 40 kHz. A microphone placed at the location of one of the monkey's head was used for calibration. Loudspeakers were further calibrated at different spatial locations to ensure that the sound level at the ear canal was the same irrespective of spatial location.

A computer running OpenEx software (System 3, TDT Inc., Alachua, FL) controlled the experiments. The sampling rate used to generate the signals (tones and noise) was 97.6 kHz. The lever state was sampled at a rate of 24.4 kHz, leading to a temporal resolution of about 40  $\mu$ s on the lever release. A full description of the apparatus used to generate the signals and masker is provided in Dylla et al. (2013).

### 2.4. Behavioral task and procedure

Monkeys were trained to release a lever to report detection of a tone in noise with positive reinforcement. Detection performance was measured only when the monkeys consistently performed the task. The behavioral procedures are described in Dylla et al. (2013) and Bohlen et al. (2014). Briefly, all trials started with the monkeys pressing down on the lever (Model 829 Single Axis Hall Effect Joystick, P3America, San Diego, CA). The signal (tone) was presented after a variable delay (400–1400 ms) following the lever press. A fluid reward was given each time the tone was correctly detected (hit) by releasing the lever within 600 ms of the tone onset. No penalty was given when a tone was presented but the monkey did not release the lever (miss). 80% of trials included a tone, and the remaining 20% no tone was played (catch trials). When the lever was released but a tone was not played (false alarm), a time-out penalty was applied (tones were not presented for a variable time between 6 and 10 s). The step size between tone levels was either 2.5 or 5 dB. Tone levels were chosen from a range spanning 90 dB. The method of constant stimuli was applied with randomized presentation order of each level, and each level was used a minimum of 15 and a maximum of 30 times. A specific masker was chosen for each block. The spatial and temporal properties of noise relative to the tone are specific to the experiment and are described in the appropriate

location.

#### 2.4.1. Temporal experiments

In the first of the *Temporal experiments*, tone detection thresholds were measured for three temporal patterns of noise: the noise was continuous (Steady State noise); the noise was gated on and off with the tone (Simultaneously Gated noise); noise was gated on when the tone was gated off, and off when the tone was gated on (Inversely Gated noise) (Fig. 1A). The duration of the tone and the bursts of noise (Simultaneously Gated noise condition) were typically 200 ms with rise/fall times (ramps) of 10 ms. The broadband noise was a flat-spectrum broadband noise (10 Hz – 40 kHz) generated by the “Random” function in OpenEx (for further details see Dylla et al., 2013). The noise levels used are specified in the results. The tone frequencies tested ranged from 0.2 to 20 kHz.

In the second experiment (Fig. 2), the onset asynchrony of the tone and the noise was varied over a 350 ms range, from –300 ms (noise onset 300 ms before the tone onset) to 50 ms (noise onset 50 ms after the tone onset). A positive onset asynchrony (12.5, 25, 50 ms) indicates that the noise was played after the tone. A negative asynchrony describes the conditions where the masker was presented before the signal (–300, –200, –100, –50). When the onset asynchrony was zero, the onsets of the noise and the tone coincided. Fig. 2A shows examples of positive onset asynchrony (when the tone onset was 100 ms before the noise onset) and negative onset asynchrony (tone onset 100 ms after the noise onset). Detection thresholds were estimated for five tone frequencies (0.5, 1, 2, 4 and 8 kHz). The duration of the tone and the noise were 200 ms. A further control condition (Fig. 2C) was used to investigate whether complete overlap of signal and masker after an onset asynchrony still resulted in full masking. We tested an onset asynchrony of –100 ms, tone and noise durations of 200 and 300 ms, respectively for this control condition.

In the third experiment, tone (200 ms duration) detection thresholds were measured in noise of varying duration (Fig. 3). This condition extended the idea of overlap between signal and noise by varying properties of the noise, as well as the overlap between the tone and the noise. Here the asynchrony between signal masker is given by either the temporal gap between the offset of the noise and the onset of the signal (negative onset asynchrony) or the gap between the offset of the tone and the onset of the signal (positive asynchrony). A temporal gap of zero (Fig. 3A) indicates that noise and signal have the same onset, whereas a temporal gap of +50 ms refers to the condition where the onset of the noise is 50 ms after the offset of the tone. The temporal gap was kept constant across different noise durations and spanned between –25 ms (noise offset was 25 ms before signal onset) and 50 ms (the onset of the noise was 50 after the offset of the tone). A 4 kHz tone was used as signal in this experiment. For noise durations less than 50 ms, the rise/fall times (ramps) used were 4 ms or 2 ms, depending on the duration of the masker.

In a fourth experiment, we investigated the overshoot effect. The duration of the masker was kept constant at 200 ms, while the onset asynchrony of short duration (6.5 ms or 12.5 ms) tones (4 or 16 kHz) was varied relative to the noise onset (Fig. 4A). The noise spectrum level was either 10 or 30 dB SPL spectrum level.

#### 2.4.2. Spatial experiments

In the *Spatial experiments* tones were played in the presence of continuous noise that originated at locations different from that of the tone. At the beginning of each block, broadband noise at 10 dB SPL spectrum level was presented for about 10 s before the first

tones were played. The location of the tone was fixed throughout the experiment (azimuth and elevation of 0°), while the noise location was varied. The noise and the tone loudspeakers were located at the same distance from the center of the monkey's head (55.9 cm). The angular separation between the signal and the noise was varied in either azimuth or elevation. In the azimuthal separation condition, the signal/noise separation was 0, 15, 30, 45, 60 or 90° (Fig. 5A) and tone frequencies of 2, 12, 16 or 24 kHz were tested. In the elevation separation condition, tone frequencies of 2, 12 and 24 kHz were tested for signal/noise separations of 0, 24.5, 36.5, 49 or 56° (Fig. 6A). Vertical signal/masker separations could not exceed 56° because we were limited by the sound booth height.

#### 2.5. Data analysis

Data were analyzed using signal detection theoretic methods (Green and Swets, 1966; MacMillan and Creelman, 2005). Behavioral accuracy was computed based on the hit rate (H) and false alarm rate (F) at each tone level using the following sequence: first, we calculated sensitivity using signal detection theory as  $d'(level) = z(H(level)) - z(F)$ , where  $z$  converts hit rate and false alarm rate into normalized units of standard deviation ( $z$ -score, norminv in MATLAB). Because we wanted these results to serve as a baseline for neurophysiological studies where we would measure (signal) and (signal + noise) representation distributions, we also calculated the probability correct (pc) at each tone level as follows:

$$pc(level) = z^{-1} \left( \frac{d'(level)}{2} \right),$$

where the inverse  $z$ -transform ( $z^{-1}$ ) converts a unique number of standard deviations of a standard normal distribution into a probability correct (normcdf in MATLAB). The conversion of  $d'$  to the  $pc$  measure was to facilitate the comparison of psychometric functions with neurometric functions obtained from neuronal responses using distribution free methods. The traditional threshold estimated at  $d' = 1$  now corresponds to  $pc(level) = 0.76$ .

All psychometric functions were fitted with an adapted Weibull cumulative distribution function (cdf), according to the following equation:

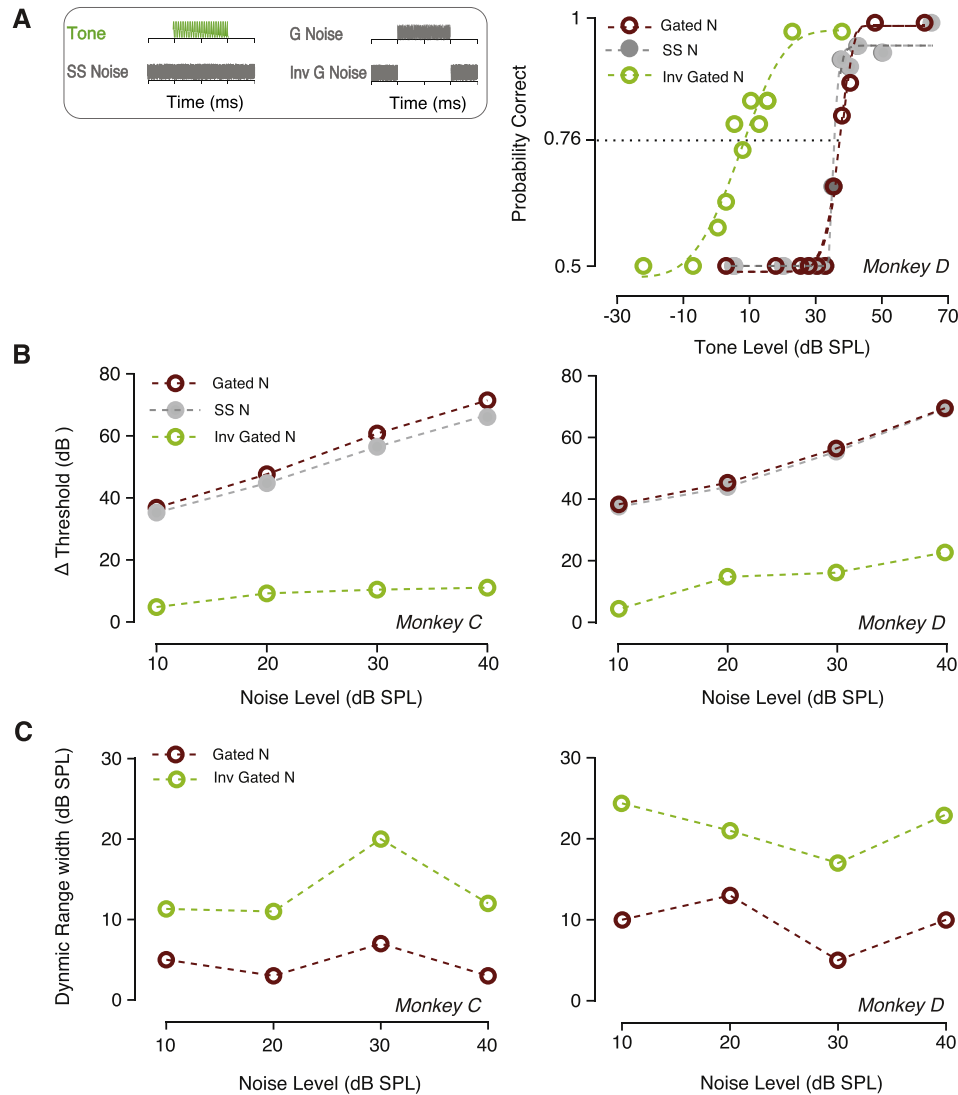
$$pc(level)_{fit} = c - d * e^{-\left(\frac{level}{\lambda}\right)^k},$$

where  $level$  indicates the tone sound pressure level (dB SPL),  $\lambda$  represents a threshold parameter.  $k$  is the slope parameter,  $c$  represents the probability correct at high sound levels, and  $d$  is the estimate of chance performance at levels well below thresholds. Threshold was estimated as the level that corresponded to  $d' = 1$ , or, equivalently,  $pc = 0.76$ .

### 3. Results

#### 3.1. Temporal experiment

Fig. 1A (left) shows the stimulus schematics for continuous, simultaneously gated, and inversely gated noise. Fig. 1A (right) shows exemplar psychometric functions ( $pc$  vs. tone level) for monkey D, for a 4 kHz tone and 10 dB SPL spectrum level noise for the three noise conditions. The circles of different colors represent different gating conditions of the noise. The tone was 200 ms in duration. The psychometric function for inversely gated noise (green circles and line) was shifted to lower levels relative to those for the simultaneously gated (red line) and continuous noise (gray line).



**Fig. 1.** Tone detectability is a function of the temporal properties characterizing noise masking. (A) Schematic illustration of three experimental conditions tested. Noise background (10–40 dB SPL) could be either continuous (Steady State noise), simultaneously gated with the tone (Simultaneously Gated noise), or gated off with onset and gated on with tone offset (Inversely Gated noise). Tones were 200 ms long, and in simultaneously gated noise conditions, noise bursts were 200 ms long. Examples of psychometric functions from monkey D show behavioral accuracy for continuous (grey line), simultaneously gated (red line), and inversely gated noise (green line) when a 4 kHz tone was presented. (B). Masked threshold shift ( $\Delta$  threshold) relative to tone threshold is plotted as a function of noise spectrum level. (C). Dynamic range of psychometric function for monkeys C and D in the simultaneously and inversely gated conditions (green and red symbols respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

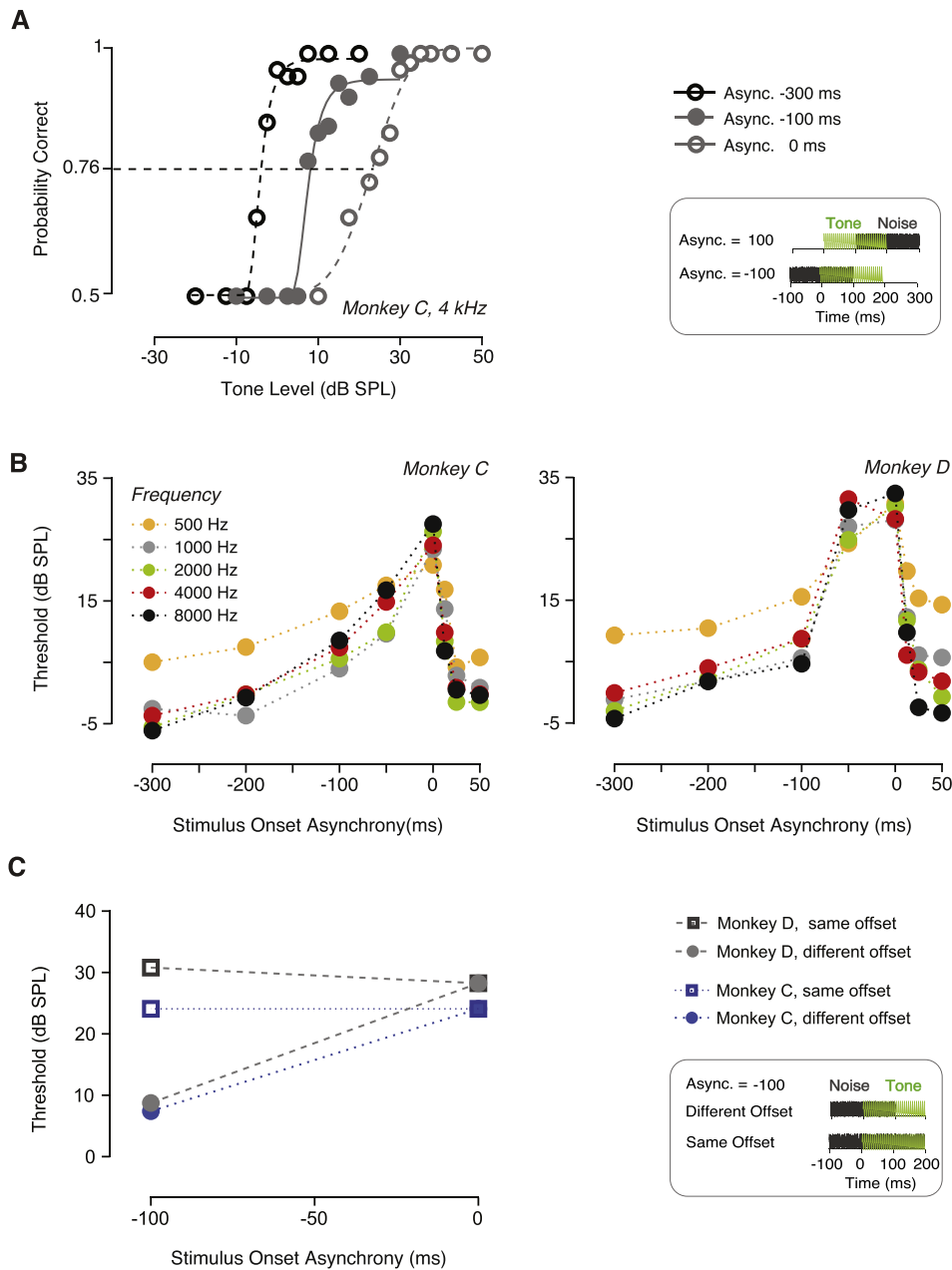
Threshold shifts ( $\Delta$  threshold, calculated as the difference between masked detection thresholds and tone alone thresholds) are shown in Fig. 1B for monkeys C (left) and D (right). These results indicate that masked detection thresholds were strongly influenced by the sound pressure level as well as the gating of the noise relative to the tone. Detection thresholds were not significantly different between continuous and simultaneously gated noise conditions. A two-way ANOVA confirmed that there was no significant difference in detection thresholds depending on whether the noise background was continuous or simultaneously gated with the tone (monkey C:  $F(1, 3) = 8.434$ ,  $P = 0.0623$ ; monkey D:  $F(1, 3) = 3.73$ ,  $P = 0.1491$ ). However, thresholds for inversely gated noise were significantly lower than those measured in both simultaneously gated (monkey C:  $F(1, 3) = 51.64$ ,  $P = 0.0056$ ; monkey D:  $F(1, 3) = 100.2$ ,  $P = 0.0021$ ) and continuous (monkey C:  $F(1, 3) = 76.48$ ,  $P = 0.0031$ ; monkey D:  $F(1, 3) = 84.04$ ,  $P = 0.0027$ ) noise. The dynamic range was calculated to estimate the steepness

of the psychometric function: it was calculated as the difference in the tone levels that evoked  $pc = 0.90$  and  $pc = 0.60$ . The dynamic range was larger for inversely gated noise compared to simultaneously gated noise (Fig. 1C). This difference was significant (two-way ANOVA, monkey D:  $F(1, 3) = 32.71$ ,  $P = 0.0106$ ; monkey C:  $F(1, 3) = 21.15$ ,  $P = 0.019$ ). The dynamic ranges were not significantly different between the steady noise and the simultaneously gated noise.

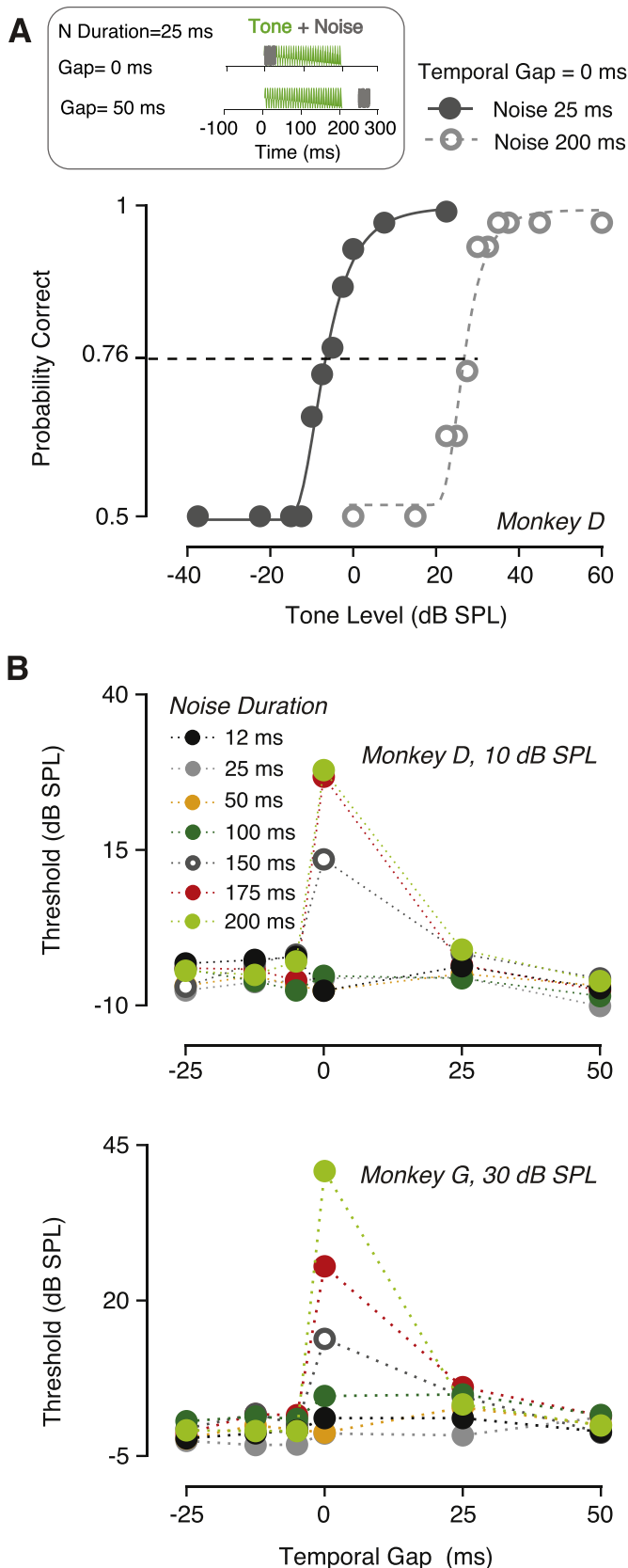
Numerous studies have shown that detection thresholds for human listeners are influenced by the temporal relationship between the tone and the noise. Fig. 2A shows the stimulus schematic, and psychometric functions (monkey C) for onset asynchronies (signal re: noise) of 0, –100 and –300 ms for a 4 kHz tone. The psychometric functions were shifted towards highest levels for onset asynchrony approached 0 ms, were shifted to intermediate levels for –100 ms onset asynchrony, and were not shifted relative to those for tone alone for –300 ms onset asynchrony. Fig. 2B

summarizes the thresholds for both positive and negative onset asynchronies showing a clear peak at 0 ms onset asynchrony, a sharp decrease in thresholds for positive asynchronies, and a more shallow decrease for negative onset asynchronies. A similar trend was observed for all tone frequencies tested. A two-way ANOVA showed that asynchrony significantly affected detection thresholds for both monkeys (monkey C:  $F(7, 28) = 53.47, P < 0.0001$ ; monkey D:  $F(7, 28) = 50.05, P < 0.0001$ ). A Dunnett's post-hoc test revealed that thresholds significantly decreased even with a short positive

onset asynchrony of 12.5 ms, but decreased more slowly for negative onset asynchronies. A significant effect of frequency was found for both monkeys (monkey C:  $F(4, 28) = 6.221, P = 0.0010$ ; monkey D:  $F(4, 28) = 7.756, P = 0.0002$ ). However, the post-hoc tests showed that only the thresholds obtained with a 500 Hz tone were significantly higher than the others. This is consistent with higher audiometric thresholds for 500 Hz tones for macaques (see [Bohlen et al., 2014](#) for audiometric thresholds of monkeys C and D).



**Fig. 2.** Onset asynchrony of simultaneously gated masker and signal affects tone detection. (A) The relative onset between target (signal) and noise (masker) was varied. Schematic shows onset asynchronies of +100 and -100 ms (tone onset preceding and following the noise onset). Tone is schematized in green, and noise in grey. Durations of both noise masker and tone (4 kHz) were held constant to 200 ms. Three psychometric functions for monkey C showed the shift in dynamic range when signal was presented with onset asynchrony of either -300 ms (black circles), -100 ms (filled circles) or 0 ms (grey circles). (B) Detection thresholds plotted as a function of the onset asynchrony between the tone and the noise masker for monkey C and D (left and right). Behavioral sensibility was measured across several tone frequencies, indicated by different colors (see legend). (C) Thresholds to 4 kHz tones when masker completely overlapped the signal along with an onset asynchrony of -100 ms (filled symbols) are compared with thresholds measured when the masker and signal had similar durations and an onset asynchrony of -100 ms (incomplete overlap, open symbols). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Effect of masker duration on detection thresholds. (A) Psychometric functions for monkey D measured when the temporal gap between masker and signal was 0 ms. A 200 ms 4 kHz tone was masked by a burst of noise, whose duration was either 25 ms (solid line) or 200 ms (dashed lines). The diagram illustrates a schematic representation of two different temporal gaps between masker and signal (noise duration was

It is possible that the threshold changes due to negative or positive onset asynchronies were due to the reduced overlap of signal and masker. This suggested that a noise with a negative onset asynchrony but longer duration such that it has greater overlap of the tone should cause a higher masked threshold. To test this hypothesis we tested 4 kHz tone detection in noise (−100 ms onset asynchrony) for two durations: 200 ms (incomplete overlap between tone and noise) and 300 ms (complete overlap between tone and noise, Fig. 2C). Data for monkeys C and D are shown in grey and blue in Fig. 2C. Thresholds in the 300 ms noise were higher than the thresholds in the 200 ms noise for both monkeys (compare squares and circles). The thresholds with complete overlap of tone and noise matched thresholds in the 0ms onset asynchrony condition (see Fig. 2C).

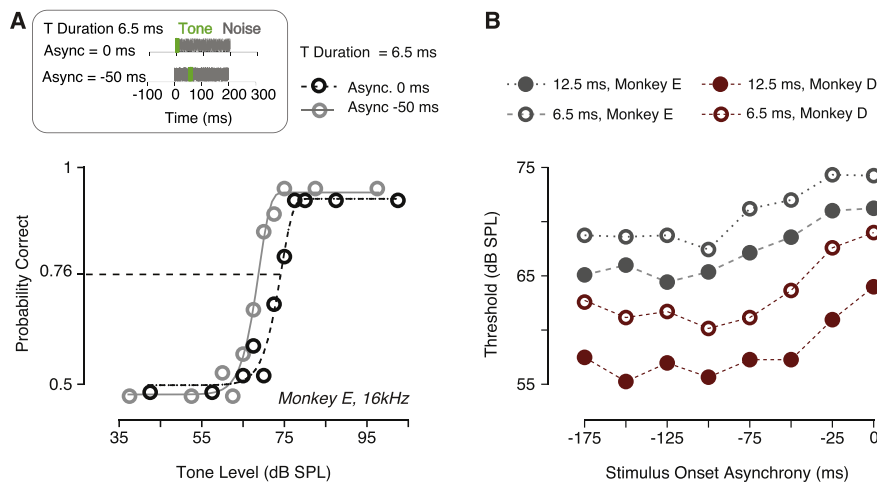
If overlap of signal and noise controlled masked thresholds, then changing the noise duration should have the same effect as onset asynchrony. Fig. 3 shows how masker duration affected detection threshold when the tone duration was held constant (200 ms). A temporal gap was constant across different masker durations. The psychometric functions for monkey D (Fig. 3A) illustrate that thresholds for a temporal gap of 0 ms (i.e. 0 ms onset asynchrony) shifted towards higher tone levels when the duration of the masker was increased from 25 ms (filled symbols) to 200 ms (open), consistent with greater overlap between tone and noise for the 200 ms duration noise.

The noise duration had an effect on masking only when the burst of noise was presented for duration longer than 100 ms, and only for an onset asynchrony of 0 ms (Fig. 3B). This result is consistent with the overlap hypothesis. Similar results were observed when the masker was presented at 25 dB SPL. A two-way ANOVA revealed that noise duration did not significantly lead to a change in thresholds (monkey D:  $F(6, 30) = 1.697, P = 0.1560$ ; monkey G:  $F(6, 30) = 3.697, P = 0.007$ ). However, Dunnett's post-hoc test showed that when the temporal gap was zero, thresholds increased significantly for noise durations greater than 100 ms. Since only relatively long noise durations seemed to significantly affect tone thresholds, a further analysis was performed to compare detection thresholds with a masker duration longer than 100 ms. A two-way ANOVA confirmed that detection thresholds significantly increased with noise durations longer than 100 ms (monkey D:  $F(1, 17) = 5.254, P = 0.0349$ ; monkey G:  $F(6, 30) = 5.407, P = 0.0327$ ). The temporal gap between signal and masker significantly affected threshold (monkey D:  $F(5, 30) = 3.712, P = 0.0098$ ; monkey G:  $F(5, 30) = 4.159, P = 0.0055$ ). Dunnett's post-hoc test showed that thresholds significantly decreased for a temporal gap of −5 ms.

Studies in humans show that thresholds of short duration tones at the onset of a noise are higher than those of tones played during the steady state portion of noise (overshoot effect). Fig. 4 shows detection performance for monkey E when the onset of a 6.5 ms, 16 kHz tone either coincided with the onset of a 200 ms noise (0 ms onset asynchrony) or was delayed of 50 ms (−50 ms onset asynchrony). The behavioral performance in the noise with onset asynchrony of −50 ms was shifted to lower sound pressure levels relative to that in noise with 0 ms onset asynchrony (Fig. 4A). This reduction in thresholds is consistent with the overshoot effect observed in humans. Fig. 4B shows detection thresholds as a function of the onset asynchrony between noise and tone. Detection performance was measured using either a 16 kHz tone masked by 30 dB SPL spectrum level noise (monkey E) or a 4 kHz tone

25 ms). Tone duration was always 200 ms while the duration of the noise varied. (B) Detection thresholds as a function of the temporal gap. The different masker durations are shown by different colors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

masked by 30 dB SPL spectrum level noise. These results suggest that detection thresholds decreased as a function of the delay between the onset of the masker and the onset of the signal. A two-way ANOVA revealed that thresholds decreased as the onset asynchrony between signal and masker increased (monkey D:  $F(7, 7) = 4.464$ ,  $P = 0.0334$ ; monkey E:  $F(7, 7) = 46.58$ ,  $P < 0.0001$ ). Dunnett's post-hoc test showed significantly decreased when the stimulus onset asynchrony was  $\geq$  of 100 ms for monkey D and  $\geq 75$  ms for monkey E.



**Fig. 4.** Overshoot effect for short tones played soon after a noise masker. (A) Schematic shows the stimulus configuration. The duration of the noise masker was held constant at 200 ms, whereas the duration of a 4 kHz or 16 kHz tone was either 6.5 ms (as shown) or 12.5 ms. Two exemplar psychometric functions from monkey E shows the horizontal dynamic range shift as a result of noise with 0 ms onset asynchrony (the onsets of noise and tone overlapped) and a negative asynchrony of 50 ms (the signal onset was delayed by 50 ms after the noise onset). (B) Tone detection thresholds are plotted as a function of the onset asynchrony for monkey D (red) and monkey E (gray). The amplitude of the noise used was either 10 (monkey D) or 30 (monkey E) SPL. Threshold functions are shown for 6.5 ms signals (open symbols) and 12.5 ms signals (filled). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Spatial experiments

Detection thresholds as a function of the horizontal spatial separation between masker and signal are summarized in Fig. 5. The location of the noise was varied along the horizontal plane, while the signal source location was constant at 0 deg azimuth and elevation. Psychometric functions (monkey B) for two angular separations (0° and 60°) show the threshold shift observed when the spatial separation between the source locations of noise and tone (16 kHz) increased (Fig. 5A). Detection thresholds for monkey A and monkey B (Fig. 5B) are plotted as a function of the angular separation between the tone and the noise, for four tone frequencies (2, 12, 16 or 24 kHz) shown in different colors. These results showed that detection thresholds decreased when the angular separation between noise and signal sources increased. A angular separation of 60° between a 16 kHz tone and a 10 dB SPL spectrum level noise caused a decrement in detection thresholds of about 15 dB relative to the condition where target and masker had angular separation of 0° (spatial release in masking). Spatial release from masking on the horizontal plane gradually increased as function of the angular separation in azimuth (either 0°, 15°, 30°, 45°, 60° or 90°). A two-way ANOVA revealed that the azimuthal separation between tone and noise caused a significant decrement in thresholds for both monkeys (monkey A:  $F(5, 15) = 15.41$ ,  $P < 0.0001$ ; monkey B:  $F(5, 15) = 25.24$ ,  $P < 0.0001$ ). The results of the Dunnett's post-hoc test indicated that thresholds significantly dropped off when the angular separation between masker and signal was equal or greater than 30 deg for monkey A and 15 deg for monkey B. The frequencies at

which tones were played significantly affected detection thresholds for monkey B (monkey A:  $F(3, 15) = 1.208$ ,  $P = 0.3409$ ; monkey B:  $F(3, 15) = 5.149$ ,  $P = 0.0120$ ). The Dunnett's post-hoc test showed that only the difference in thresholds between 12 kHz and 16 kHz was significant for monkey B.

A similar experiment was conducted by varying the location of the noise masker over the vertical plane. Behavioral performance was measured (Fig. 6) for 5 different elevations (either 0, 24.5, 36.5, 49 or 56 deg). Fig. 6A shows two psychometric functions for a 12 kHz tone that was separated in elevation from the masker by

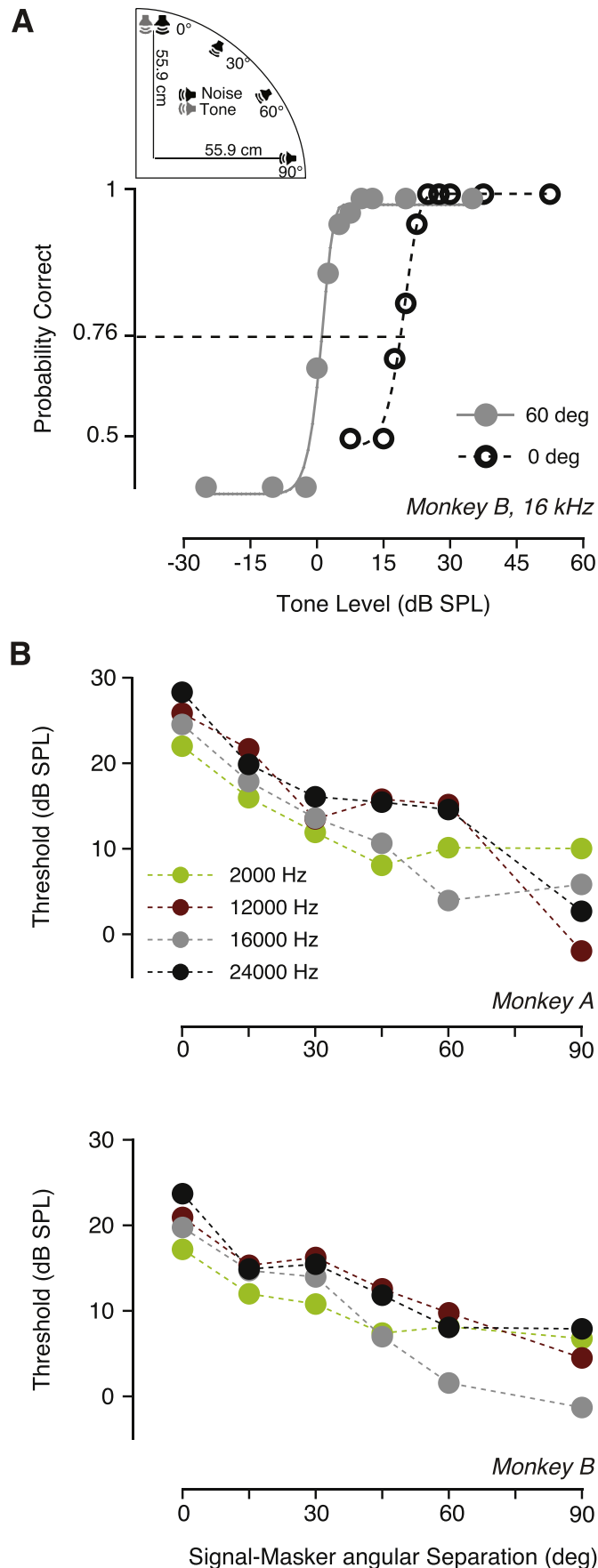
24.5° and 56°. These results showed that when the angular separation between tone and noise was 56°, detection thresholds decreased of about 6 dB compared to when the angular separation was 24.5°. Tone detectability improved as a function of the angular separation in elevation. Although detection thresholds decreased depending on the spatial relationship between tone and noise for both monkeys (Fig. 6B), this effect was smaller than that observed for the azimuthal plane. A two-way ANOVA revealed a significant decrement in thresholds only for monkey A (monkey A:  $F(4, 8) = 12.52$ ,  $P = 0.0016$ ; monkey B:  $F(4, 8) = 2.196$ ,  $P = 0.1596$ ). The Dunnett's post-hoc test showed that thresholds for monkey A were significantly different when the minimum angular separation between masker and signal was 36.5. Signal frequency significantly affected thresholds for monkey A (monkey A:  $F(2, 8) = 24.41$ ,  $P = 0.0004$ ). The Dunnett's post-hoc test revealed a significant difference in detection performance (for monkey A) between 24 kHz and the other frequencies played (2 kHz and 12 kHz).

## 4. Discussion

The results indicate how the detectability of a tone is affected by varying its temporal and spatial relationship to a noise masker.

### 4.1. Temporal cues within acoustic scenes affect tone detectability

Differences in tone detection due to the characteristics of the noise masker may provide crucial information concerning perceptual adaptation. Physiological work (Costalupes et al., 1984)



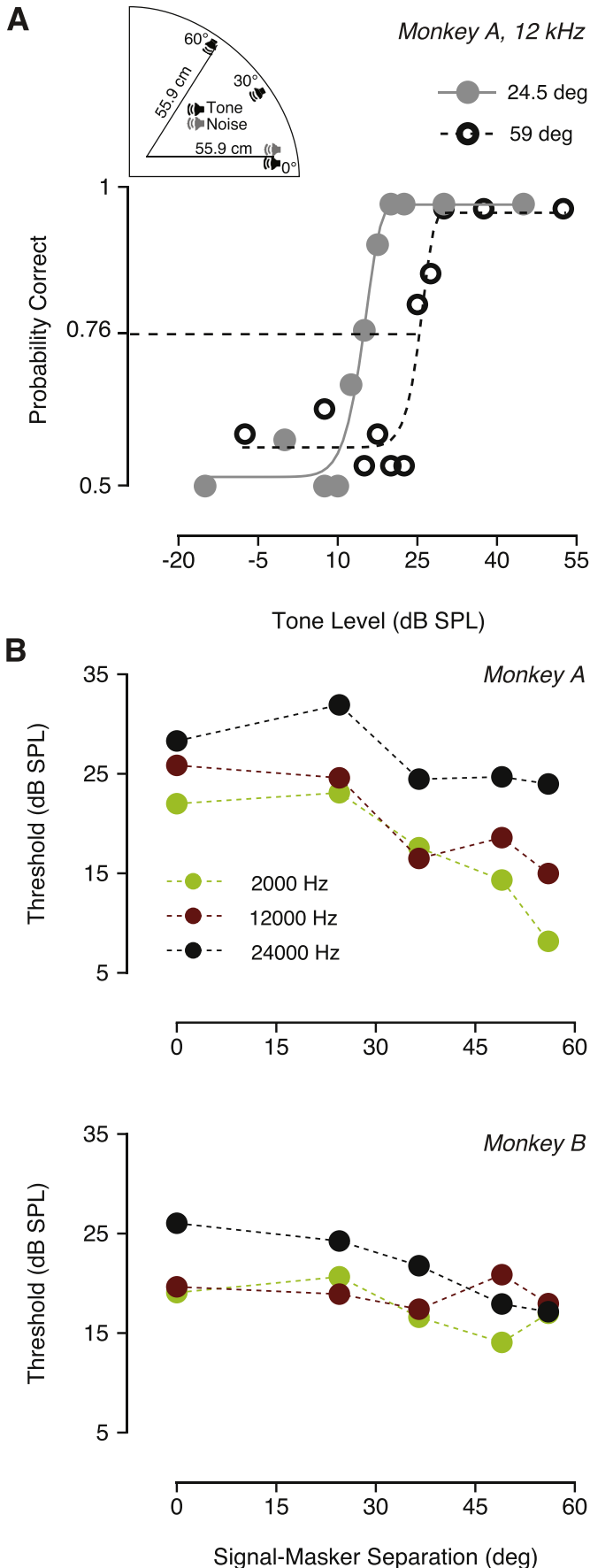
has compared the effects of continuous and gated broadband noise on the average discharge rate to best frequency (BF) tones recorded from auditory nerve fibers of anesthetized cats. Although Costalupes et al. (1984) observed a reduced discharge rate of auditory nerve fibers when noise was continuous, the horizontal shift of the neuronal dynamic range was similar for gated and continuous noise. These findings are consistent with the results described by Rees and Palmer (1988) in a study where the dynamic range shift of neurons in the inferior colliculus (IC) of anesthetized guinea pigs was observed to be approximately the same regardless of the characteristics of the noise background (either simultaneously gated or continuous). Although adaptation to previous noise caused the same rate function for continuous and inversely gated noise, a significant horizontal dynamic range shift towards higher sound pressure levels was observed in the presence of continuous noise (Costalupes et al., 1984; Rees and Palmer, 1988). These results led Costalupes et al. (1984) to hypothesize that suppression/inhibition mechanisms could explain the adjustment of the dynamic range observed during masking.

The first experiment investigated the effect of simultaneous vs. non-simultaneous maskers (Fig. 1). Detection thresholds did not vary significantly depending on whether the noise background was continuous or simultaneously gated with the tone, consistent with physiological work (VIII nerve: Costalupes et al., 1984; Inferior colliculus: Rees and Palmer, 1988). Psychophysical work (e.g. Leshowitz and Cudahy, 1975; Zwicker and Fastl, 1972) has shown that when using sinusoidal stimuli, masking produced by simultaneously gated maskers was larger than for continuous maskers. Wier et al. (1977) employed an adaptive two-interval forced-choice (2IFC) task where tone signals were embedded in a broadband noise (100–10,000 Hz), and found that gated noise produced larger masking compared to continuous noise. However, the thresholds for gated noise were only 1–3 dB higher than those for continuous noise, and the variability across subject was high. The disparity between masking effects produced by gated and continuous noise described in early psychophysical studies (Leshowitz and Cudahy, 1975; Sherrick and Mangabeira-Albernaz, 1961; Zwicker and Fastl, 1972) could be due to differences in the stimuli used. The tone duration employed here to compare detection thresholds across different noise backgrounds (continuous versus gated noise) is longer than the tone duration tested in previous studies (e.g. Leshowitz and Cudahy, 1975; Zwicker and Fastl, 1972). However, a pilot experiment revealed that the similarity in detection thresholds did not change as a function of the type of noise masking used even when a shorter duration signal was presented. Another explanation that might account for this discrepancy could be related to differences in the methodology. Differently from the Go/no Go detection task employed here, when a two-alternative forced-choice procedure is used listeners are always required to perform also a discrimination task, since they are asked to identify which interval contains the signal.

Non-simultaneous masking effects provide salient information concerning how the auditory system segregates a tone target from a masker. The onset envelope of the signal represents an

**Fig. 5.** Spatial release from masking along the azimuthal plane. (A) Schematic shows that a tone was presented in front of the monkeys' head, whereas the noise masker (10 dB SPL spectrum level) was presented from different azimuths. Exemplar psychometric functions from monkey B (16 kHz, 0° azimuth) for maskers arising from two azimuths (0° and 60°) showed the shift in dynamic range due to the spatial separation between tone target and noise masker. (B) Detection threshold as a function of azimuthal angular separation between signal and masker for monkeys A (top) and B (bottom). Different colors show thresholds for different frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)





informative cue to detect sounds within noisy backgrounds. In this study, we systematically manipulated the temporal relationship between the onset of the signal and the onset of the noise. Non-simultaneous masking was observed both when the signal onset followed the masker onset (forward masking) and when the signal onset preceded the masker onset (backward masking). Forward masking has been extensively studied in humans (e.g., de Maré, 1940; Lüscher and Zwislocki, 1947; Munson and Gardner, 1950; Zwislocki et al., 1959; Plomp, 1964; Elliott, 1971; Widin and Viemeister, 1979; Jesteadt et al., 1982; Moore and Glasberg, 1983; Nelson, 1991; Plack and Oxenham, 1998). Fig. 2 shows that forward masking significantly decreased when the delay between the onset of the signal and the masker (negative asynchrony) increased, consistent with previous studies noted above. Our results showed that the temporal window over which forward masking effects can be observed was about 200 ms, consistent with results in humans showing a masking effect decrement to zero after about 200 ms. Forward masking was affected by the proportion of overlapping signal and noise. When the noise preceded the signal by 100 ms (–100 ms onset asynchrony) but had the same offset of the tone (complete overlap of signal and noise), no release from masking was observed (Fig. 2 C). When the last 100 ms of tones were left unmasked (the masker and the signal had different offsets) thresholds were lower. This result suggests that the auditory system might integrate information over the entire duration of both signal and masker.

Less attention is generally given to the effects of backward masking. However, previous studies have reported smaller backward masking effects in human subjects (e.g.; Miyazaki and Sasaki, 1981; Oxenham and Moore, 1994). Consistently with previous work, Fig. 2 showed that the effect of backward masking was smaller compared to that observed with forward masking. Although there is a general agreement concerning the existence of a larger forward masking effect, the results presented in this paper suggested that the shape of temporal filters (such as those in Oxenham and Moore, 1994) might vary as a function of the duration of the signal and the noise masker presented (Fig. 3). Further work is needed in order to clarify these results. Fig. 3 shows that masker duration might affect tone detectability. These results are consistent with early psychophysical work (Fastl, 1976) that found a decrement in forward masking due to the duration of a broadband noise masker. However, the masking effect in our study was significant only for durations longer than 100 ms, and only for onset asynchronies of 0ms. The results summarized in Fig. 3 suggest that the decrement in forward and backward masking is quicker when signal and masker are set apart by a temporal gap (the noise does overlap with the signal only when the temporal gap is zero) compared to when a proportion of the masker overlaps with the signal (Fig. 2). This confirms previous findings (Fig. 2C), which suggested that the masking is related to the temporal overlap between the signal and the masker.

A brief tone signal presented soon after the onset of a masker has higher threshold compared to when the tone onset was delayed from the noise onset (Zwicker, 1965). In humans, the overshoot effect of 8 dB is well matched with the data obtained in

**Fig. 6.** Spatial release from masking along elevation. (A) Schematic shows that the location of the signal was fixed at 0° azimuth and elevation, whereas the noise (10 dB SPL spectrum level) was presented at different elevations (0° azimuth). Examples of two psychometric functions from monkey A (12 kHz) with noise originating at two elevations (24.5 and 60 deg) show the shift in dynamic range due to masker-signal angular separation. (B) Detection threshold plotted as a function of the elevation angular separation between signal and masker. Different colors show thresholds for different frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

macaques (Fig. 4). The overshoot has been previously explained in terms of short-term neural adaptation (e.g. Smith and Zwislocki, 1975; Smith, 1979; Westerman and Smith, 1984), and is a nonlinearity that can only be explained by processes beyond the auditory nerve. Recent studies in animal models show cortical responses that are suggestive of overshoot (Gay et al., 2014). Some previous studies (e.g. von Klitzing and Kohlrausch, 1994; Strickland, 2001, 2004; 2008; Keefe et al., 2009; Walsh et al., 2010a, 2010b; McFadden et al., 2010) suggested that the olivocochlear efferent system could be involved, suggesting a substrate for the suppression. In a recent work, McFadden et al. (2010) explored the mechanisms underlying the overshoot effect by asking to human listeners to detect a 10 ms tone masked by 400 ms wideband noise, observing that detectability remained approximately constant over the first 20–30 ms of signal delay (hesitation) before behavioral sensitivity started to improve. This finding led McFadden et al. (2010) to assume that because of the similarity between this hesitation interval and the hesitation observed in physiological measurements conducted in the medial olivocochlear (MOC) system, the overshoot effect might mainly depend on such a system. Our results (Fig. 4) showed a hesitation interval of about 75–100 ms in the macaques, which is larger than that observed by McFadden et al. (2010). This difference might be due either to species differences, the duration of the tone and noise, or to the methodology adopted.

The results outlined in this manuscript showed that the amount of overlap between signal and masker is an important feature of masked detection. This is consistent with a model of auditory detection previously developed (Bohlen et al., 2014), that implicates accumulation/integration, and it is generally consistent with models of sensory processing and sensorimotor integration (Vickers, 1970; Usher and McClelland, 2001). Models of central neurons have also incorporated accumulation as an important aspect of computing decision (e.g., Wang, 2008). The window obtained in Fig. 2 may be a first approximation of temporal windows that have been well described for human hearing (e.g., Oxenham and Moore, 1994).

The results outlined in the present study support the hypothesis that a neuronal suppression and adaptation might be the dominant mechanism underlying tone detection thresholds in simultaneous and non-simultaneous masking contexts. Tone detectability was significantly affected by the onset asynchrony between the signal and the masker. The proportion of noise energy present at the same time of the signal (e.g. simultaneously gated noise condition or onset asynchrony conditions where signal and masker overlapped) significantly affected detection threshold. Masking effects caused by simultaneously gated noise did not vary when previous noise exposure (i.e. adaptation to continuous noise played before the signal) was present. In a recent study Nelson et al. (2009) recorded neuronal activity from the central nucleus the inferior colliculus of awake marmosets, and concluded that forward masking could be better explained in terms of inhibitory mechanisms either in the IC or at lower sites of the auditory pathways. A possible interpretation is that the inferior colliculus might receive inhibitory inputs from the superior paraolivary nucleus (SPON), that would result in a wide dynamic range inhibitory mechanism locked to sound offset (Gai, 2016). However, the large variability in forward masking in their dataset makes difficult to reconcile their neuronal threshold shifts with psychophysical threshold shifts. Alves-Pinto et al. (2010) explored the magnitude of forward masking by recording from neurons on the auditory cortex of anesthetized guinea pigs. They found that threshold shifts were larger in the cortex compared to those observed in subcortical areas, suggesting that cortical processes might be involved in forward masking.

#### 4.2. Tone detectability is affected by spatial cues: free-field release from masking in non-human primates

In the last two experimental conditions summarized in this study, we explored to what extent spatial location cues employed to segregate signals from the background lead to spatial release from masking (SRM) within auditory scenes. Sound localization relies on both monaural and binaural spatial cues characterizing any natural acoustic environment (Yost and Gourevitch, 1987). It is generally accepted that the human auditory system weights ITDs and ILDs binaural cues for processing low frequencies (below about 1–2 kHz) and high frequencies (above about 1–2 kHz), respectively (Strutt, 1907). However, to discriminate the elevation of sound signals from maskers, the auditory system might have to rely on monaural spectral cues arising from the interaction of sound sources with the pinnae (Wightman and Kistler, 1989). Although this topic has been fairly largely investigated in humans, a relative lack of research on nonhuman species leads to a poor understanding of the neural mechanisms underlying spatial release from masking, and even a question of whether spatial release from masking occurs in animal models. In fact, Sollini et al. (2016) suggest that under some conditions, ferrets do not show spatial release from masking.

The first experimental condition described here aimed to explore non-human primates' ability to correctly segregate a target tone from a broadband white noise masker along the azimuthal plane (Fig. 5). The experimental design adopted here was similar to the one previously used by Saberi et al. (1991) to test free field release from masking in humans in which they concluded that detectability of the signal improves as the separation between clicks and noise is increased. However, the experiments conducted by Saberi et al. (1991) involved broadband signals while the target signal in the present work was always a pure tone.

The results summarized here (Fig. 5) are consistent with previous work in humans on free field signal detection and localization in noise (Kidd et al., 1998; Good and Gilkey, 1996; Lorenzi et al., 1999; Arbogast et al., 2005; Lingner et al., 2012), suggesting that monkeys show spatial release from masking, and that detection performance improves with increasing the angular separation between masker and signal. Fig. 5 shows a significant gradual decrement in tone detection thresholds as a function of the angular separation between signal and noise. Our results showed a significant effect of frequency only for one monkey. However, this was probably due to the range of frequencies used, which was composed of relatively high frequencies (equal of above 2 kHz).

We also explored whether angular separation in elevation could lead to spatial release in masking, and whether macaques can use spectral cues to detect masked signals. To the best of our knowledge, this is the first report of SRM measured with tonal signals. Masker locations were varied along elevation (0° azimuth). The results described in this study (Fig. 6) showed that spatial release from masking could be observed also along the vertical plane, at least in one subject. The mechanisms underlying spatial release from masking across the vertical plane are still unclear. Spectral cues are considered a good “cue candidate” used by the auditory system to detect changes in sound elevation (Khun, 1987). This finding is consistent with results of human studies (e.g. Saberi et al., 1991; Grantham et al., 2003), and might suggest that non-human primates can use spectral cues to detect sound locations when only such information is available. However, the magnitude of the effect in elevation reported here was smaller compared to that measured by separating signal and noise in azimuth. Spatial separation in elevation between signal and masker led to a significant change in thresholds only for one monkey. This may suggest individual differences in elevation decoding in macaques that needs

further exploration.

The mechanism underlying spatial release from masking in elevation may be a simple signal to noise detection at the tone frequency. If the tone frequency coincided with the notch created by the head related transfer function, tone thresholds will be lower re: noise from another location that did not show a notch at the tone frequency. Head related transfer functions have been explored in macaques (Spezio et al., 2000), who found that spectral notches occur at frequencies >10 kHz in macaques; further these notches occur at higher frequencies as the sound progressed from lower elevations to higher elevations. Spezio et al. (2000) explored notches only up to 15 kHz. Our studies show that notches are created at higher frequencies (up to 24 kHz) for elevations up to 60°.

The behavioral results outlined in this paper represent the basis for future physiological investigations concerning free field release from masking in non-humans primate. The hypothesis of two different neuronal mechanisms accounting for binaural and monaural spatial cues is not controversial. It has been hypothesized that the monaural pathways might originate at the dorsal cochlear nucleus (e.g., Young et al., 1992), whereas the superior olivary complex represents the first site of the binaural pathway (Goldberg and Brown, 1969; Tsuchitani and Boudreau, 1969). There are studies suggesting that these two pathways merge at the level of the inferior colliculus (e.g., Majorossy and Kiss, 1994; Oliver et al., 1997). The existence of two separate streams has been proposed also by a recent neuroimaging work (Thompson et al., 2006), showing that lateralized fMRI responses in the midbrain switch sides when ITD is increased. Despite the general agreement concerning the existence of two separate pathways for processing binaural and monaural spatial cues (Schwartz, 1992), it is not clear to what extent these two possible “channels” are independent, when spatial cues arising from acoustic images have to be combined to reliably perform a task. It could be hypothesized the final acoustic “spatial percept” could always arise from a combination of binaural and monaural spatial cues at the level of the inferior colliculus or the auditory cortex (Heffner and Masterton, 1975; Heffner and Heffner, 1990).

## Acknowledgments

The authors would like to thank Mary Feurtado for surgical assistance, and Bruce and Roger Williams for building hardware. Srikanth Reddy, Courtney Timms, and Abigail Bernard helped collect parts of the data. Research was supported by the National Institute of Deafness and Communication Disorders (NIDCD) of the National Institutes of Health (NIH) under award number R01 DC11092.

## References

- de Maré, G., 1940. Fresh observations on the so-called masking effect of the ear and its possible diagnostic significance. *Acta Oto-Laryngol.* 28, 314–316.
- Alves-Pinto, A., Baudoux, S., Palmer, A.R., Sumner, C.J., 2010. Forward masking estimated by signal detection theory analysis of neuronal responses in primary auditory cortex. *JARO* 11, 477–494.
- Arbogast, T.L., Mason, C.R., Kidd Jr., G., 2005. The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 117, 2169–2180.
- Bohlen, P., Dylla, M., Timms, C., Ramachandran, R., 2014. Detection of modulated tones in modulated noise by non-human primates. *JARO* 15 (5), 801–821. <http://dx.doi.org/10.1007/s10162-014-0467-7>.
- Bregman, A.S., 1990. *Auditory Scene Analysis. The Perceptual Organization of Sound.* MIT Press, Cambridge.
- Brosch, M., Schreiner, C.E., 1997. Time course of forward masking tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 77, 923–943.
- Christison-Lagay, K.L., Cohen, Y.E., 2014. Behavioral correlates of auditory streaming in rhesus macaques. *Hear. Res.* 309, 17–25.
- Costalupes, J.A., Young, E.D., Gibson, D.J., 1984. Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. *J. Neurophysiol.* 51, 1326–1344.
- Duifhuis, H., 1973. Consequences of peripheral frequency selectivity for nonsimultaneous masking. *J. Acoust. Soc. Am.* 54, 1471–1488.
- Dylla, M., Hrnicek, A., Rice, C., Ramachandran, R., 2013. Detection of tones and their modification by noise in nonhuman primates. *J. Assoc. Res. Otolaryngol.* 14, 547–560.
- Elliott, L.L., 1971. Backward and forward masking. *Audiology* 10, 65–76.
- Fastl, H., 1976. Temporal masking effects: I. Broadband noise masker. *Acustica* 35, 287–302.
- Gai, Y., 2016. ON and OFF inhibition as mechanisms for forward masking in the inferior colliculus: a modeling study. *J. Neurophysiol.* 115 (5), 2485–2500.
- Gatehouse, R.W., 1986. Spatially dependent free field masked thresholds. In: *Proceedings of the 12th International Congress on Acoustics*, pp. B2–B7 (Canada).
- Gatehouse, R.W., 1987. Further research on free field masking. *J. Acoust. Soc. Am.* 82 (Suppl. 1), 108.
- Gay, J.D., Voytenko, S.V., Galazyuk, A.V., Rosen, M.J., 2014. Developmental hearing loss impairs signal detection in noise: putative central mechanisms. *Front. Syst. Neurosci.* 8 (162) <http://dx.doi.org/10.3389/fnsys.2014.00162>.
- Goldberg, J.M., Brown, P.B., 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* 32, 613–636.
- Good, M., Gilkey, R., 1996. Sound localization in noise: the effect of signal-to-noise ratio. *J. Acoust. Soc. Am.* 99, 1108–1125.
- Grantham, D.W., Hornsby, B.W., Erpenbeck, E.A., 2003. Auditory spatial resolution in horizontal, vertical, and diagonal planes. *J. Acoust. Soc. Am.* 114, 1009–1022.
- Green, D.M., Swets, J.A., 1966. *Signal Detection Theory and Psychophysics.* Wiley, New York.
- Harris, D., Dallos, P., 1979. Forward masking of auditory nerve fiber responses. *J. Neurophysiol.* 42, 1083–1107.
- Heffner, H.E., Heffner, R.S., 1990. Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J. Neurophysiol.* 64, 915–931.
- Heffner, H., Masterton, B., 1975. Contribution of auditory cortex to sound localization in the monkey (*Macaca mulatta*). *J. Neurophysiol.* 38, 1340–1358.
- Jesteadt, W., Bacon, S.P., Lehman, J.R., 1982. Forward masking as a function of frequency, masker level, and signal delay. *J. Acoust. Soc. Am.* 71, 950–962.
- Keefe, D.H., Schairer, K.S., Ellison, J.C., Fitzpatrick, D.F., Jesteadt, W., 2009. Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot. *J. Acoust. Soc. Am.* 125, 1595–1604.
- Khun, G., 1987. Physical acoustics and measurements pertaining to directional hearing. In: Yost, William A., Gourevitch, George (Eds.), *Directional Hearing.* Springer-Verlag, New York.
- Kidd, G., Mason, C.R., Rohtla, T.L., Deliwal, P.S., 1998. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Acoust. Soc. Am.* 104, 422–431.
- Leshowitz, B., Cudahy, E., 1975. Masking patterns for continuous and gated sinusoids. *J. Acoust. Soc. Am.* 58, 235–242.
- Lingner, A., Wiegand, L., Grothe, B., 2012. Sound localization in noise by gerbils and humans. *J. Assoc. Res. Otolaryngol.* 13, 237–248.
- Lorenzi, C., Gatehouse, S., Lever, C., 1999. Sound localization in noise in normal-hearing listeners. *J. Acoust. Soc. Am.* 105, 1810–1830.
- Lüscher, E., Zwislocki, J.J., 1947. The delay of sensation and the remainder of adaptation after short pure-tone impulses on the ear. *Acta Oto-Laryngol.* 35, 428–455.
- MacMillan, N.A., Creelman, C.D., 2005. *Detection Theory: a User's Guide*, second ed. Lawrence Erlbaum Associates, Mahwah.
- Majorossy, K., Kiss, A., 1994. Convergence of topographic projections to the inferior colliculus from the auditory subcollicular nuclei. *Acta Biol. Hung.* 45, 347–359.
- McFadden, D., Walsh, K.P., Pasanen, E.G., Grenwelle, E.M., 2010. Overshoot using very short signal delays. *J. Acoust. Soc. Am.* 128, 1915–1921.
- Miyazaki, K., Sasaki, T., 1981. Auditory masking patterns in forward-, and backward-, and simultaneous-masking situations. *Jpn. J. Physiol.* 52, 106–112.
- Moore, B.C.J., Glasberg, B.R., 1983. Growth of forward masking for sinusoidal and noise maskers as a function of signal delay: implications for suppression in noise. *J. Acoust. Soc. Am.* 73, 1249–1259.
- Munson, W.A., Gardner, M.B., 1950. Loudness patterns—a new approach. *J. Acoust. Soc. Am.* 22, 177–190.
- Nelson, D.A., 1991. High-level psychophysical tuning curves: forward masking in normal-hearing and hearing-impaired listeners. *J. Speech Hear. Res.* 34, 1233–1249.
- Nelson, D.A., Swain, A.C., 1996. Temporal resolution within the upper accessory excitation of a masker. *Acust. Acta Acust.* 82, 328–334.
- Nelson, P.C., Smith, Z.M., Young, E.D., 2009. Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. *J. Neurosci.* 29, 2553–2562.
- Oliver, D.L., Beckius, G.E., Bishop, D.C., Kuwada, S., 1997. Simultaneous anterograde labeling of axonal layers from lateral superior olive and dorsal cochlear nucleus in the inferior colliculus of cat. *J. Comp. Neurol.* 382, 215–229.
- Osmanski, M.S., Song, X., Wang, X., 2013. The role of harmonic resolvability in pitch perception in a vocal nonhuman primate, the common marmoset (*Callithrix jacchus*). *J. Neurosci.* 33 (21), 9161–9168.
- Oxenham, A.J., Moore, B.C.J., 1994. Modeling the additivity of nonsimultaneous masking. *Hear. Res.* 80, 105–118.
- Oxenham, A.J., Moore, B.C.J., 1995. Additivity of masking in normally hearing and hearing-impaired subjects. *J. Acoust. Soc. Am.* 98, 1921–1935.
- Oxenham, A.J., Moore, B.C.J., 1997. Modeling the effects of peripheral nonlinearity in

- listeners with normal and impaired hearing. In: Jesteadt, W. (Ed.), *Modeling Sensorineural Hearing Loss*. Erlbaum, Hillsdale, NJ, pp. 273–288.
- O'Connor, K.N., Johnson, J.S., Niwa, M., Noriega, N.C., Marshall, E.A., Sutter, M.L., 2011. Amplitude modulation detection as a function of modulation frequency and stimulus duration: comparisons between macaques and humans. *Hear. Res.* 277 (1), 37–43.
- Plack, C.J., Oxenham, A.J., 1998. Basilar-membrane nonlinearity and the growth of forward masking. *J. Acoust. Soc. Am.* 103, 1598–1608.
- Plomp, R., 1964. The rate of decay of auditory sensation. *J. Acoust. Soc. Am.* 36, 277–282.
- Plomp, R., 1976. Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise). *Acta Acustica United Acustica* 34 (4), 200–211.
- Plomp, R., Mimpen, A.M., 1981. Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech-reception threshold for sentences. *Acta Acustica United Acustica* 48 (5), 325–328.
- Ramachandran, R., Lisberger, S.G., 2005. Normal performance and expression of learning in the vestibulo-ocular reflex (VOR) at high frequencies. *J. Neurophysiol.* 93, 2028–2038.
- Rees, A., Palmer, A.R., 1988. Rate-intensity functions and their modification by broadband noise for neurons in the Guinea pig inferior colliculus. *J. Acoust. Soc. Am.* 83, 1488–1498.
- Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., Perrott, D.R., 1991. Free-field release from masking. *J. Acoust. Soc. Am.* 90, 1355–1370.
- Santon, F., 1987. Detection of a pure sound in the presence of masking noise, and its dependence on the angle of incidence of the noise. *Acustica* 63, 222–228.
- Schwartz, I.R., 1992. The superior olivary complex and lateral lemniscal nuclei. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy*. Springer, New York, pp. 117–167.
- Sherrick Jr., C.E., Mangabeira-Albernaz, P.L., 1961. Auditory threshold shifts produced by simultaneously pulsed contralateral stimuli. *J. Acoust. Soc. Am.* 33, 1381–1385.
- Smith, R.L., 1979. Adaptation, saturation, and physiological masking in single auditory-nerve fibers. *J. Acoust. Soc. Am.* 65, 166–178.
- Smith, R.L., Zwislocki, J.J., 1975. Short-term adaptation and incremental responses of single auditory-nerve fibers. *Biol. Cybern.* 17, 169–182.
- Sollini, J., Alves-Pinto, A., Sumner, C.J., 2016. Relating approach-to-target and detection tasks in animal psychoacoustics. *Behav. Neurosci.* 130, 393–405.
- Spezio, M.L., Keller, C.H., Marrocco, R.T., Takahashi, T.T., 2000. Head related transfer functions of the Rhesus monkey. *Hear. Res.* 144, 73–88.
- Strickland, E.A., 2001. The relationship between frequency selectivity and overshoot. *J. Acoust. Soc. Am.* 109, 2062–2073.
- Strickland, E.A., 2004. The temporal effect with notched-noise maskers: analysis in terms of input-output functions. *J. Acoust. Soc. Am.* 115, 2234–2245.
- Strickland, E.A., 2008. The relationship between precursor level and the temporal effect. *J. Acoust. Soc. Am.* 123, 946–954.
- Strutt, J.W., 1907. On our perception of sound direction. *Philos. Mag.* 13, 214–232.
- Thompson, S.K., von Kriegstein, K., Deane-Pratt, A., Marquardt, T., Deichmann, R., Griffiths, T.D., McAlpine, D., 2006. Representation of interaural time delay in the human auditory midbrain. *Nat. Neurosci.* 9, 1096–1098.
- Tsuchi, C., Boudreau, J.C., 1969. Stimulus level of dichotically presented tones and cat superior olive S-segment cell discharge. *J. Acoust. Soc. Am.* 46, 979–988.
- Usher, M., McClelland, J.L., 2001. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108 (3), 550.
- Vickers, D., 1970. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics* 13 (1), 37–58.
- von Klitzing, R., Kohlrausch, A., 1994. Effect of masker level on overshoot in running- and frozen-noise maskers. *J. Acoust. Soc. Am.* 95, 2192–2201.
- Walsh, K.P., Pasanen, E.G., McFadden, D., 2010a. Properties of a nonlinear version of the stimulus-frequency otoacoustic emission. *J. Acoust. Soc. Am.* 127, 955–969.
- Walsh, K.P., Pasanen, E.G., McFadden, D., 2010b. Overshoot measured physiologically and psychophysically in the same human ears. *Hear. Res.* 268, 22–37.
- Wang, X.J., 2008. Decision making in recurrent neuronal circuits. *Neuron* 60 (2), 215–234.
- Westerman, L.A., Smith, R.L., 1984. Rapid and short-term adaptation in auditory nerve responses. *Hear. Res.* 15, 249–260.
- Widin, G.P., Viemeister, N.F., 1979. Intensive and temporal effects in pure-tone forward masking. *J. Acoust. Soc. Am.* 66, 388–439, 5.
- Wier, C.C., Green, D.M., Hafter, E.R., Burkhardt, S., 1977. Detection of a tone burst in continuous- and gated-noise maskers; defects of signal frequency, duration, and masker level. *J. Acoust. Soc. Am.* 61, 1298–1300.
- Wightman, F.L., Kistler, D.J., 1989. Headphone simulation of free-field listening. II: psychophysical validation. *J. Acoust. Soc. Am.* 85, 868–878.
- Yost, W.A., Gourevitch, G., 1987. *Directional Hearing*. Springer-Verlag, New York.
- Young, E.D., Spirou, G.A., Rice, J.J., Voigt, H.F., 1992. Neural organization and responses to complex stimuli in the dorsal cochlear nucleus. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 336, 407–413.
- Zwicker, E., 1965. Temporal effects in simultaneous masking by white-noise bursts. *J. Acoust. Soc. Am.* 37, 653–663.
- Zwicker, E., Fastl, H., 1972. On the development of the critical band. *J. Acoust. Soc. Am.* 52, 699–702.
- Zwislocki, J., Pirodda, E., Rubin, H., 1959. On some post-stimulatory effects at the threshold of audibility. *J. Acoust. Soc. Am.* 31, 9–14.