Behavioral/Systems/Cognitive

Illusory Sound Perception in Macaque Monkeys

Christopher I. Petkov,[1] **Kevin N. O'Connor,**[2] **and Mitchell L. Sutter**[1,2]

[1]Center for Neuroscience and [2]Section of Neurobiology, Physiology and Behavior, University of California, Davis, California 95616

In most natural listening environments, noise occludes objects of interest, and it would be beneficial for an organism to correctly identify those objects. When a sound of interest ("foreground" sound) is interrupted by a loud noise, subjects perceive the entire sound, even if the noise was intense enough to completely mask a part of it. This phenomenon can be exploited to create an illusion: when a silent gap is introduced into the foreground and high-intensity noise is superimposed into the gap, subjects report the foreground as continuing through the noise although that portion of the foreground was deleted. This phenomenon, referred to as auditory induction or amodal completion, is conceptually similar to visual induction, fill-in, illusory motion, and illusory contours. Two rhesus macaque monkeys performed a task designed to assess auditory induction. They were trained to discriminate complete stimuli from those containing a silent gap in the presence of two types of noise. Interrupting noise temporally coincided only with the gap, and in humans this causes induction. Surrounding noise temporally encompassed the entire foreground, and in humans this causes masking without auditory induction. Consistent with previous human psychophysical results, macaques showed better performance with surrounding masking noise than interrupting noise designed to elicit induction. These and other control experiments provide evidence that primates may share a general mechanism to perceptually complete missing sounds.

Key words: auditory; induction; scene analysis; psychophysics; segmentation; speech; illusion

Introduction

How can sounds be discerned in noisy, complex listening environments? For example, how can people follow a conversation in a lively "cocktail party" or monkeys discriminate another monkey's vocalization in natural environments, like a jungle or forest, with many other animals and sound sources?

Auditory induction provides an example of the nervous system extracting a sound of interest (i.e., a "foreground" sound) from noise introduced by other sound-producing objects. When a foreground is interrupted by a brief noise, the entire foreground can be heard continuing through the noise, even with noise intense enough to completely mask the underlying foreground sound (Warren, 1970). That interrupting noise does not make a foreground sound less perceptible seems paradoxical but makes sense if we consider that the nervous system restores missing information when a foreground is interrupted by intense noise. Warren (1970) exploited this perceptual restoration to create an illusion. When silent gaps were interspersed in speech with high-intensity noise superimposed during the gaps, subjects reported the speech as continuing through the noise, although the speech signal that should have occurred during the noise was deleted. Without noise, the missing parts of speech were reported as discontinuous, and the speech was reported as incomprehensible. Similar illusory induction has been demonstrated for non-speech sounds [tones (Warren et al., 1972, 1988; Bennett et al., 1984;

frequency modulated (FM) sweeps (Dannenbring and Bregman, 1976; Ciocca and Bregman, 1987; Kluender and Jenison, 1992)], indicating that perceptual completion of sounds interrupted by noise generalizes to different types of foregrounds.

These results suggest that the auditory system forms a "model" of the foreground sound that influences perception. Unless evidence is provided to the contrary, the auditory system maintains the foreground perceptual model. When interrupted by intense, short noise, the foreground information occurring during the noise is unavailable to the auditory system, so it relies on the foreground model. Alternatively, for less intense noises and longer gaps, the break in the foreground can be detected with confidence and thus the model can be rejected and overridden.

Determining whether animals share similar mechanisms, a first step in understanding the neural underpinnings of auditory induction, will further the understanding of how the brain analyzes acoustically complex environments. Investigation of the tendency of cotton-top tamarins to vocalize in response to complete vocalizations provided a behavioral correlate of induction (Miller et al., 2001). Evidence for neural correlates of auditory induction (Sugita, 1997), visual induction (Rossi and Paradiso, 1999), and inferred visual motion (Assad and Maunsell, 1995) indicates that animals share some gestalt properties inherent in analyzing a scene.
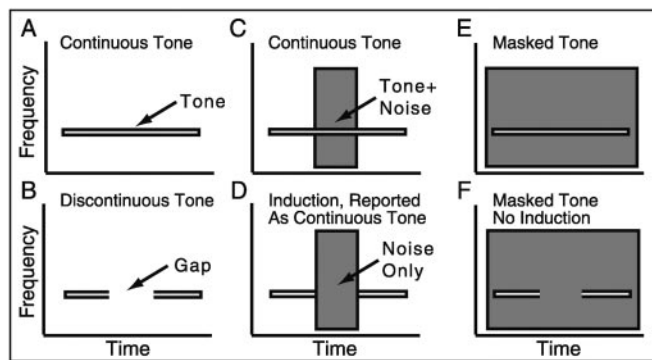
We used standard operant techniques to measure the ability of macaque monkeys to detect the continuity of foreground sounds in the presence of different types of noise. Some conditions were chosen to cause induction and others to cause masking, allowing us to compare macaque thresholds with human thresholds that are linked to auditory induction. The results are consistent with human results (Kluender and Jenison, 1992), indicating that induction is present in macaques.

Figure 1. Schematized spectrograms of induced and masked foregrounds. A single frequency tone is depicted as a horizontal bar with a narrow frequency range (*A*), but noise has a broader frequency spectrum (see Noise Only in *D*). Higher intensity components are represented by darker shading. *A*, A complete tone segment is reported as continuous by human listeners. *B*, A tone with a silent portion (gap) is reported as sounding discontinuous by listeners who readily detect the gap; however, when a higher intensity interrupting noise fills this gap (*D*), most listeners report that the tone was continuous throughout as if the tone were present in the noise (auditory induction). In this case they report that the stimulus in *D* sounds like the stimulus shown in *C*. *E*, When intense noise is changed to completely "surround" the foreground, the foreground is masked and continuity cannot be determined. Note, when lower intensity "surrounding" noise is used, no masking occurs, and the entire foreground is heard correctly. *F*, When a foreground with a gap is used with intense surrounding noise, the entire foreground is masked and continuity cannot be determined; however, with lower intensity noise, no masking occurs, and the discontinuity in the foreground can be readily identified. Note that there are intermediate intensities at which induction can occur without masking such that *D* would be reported as continuous and *F* would be reported as discontinuous.

Materials and Methods

In general, two monkeys were trained to detect the continuity/discontinuity of foreground sounds presented under conditions known to cause either auditory induction (Fig. 1 *D*) or masking (Fig. 1 *F*) in humans. The foreground sounds were tones, FM sweeps, or a vocalization; on some trials they were continuous, and on other trials they were discontinuous. The goal was to determine whether the animals responded to discontinuous foreground sounds as if they were continuous under conditions that cause induction.

Behavioral techniques

Experiments were performed with two adult male macaque monkeys (*Macacca mulatta*) and conformed to the policy of the National Institutes of Health on experimental animal care and a protocol approved by the University of California Davis animal care and use committee. Training occurred with macaques seated in a primate chair designed to be "acoustically transparent" within a sound-attenuated chamber. A drinking tube and a response lever were mounted on the chair in front of the macaque. Using standard operant conditioning techniques with juice serving as positive reinforcement, macaques were trained gradually to perform a go no-go task. Subjects were trained initially to detect gaps in tones without noise. Within a session, gap duration was varied. We then gradually introduced noise (first low-intensity noise), still varying gap duration within a session. On each session the noise intensity was fixed, and gap duration varied. On successive sessions higher-intensity noise was introduced. During initial training, within a daily session we interleaved surrounding and interrupting noise to avoid overexposure to either noise type.

In the go no-go task (Fig. 2), the animal had to press a lever to begin a trial and release a lever if the second of two sounds differed from the first (in which case the second sound is called a target) but had to keep holding the lever down if the second and first sound were identical (in which case the second sound is called a standard). The interstimulus interval was 400 msec, and intertrial interval varied (depending on how quickly the subject initiated a new trial) but was at least 1 sec. If the subject failed to release the lever within 800 msec after a target presentation, a "miss" was scored. If the animal correctly released within 800 msec, a "hit" was scored. Approximately one-fifth of the trials were "catch" trials in which

two standards were presented, and the lever was not to be released for a "correct rejection" to be scored. If subjects released the lever during these trials a "false alarm" was scored. Hits and correct rejections were reinforced with juice. No juice followed an incorrect response (miss or false alarm). False alarm responses were also penalized with a time-out.

Stimuli

Stimuli consisted of foreground sounds presented either with or without background noise. Target stimuli were identical to standards except that they contained a gap or silent portion in the foreground sound. Correct detection of gaps and correct rejection of continuous foregrounds were reinforced with juice. This acted as a conservative measure of induction because macaques were trained to detect the gaps in the foregrounds and could maximize reinforcement by attempting to overcome the effects of induction.

One of three different foregrounds could be used. One foreground was a 400 msec, 2 kHz tone, presented at 45 dB sound pressure level (SPL) (unfiltered calibration; Bruel & Kjaer 2231) and sampled at 50 kHz. Another foreground was a 400 msec, positively sweeping linear FM tone centered at 2.03 kHz (range, 1680–2380 Hz), presented at 45 dB SPL, and sampled at 50 kHz. The third foreground, provided by M. D. Hauser and described in Hauser (1991), was a complete 512 msec "coo," presented at 45 dB SPL and sampled at 25 kHz. The beginning (onset) and end (offset) of the entire foreground was cosine ramped with an 8 msec rise–fall time; transitions into and out of gaps, which were temporally centered in the stimulus, were cosine ramped with 3 msec rise–fall times. The gap duration was defined as the silent portion plus the cosine ramp.

Either interrupting (designed to elicit auditory induction) or surrounding noise (a masking control) was used. All noise components were calibrated in RMS level (SPL, unfiltered; Bruel & Kjaer 2231) and were "frozen" within a given trial so that the only difference between standard and target stimuli was the presence or absence of a silent gap in the foreground. Interrupting noise was not ramped and was temporally centered over the foreground; when a gap was present (target stimulus), this also corresponded to a noise that completely overlapped the gap (including the 3 msec ramps) but not the foreground (Figs. 1 *D*, 2). Surrounding noise completely surrounded the foreground in time and included an extra 25 msec onset and offset ramp, such that the noise reached the plateau of its ramping when the foreground began, and began to transition off when the foreground was completed (schematically depicted in Fig. 1, *E* and *F*; ramps not shown). Surrounding noise masks the foreground but does not elicit auditory induction because it is temporally mismatched from the silent gap (Bregman, 1990; Kluender and Jenison, 1992). This allowed for comparisons of performance under conditions that elicit masking and induction in humans (Kluender and Jenison, 1992). If macaque performance is comparable with human performance that is closely linked with masking and induction, then the hypothesis that auditory induction occurs in macaques is supported.

Notched noise was also used. Here we removed the spectral energy of the interrupting or surrounding noises for a 1 octave region at 2 kHz (for the 2 kHz and FM foregrounds) or at 8 kHz (for only the 2 kHz tonal foreground). These stimuli were designed to reduce masking and induction with the matched spectral notch (at 2 kHz) but not with the unmatched 8 kHz notch stimuli, because a requirement of auditory induction and masking is that the noise contains energy near the frequency of interest (Warren et al., 1972).

Experimental variables

Two different variables, intensity and gap duration, were parametrically manipulated to provide two independent assessments of induction versus masking. Five to eight different values of each variable were used for each type of experiment. Noise intensity varied linearly in 3–6 dB increments (intensity experiments), and gap durations were varied in equal logarithmic intervals from 6 to 206 msec and linearly from 0 to 6 msec (gap duration experiments).

Intensity experiments. For the noise intensity experiments, the intensity of the noise was the independent variable. The gap centered over the stimulus was fixed at 56 msec. Sessions were composed of ~1000 trials (~3–4 hr), with conditions blocked by noise type. Subjects were coun-
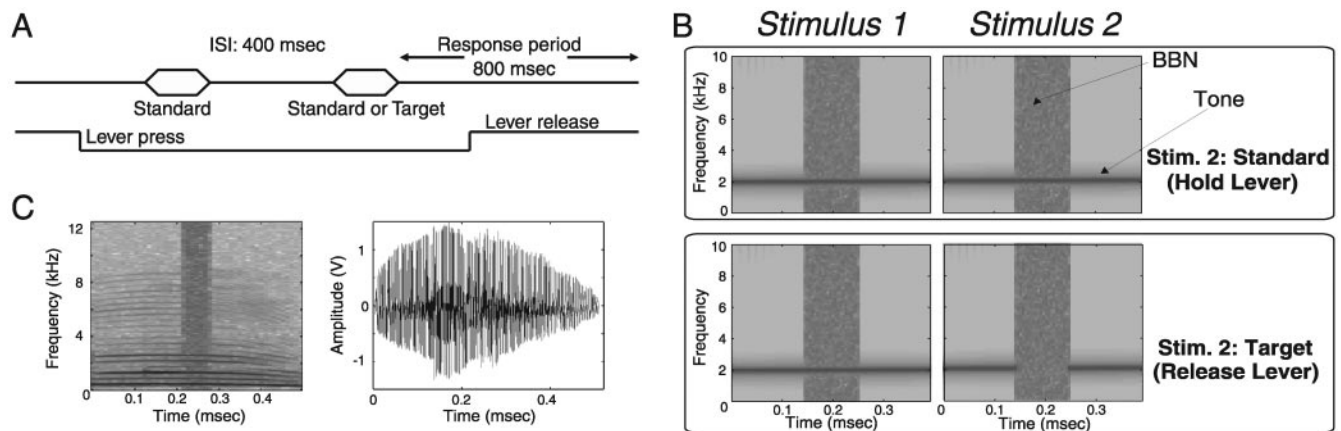
Figure 2. Behavioral task and spectrograms of some of the stimuli used. *A*, Schematized task. After a lever press two sound stimuli were played (hexagons). A lever release was to be made within 800 msec after a target (different from first sound) but not a standard (same as first sound) stimulus for reward. *B*, Spectrograms of 2 kHz tonal foregrounds with interrupting broadband noise (BBN). Macaques were trained to let go of the lever only if the second stimulus was a target (all targets were identical to standards except that they contained a silent gap in the foreground). *C* shows a spectrogram (left) and an amplitude plot (right) of the coo vocalization with interrupting noise.

terbalanced across interrupting and surrounding noise conditions for the noise intensity experiments that were performed with blocked trials, i.e., one macaque began with interrupting and the other began with surrounding noise. To check the consistency of performance as a function of how the stimuli were blocked over sessions, we occasionally interleaved interrupting and surrounding noise stimuli. Performance was not significantly different under these conditions [repeated-measures (RM)-ANOVA (see below); $p > 0.5$]. When possible we repeated the starting condition after an experimental run to verify that thresholds were stable and performance was not statistically different. For example, if we started with interrupting noise blocks and then followed with surrounding noise, in the end we would reconfirm the initial interrupting noise thresholds.

Gap duration experiments. In these experiments, the intensity of the noise was fixed at 40 dB SPL, and the gap duration was varied. Gap duration experiments were also conducted to assess the abilities of macaques to detect small gaps in tonal foregrounds without noise. Noise type (e.g., broad-band, notched, surrounding, interrupting, and no noise) was randomly interleaved within a given session for the gap duration experiments. Finally, experiments were also conducted to assess the abilities of macaques to detect small gaps in broadband noise. This is similar to standard "gap detection" experiments. In these experiments, only a broadband noise (no tones, FMs, or vocalizations) was used, and the animals had to detect differences between continuous noise standards and gap-in-noise targets. This 400 msec, 45 dB noise had 8 msec onset–offset ramps and no ramps to transition into the gap.

Experimental parameters (noise intensity or gap duration) were adjusted using the method of constant stimuli to obtain sigmoidal psychometric functions.

Experimental apparatus

Acoustic stimuli were constructed and digitally edited with Matlab software (MathWorks), presented with 16 bit output resolution (TDT Systems), attenuated (TDT; Leader), amplified (Radio Shack MPA-200), and then delivered through a speaker (Radio Shack PA-110, 10 inch woofer and piezo horn tweeter: 38–27,000 Hz) positioned at ear level 1.5 m in front of the subject. Experimental sessions were conducted in a double-walled IAC, sound-attenuating booth (9.5 × 10.5 × 6.5 feet, internal), lined with foam to reduce echoes.

Data analysis

For each experimental condition, data analysis was based on at least four daily sessions after the animal reached asymptotic performance. Using signal detection theory (Green and Swets, 1974), thresholds were determined using the log-odds ratio (LOR) (Luce et al., 1963). The LOR is a criterion-free measure of sensitivity comparable with d' (Macmillan and Creelman, 1991). The LOR is based on a logistic probability distribution

and is closely correlated with d', which is based on a Gaussian distribution (McNicol, 1972). Our comparisons confirmed the near equivalence of LOR and d' values showing that they are almost precisely directly proportional, having correlations very close to 1 (typically within 1 or 2%). Traditionally a d' value of 1 is used as threshold; therefore, threshold was defined as an LOR of 1.

The major advantage of LOR for analysis is its ease of computation relative to d'. Unlike d', which is typically computed using standardized (z) scores from the normal distribution, the LOR is readily computed using the logit (or log-odds transformation), a calculation analogous to that of deriving a z-score difference, but simpler computationally. Contrasting d', the LOR can easily be calculated from the proportion of hits (H), false alarms (FA), correct rejections (CR), and misses (M) with one formula:

$$\text{LOR} = \ln \{[(H/M)^*(CR/FA)]^{1/2}\}, \qquad (1)$$

Other mathematically equivalent formulations of this equation can be found in the literature using algebraic manipulations and the fact that $H = 1 - M$ and $FA = 1 - CR$ (Macmillan and Creelman, 1991).

To calculate thresholds, LOR values were plotted against noise intensity or gap duration, and the following sigmoidal equation was fit (McNicol, 1972; Macmillan and Creelman, 1991):

$$\text{Predicted LOR} = y_0 + a/\{1 + \exp(-(x_m - x_0)/b)\}, \qquad (2)$$

where x_m is the noise intensity level or gap duration and $a, x_0, y_0,$ and b are free parameters (see Fig. 3). Where these curves crossed an LOR value of 1 was defined as threshold.

Significance testing was conducted using a repeated-measures ANOVA on LOR thresholds determined by fitting Equation 2 to psychometric functions. Thresholds were collected across different sessions; at least four sessions for each experimental condition were used, and they were randomly sampled, when necessary, to balance the samples across experimental conditions. We hypothesized that performance would differ across experimental conditions (e.g., interrupting vs surrounding noise for the intensity experiment) but that subjects would not differ in their performance. Thus the design of the RM-ANOVA incorporated a between-subjects factor of subject (macaque) and a within-subjects factor of experimental condition. This necessitated that we collapse across sessions to analyze the noise conditions because there were insufficient degrees of freedom to add this as a variable into the RM-ANOVA model. For the results in this report we determined that LOR thresholds did not differ by session (RM-ANOVA) and were not a significant predictor of LOR threshold variance (linear regression), although experimental condition was. This confirmed that we had adequately stable performance
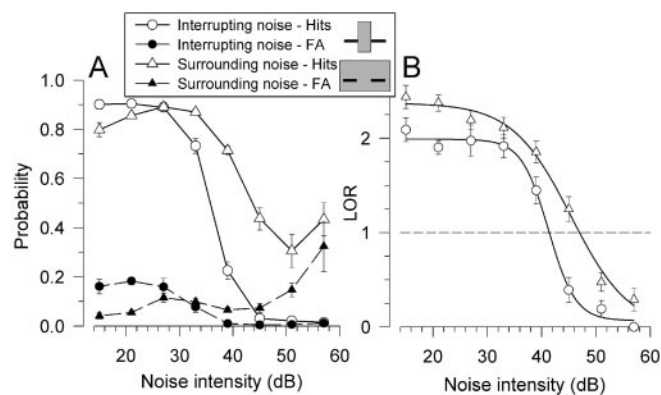
Figure 3. Psychometric functions and threshold determination. Plots show representative noise-intensity experiment psychometric functions from subject X for interrupting (circles) and surrounding (triangles) noise. Spectrograms of target stimuli are schematized in the legend. *A* shows the probability of a hit (solid lines and open symbols) or false alarm (FA) (dashed lines and black-filled symbols) for interrupting and surrounding noise. *B* shows the transformation to a criterion-free measure LOR. Curved lines represent sigmoidal functions (Eq. 2) fit to the points. The dashed line corresponds to the threshold level of LOR = 1. Error bars (SEM) designate the between-session variance in performance.

Figure 4. Box plots of the thresholds of two macaques on the intensity experiment. Noise intensity thresholds for interrupting and surrounding noise are shown using 2 kHz tonal (*A, B*) and coo vocalization (*C, D*) foregrounds. Spectrograms on bottom schematize the target stimuli. Shaded circles represent macaque X and open circles represent Z; horizontal lines bisecting the rectangles represent the mean.

across sessions and justified the use of the subject by experimental condition RM-ANOVA model.
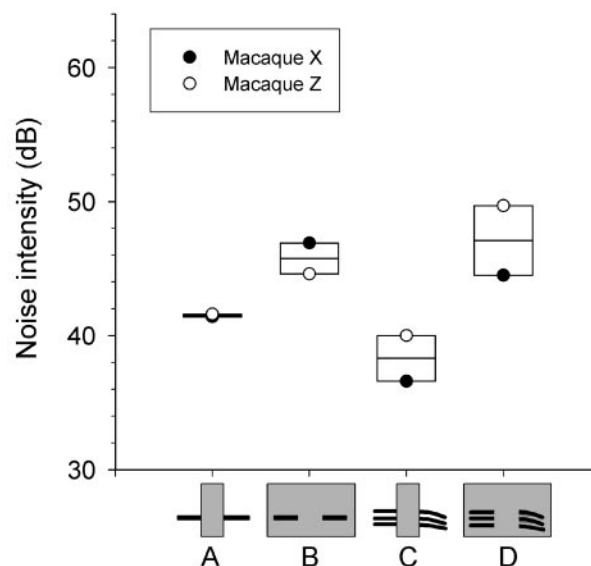
The results did not change when the data for the model were the points on the psychometric function averaged across sessions. This was the case, although the psychometric functions being compared occasionally had few intensity or gap duration values in common (e.g., when performance differed greatly across compared conditions as in the notched-noise or gap duration experiments) and thus reduced sample size.

Results

Intensity experiments

A representative psychometric function from the noise intensity experiments is shown in Figure 3. As expected, for both interrupting and surrounding noise, performance drops as noise intensity increases; however, the manifestation of worsening performance is different for the two types of noise, suggesting differences in how the two different stimuli are perceived (see Discussion). For surrounding noise (Fig. 3, triangles), which causes masking in humans, as noise intensity increases the hit rate decreases, whereas the false alarm rate increases. In other words, for surrounding noise the monkey is more likely to respond to targets by withholding responses at higher compared with lower noise intensities, i.e., to fail to detect a gap at high surrounding noise intensities. For standard trials, however, in which the second sound does not contain a gap, at higher noise intensities the monkey is more likely to respond as if a gap were detected (i.e., to false alarm). These results suggest that, at higher surrounding noise intensities, there is less certainty about the presence of a gap caused by masking and a "guessing" strategy is used. The changes in performance for interrupting noise (Fig. 3, circles), which causes induction in humans, are quite different. For interrupting noise, as noise intensity increases, both hit and false alarm rates decrease. In other words, for intense interrupting noise the monkey is more likely to respond to both gap-containing targets and continuous standards as if they were continuous, consistent with induction of the missing tone segment.

Consistent with the example of Figure 3*B* and human results, macaque performance was worse for interrupting noise that causes auditory induction than for masking surrounding noise. This is manifested by lower intensity thresholds for interrupting noise (Fig. 3*B*, right). In this case, a worse threshold corresponds

to a lower intensity because performance drops with increasing noise intensity and therefore performance drops below threshold (Fig. 3*B*, dashed line at LOR = 1) at a lower intensity interrupting than surrounding noise, showing that a less intense interrupting noise disrupted continuity detection more. This result held for both tonal and vocalization foregrounds (Fig. 4). For tonal foregrounds the interrupting noise thresholds (Fig. 4*A*) on average were 4.1 dB worse than for surrounding noise (Fig. 4*B*). This effect was significant ($F_{(1,14)} = 23.07$; $p < 0.001$). For the coo vocalization, interrupting noise thresholds were 8.0 dB worse than surrounding noise. This effect (Fig. 4*C,D*) was also significant ($F_{(1,6)} = 29.35$; $p < 0.01$).

Differences in performance between the two monkeys (Fig. 4*A*, open and filled circles) were small, particularly for tonal foregrounds (note how close the two data points are in each plot). Neither intersubject nor intersession differences in performance reached significance.

Because our subjective experience was that induction was more pronounced for speech and vocalizations than tones, we hypothesized that interrupting noise may affect macaque performance more with the coo than with the 2 kHz tone. The result that masking thresholds were similar (Fig. 4*B,D*) but interrupting noise thresholds were worse with coos (Fig. 4*C*) than tonal foregrounds (Fig. 4*A*) is consistent with this notion. The difference in thresholds between coo and tonal foregrounds with masking surrounding noise (Fig. 4*B,D*) did not reach significance, but the difference between thresholds with coo and tonal foregrounds using interrupting noise (Fig. 4*A,C*) did ($F_{(1,6)} = 5.24$; $p < 0.05$).

Because induction of a tone is weaker in humans when frequencies in common with the tone are removed or "notched" from interrupting noise (Warren et al., 1972), we tested for a similar relationship in monkeys. This is also important in controlling for changes in performance that might be introduced by response biases caused by the saliency of an intense noise if that noise would disrupt performance. If performance were recovered
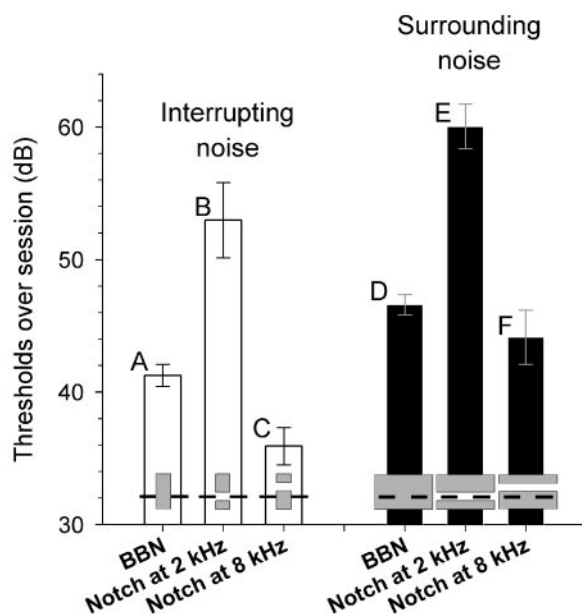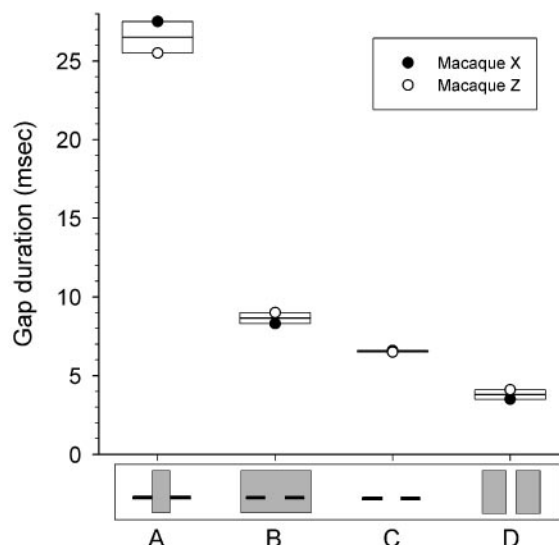
Figure 5. Intensity experiment thresholds (over session) using notched-noise with 2 kHz tonal foregrounds (subject X only). White bars depict interrupting noise results; black bars depict surrounding noise results. Schematized spectrograms embedded in the bars show the type of noise used—broadband noise (BBN), BBN with a Notch at 2 or 8 kHz—relative to target tones. Error bars designate SE of the mean between sessions.
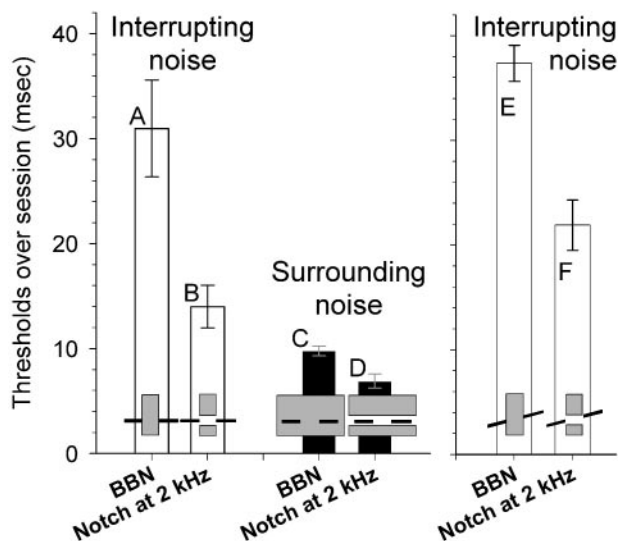

Figure 6. Gap duration experiment thresholds. Thresholds for interrupting (A), surrounding (B), and no noise (C) with 2 kHz foregrounds. D, Gap in broadband noise thresholds. Schematized spectrograms of the stimuli are shown on bottom. Format is the same as Figure 4.

by notching the noise around the frequency of the tone, this would argue against the saliency of intense noise causing the drop in performance. We also used noise notched at 8 kHz, but energetically matched to the notch at 2 kHz, to control for the possibility that a notch by itself would influence performance. The threshold of macaque X improved by 11.7 dB when the unnotched interrupting noise was notched at 2 kHz and by 13.4 dB when the surrounding noise was notched at 2 kHz (Fig. 5). These improvements were significant (interrupting noise: $F_{(1,3)} = 12.80$, $p < 0.05$; surrounding noise: $F_{(1,4)} = 158.79$, $p < 0.001$) and ruled out the possibility that the saliency of intense noise was causing the drop in performance. Performance dropped to below broadband noise levels when the notch was moved to 8 kHz (Fig. 5); the difference between broadband noise and broadband noise with an 8 kHz notch (Fig. 5) did not reach significance for interrupting or surrounding noise using the RM-ANOVA analysis ($p = 0.188$ for interrupting noise; $p = 0.673$ for surrounding noise); however, the small error bars suggest an effect for interrupting noise. The observation with respect to the error bars is supported by an independent samples t test, which found that interrupting broadband noise and 8 kHz notched noise performances were significantly different ($p < 0.01$). Finally, it should be noted that performance was still worse for interrupting (Fig. 5B) than surrounding (Fig. 5E) noise with a 2 kHz notch ($F_{(1,3)} = 11.67$; $p < 0.05$) (see Discussion).

Gap duration experiments

Using a 2 kHz single frequency tone and a linear FM ramp, these experiments varied the gap duration (and interrupting noise duration, when present) with the noise intensity (when present) fixed at 40 dB. Tones and FMs were 45 dB. We expected that a longer gap would be required for detection with interrupting (inducing) than surrounding (masking) noise, as in human studies. Results were consistent with this prediction. For noise-interrupted tonal foregrounds, only gaps >26.5 msec mean du-

ration could be reliably detected (Fig. 6A). For surrounding noise, gaps as short as 8.7 msec mean duration could be detected (Fig. 6B). Gaps in tones that were between 8.7 and 26.5 msec long could be detected with surrounding noise but could not be detected with interrupting noise. The differences between interrupting and surrounding noise were significant ($F_{(1,10)} = 18.55$; $p < 0.01$).

We also measured the ability of the macaques to detect gaps in tones when noise was not present (Fig. 6C). The mean threshold of 6.6 msec for detecting gaps in tones without noise (Fig. 6C) was shorter than the 8.7 msec threshold for surrounding noise (Fig. 6B). The difference, although small, was significant ($F_{(1,10)} = 20.01$; $p < 0.01$). We also measured the ability of macaques to detect gaps in noise foregrounds (Fig. 6D). This was done to allow for comparisons with standard techniques measuring the temporal resolution of the auditory system in humans and other animals. Macaque thresholds for this form of gap detection, detecting gaps in broadband noise, was 3.8 msec, which was significantly better than the ability to detect gaps in tones ($F_{(1,10)} = 149.13$; $p < 0.001$).

Similar results were obtained using an FM foreground with macaque X. Experiments with FM foregrounds also yielded significant differences between detection thresholds for silent gaps with interrupting (37.4 ± 1.7 msec) and surrounding (6.4 ± 0.1 msec) noise ($F_{(1,5)} = 264.3$; $p < 0.001$); however, the ability to detect gaps in FMs without noise (7.2 ± 0.5 msec) was not significantly different from the ability to detect gaps in surrounding noise.

We also investigated how spectrally notched noise affected gap duration thresholds. Using broadband noise with a 1 octave notch at 2 kHz lowered the gap duration thresholds and improved performance for interrupting and surrounding noise (Fig. 7). For interrupting noise with tonal foregrounds (Fig. 7A,B), thresholds were 17.0 msec better for notched than unnotched noise; this effect only approached significance ($F_{(1,3)} = 7.64$; $p = 0.07$). For surrounding noise with tonal foregrounds (Fig. 7C,D), thresholds were 2.9 msec better for notched than unnotched noise, but this effect was not significant. Finally, for interrupting noise with FM foregrounds (Fig. 7E,F), thresholds

Figure 7. Gap duration experiment thresholds (over session) with notched-noise (subject X only). Format is as in Figure 5.

were 15.5 msec better for notched than unnotched noise, an effect that was significant ($F_{(1,3)} = 41.33$; $p < 0.05$).

Discussion

Evidence that auditory induction occurs in animals

Although induction has been characterized most fully in humans, only two papers, to our knowledge, address auditory induction in other animals. Our data using standard psychophysical techniques, when combined with the other published behavioral data, strongly support the notion that induction is a general property of auditory systems.

Sugita (1997) showed in cats that detecting a gap in an FM glide was more difficult as interrupting noise intensity increased; however, this was limited to one interrupting noise condition, and important controls for the timing and spectrum of the noise and for masking were not presented in that short report. Therefore, because it has been shown that only intense "interrupting" noise impaired performance, there is only weak evidence for induction in cats.

Stronger evidence for induction was provided in behavioral studies of cotton-top tamarin monkeys (Miller et al., 2001). Tamarins vocalize in response to species-specific vocalizations (calls) but not, generally, to other sounds. Miller et al. (2001) found that tamarins tended not to vocalize in response to calls with silent gaps but did when the gaps were filled with noise. This is consistent with the noise allowing for the missing information to be restored. Importantly, they also manipulated the timing of the noise relative to the gap to show that when the noise was not temporally superimposed into the gap, the vocalization rate of tamarins decreased. This demonstrated that the monkey vocalized under conditions that elicit induction and ruled out the possibility that the intense noise by itself caused the animals to vocalize. These studies cleverly exploited a natural behavior to tap into properties of the auditory system, but the techniques are not amenable for making direct comparisons with humans.

The results presented here extend the evidence of auditory induction in animals by incorporating control experiments for the structure and timing of the noise and by using a psychophysical approach similar to that used in humans (Warren et al., 1972; Kluender and Jenison, 1992). Our data further support the occurrence of induction in monkeys because stimuli that manipu-

late thresholds related to masking and induction in humans also similarly change thresholds in macaques. Interrupting noise, known to elicit auditory induction in humans, makes gap detection more difficult (Warren et al., 1972; Bregman, 1990) than surrounding noise, which causes masking (Kluender and Jenison, 1992). By manipulating the temporal parameters of inducing noise and noise designed to cause masking, similar differences between induction and masking are found for speech recognition in humans (Stuart et al., 1995; Stuart and Phillips, 1997); that is, at similar noise intensities, word recognition is better with interrupting noise because of induction than with masking noise. Differences in macaque thresholds for detecting discontinuities in interrupting and surrounding noise are consistent with human differences and support the notion that induction occurs with interrupting noise in macaques.

Results relating to false alarm rates also support auditory induction in macaques. Macaque hit and false alarm rates both decline with increasing interrupting noise intensity (Fig. 3A), demonstrating that macaques respond to stimuli both with and without gaps as if they were continuous. This supports the interpretation that the macaques have high confidence that the second presented (target) foreground was the same as the first presented continuous (standard) foreground even when there was a gap in the target. This indicates that induction and perceptual restoration of the tone occur during the noise-filled gap. Alternatively, when surrounding noise was used the false alarm rate increased and the hit rate decreased with increasing noise intensity (Fig. 3A). This suggests that for surrounding noise, macaques have less confidence in reporting what occurred during the gap because of masking and, accordingly, used a guessing strategy.

A similar result has been reported in humans. Kluender and Jenison (1992) asked subjects to detect the presence of a continuous foreground FM glide in interrupting or surrounding noise. Subjects tended to false alarm more to interrupting than surrounding noise. Because their subjects were asked to respond to continuity, this was suggested to reflect induction for interrupting noise because higher false-alarm rates (more reports of continuity) were observed for interrupting than surrounding noise. These results might seem different from ours because in our study false-alarm rate decreased for interrupting noise; however, because our subjects reported discontinuity, this lower false alarm rate corresponds to less detection of discontinuity and therefore more detection of continuity, consistent with induction.

Differences in thresholds caused by introducing spectral notches in noise

Warren et al. (1972) investigated how the frequency (spectral) content of noise affected induction and masking in humans. When bandpassed noise was introduced with a frequency range outside of the frequency of a foreground tone, both masking and induction were less pronounced than when noise energy was present at the frequency of the tone. Similarly, when a spectral notch was placed in a noise and the notch was centered at the frequency of the foreground tone, masking and induction thresholds for the tone were reduced in fairly equal amounts, so that induction still occurred at lower noise intensities than masking. Our results with macaques show similar threshold shifts for notched interrupting (Fig. 5B) and masking (Fig. 5E) noise. Also, consistent with human results, macaque performance was worse for interrupting (Fig. 5B) than surrounding (Fig. 5E) noise with a 2 kHz notch. Once again the strong similarity in how macaque

and human thresholds change when the noise is manipulated suggests similar mechanisms.

Another interesting aspect of the results was that the 8 kHz notch in interrupting noise appeared to enhance the induction effect when compared with broadband noise. Although speculative because of the conflicting statistical results, if true this enhancement would be interesting because it suggests that the 8 kHz notch is providing some kind of release from lateral inhibition in the 2 kHz band, thereby making the noise more effective in the 2 kHz band of interest.

Induction for different types of foregrounds

To our knowledge there have been no direct comparisons of induction thresholds with different types of foregrounds. Because our subjective experience was of stronger induction for vocalizations than tones, we tested whether macaque thresholds support stronger induction for coo vocalizations. This indeed was the case. Because thresholds for surrounding noise presented with the two foregrounds were roughly the same (Fig. 4*B*,*D*), this supports the notion that more complex or naturalistic foregrounds result in equal masking; however, the worse (lower intensity) thresholds for vocalizations (compared with tones) with interrupting noise support stronger induction for the vocalizations (Fig. 4*A*,*C*). One explanation is that sounds with higher levels of complexity (variation along multiple parameters) and more redundancy across these parameters are stronger inducers.

Mechanisms of auditory induction

Induction, a process that follows rules of auditory scene analysis (Bregman, 1990), is remarkable because of its illusory nature. Such phenomena provide an excellent opportunity to study perception and brain mechanisms for analyzing complex sounds in naturalistic settings. An illusion creates a point of departure between the physical form of a sound and its perception; therefore, illusory phenomena can help to determine where in the brain activity is more closely related to the physical sound and where it more closely follows the percept. This opportunity to "trick" the brain into responding to an illusory stimulus feature can help us to determine the neural mechanisms underlying the percept of that stimulus as it is transformed from a veridical representation of the world to an illusory one shaped by the rules that the system uses to interpret the world.

Although we have gained insight into induction from the innovative work in tamarins (Miller et al., 2001), the call-back paradigm is not amenable to future physiological recording because single-neuron responses to sounds cannot be interpreted if the animal is simultaneously vocalizing. Also less is known about whether tamarins are suitable for physiological study. Recording single neurons from macaques while they perform tasks like the one used in our study is a well established technique (Evarts, 1968; Mountcastle et al., 1972) that has provided tremendous insight into visual perception, including mechanisms underlying illusory motion (Assad and Maunsell, 1995); therefore, this psychophysical task in macaques will provide an excellent animal model for studying the neural mechanisms of auditory perception.

Summary

In conclusion, we performed two independent experiments in macaques with multiple manipulations of noise to investigate auditory induction, controlling for masking. Results were consistent with published studies of induction in humans. We also report macaque gap detection thresholds using gaps in broadband noise that are comparable with those reported for humans (Penner, 1977) and other species (Fay, 1988). In combination with the published work, our results provide support for the hypothesis that animals share similar mechanisms for auditory induction, a process that may act in the formation of stable representations of sounds encountered in acoustically complex environments.

References

Assad JA, Maunsell JH (1995) Neuronal correlates of inferred motion in primate posterior parietal cortex. Nature 373:518–521.

Bennett KB, Parasuraman R, Howard Jr JH, O'Toole AJ (1984) Auditory induction of discrete tones in signal detection tasks. Percept Psychophys 35:570–578.

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: MIT.

Ciocca V, Bregman AS (1987) Perceived continuity of gliding and steady-state tones through interrupting noise. Percept Psychophys 42:476–484.

Dannenbring GL, Bregman AS (1976) Effect of silence between tones on auditory stream segregation. J Acoust Soc Am 59:987–989.

Evarts EV (1968) Relation of pyramidal tract activity to force exerted during voluntary movement. J Neurophysiol 31:14–27.

Fay RR (1988) Hearing in vertebrates: a psychophysics data book. Winnetka, IL: Hill-Fay Associates.

Green DM, Swets JA (1974) Signal detection theory and psychophysics. Huntington, NY: Robert E. Krieger.

Hauser MD (1993) Food-associated calls in rhesus macaques (*Macaca mulatta*). I. Socioecological factors influencing call production. Behav Ecol 4:194–205.

Kluender KR, Jenison RL (1992) Effects of glide slope, noise intensity, and noise duration on the extrapolation of FM glides through noise. Percept Psychophys 51:231–238.

Luce RD, Bush RR, Galanter E (1963) Handbook of mathematical psychology. New York: Wiley.

Macmillan NA, Creelman CD (1991) Detection theory: a user's guide. Cambridge, UK: Cambridge UP.

McNicol D (1972) A primer of signal detection theory. London: Allen and Unwin.

Miller CT, Dibble E, Hauser MD (2001) Amodal completion of acoustic signals by a nonhuman primate. Nat Neurosci 4:783–784.

Mountcastle VB, LaMotte RH, Carli G (1972) Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. J Neurophysiol 35:122–136.

Penner MJ (1977) Detection of temporal gaps in noise as a measure of the decay of auditory sensation. J Acoust Soc Am 61:552–557.

Rossi AF, Paradiso MA (1999) Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. J Neurosci 19:6145–6156.

Stuart A, Phillips DP (1997) Word recognition in continuous noise, interrupted noise, and in quiet by normal-hearing listeners at two sensation levels. Scand Audiol 26:112–116.

Stuart A, Phillips DP, Green WB (1995) Word recognition performance in continuous and interrupted broad-band noise by normal-hearing and simulated hearing-impaired listeners. Am J Otolaryngol 16:658–663.

Sugita Y (1997) Neuronal correlates of auditory induction in the cat cortex. NeuroReport 8:1155–1159.

Warren RM (1970) Perceptual restoration of missing speech sounds. Science 167:392–393.

Warren RM, Obusek CJ, Ackroff JM (1972) Auditory induction: perceptual synthesis of absent sounds. Science 176:1149–1151.

Warren RM, Wrightson JM, Puretz J (1988) Illusory continuity of tonal and infratonal periodic sounds. J Acoust Soc Am 84:1338–1342.