Published in final edited form as:

Eur J Neurosci. 2013 October ; 38(7): 3056–3070. doi:10.1111/ejn.12325.

The inferior colliculus encodes the Franssen auditory spatial illusion

Abigail Z. Rajala¹, Yonghe Yan², Micheal L. Dent³, and Luis C. Populin^{1,2,4,5}

¹Neuroscience Training Program, University of Wisconsin-Madison, Madison, WI, USA

²Department of Neuroscience, B385 MSC, University of Wisconsin-Madison, 1300 University Ave, Madison, WI 53706, USA

³Department of Psychology, University of Buffalo, The State University of New York, Buffalo, NY, USA

⁴Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

⁵McPherson Eye Research Institute, University of Wisconsin-Madison, Madison, WI, USA

Abstract

Illusions are effective tools for the study of the neural mechanisms underlying perception because neural responses can be correlated to the physical properties of stimuli and the subject's perceptions. The Franssen illusion (FI) is an auditory spatial illusion evoked by presenting a transient, abrupt tone and a slowly rising, sustained tone of the same frequency simultaneously on opposite sides of the subject. Perception of the FI consists of hearing a single sound, the sustained tone, on the side that the transient was presented. Both subcortical and cortical mechanisms for the FI have been proposed, but, to date, there is no direct evidence for either. The data show that humans and rhesus monkeys perceive the FI similarly. Recordings were taken from single units of the inferior colliculus in the monkey while they indicated the perceived location of sound sources with their gaze. The results show that the transient component of the Franssen stimulus, with a shorter first spike latency and higher discharge rate than the sustained tone, encodes the perception of sound location. Furthermore, the persistent erroneous perception of the sustained stimulus location is due to continued excitation of the same neurons, first activated by the transient, by the sustained stimulus without location information. These results demonstrate for the first time, on a trial-by-trial basis, a correlation between perception of an auditory spatial illusion and a subcortical physiological substrate.

Keywords

auditory spatial illusion; inferior colliculus; monkey

Introduction

Illusions, distortions of objective perception of some physical aspect of stimuli, are useful tools for the study of the mechanisms underlying perception because one can correlate neural responses to both the physical properties of the stimuli and the reports of the subjects that perceive them (Logothetis, 1998). Visual illusions are the most widely known and studied; however, much less is known about spatial auditory illusions and the mechanisms that underlie them.

The Franssen illusion (FI; Franssen, 1962) is an example of the latter. It is evoked by simultaneously presenting two pure tones of the same frequency, one transient and abrupt and the other slowly rising and sustained, from transducers located to the left and right of the subject in a reverberant room, and it consists of hearing a single sound, of long duration, on the side where the transient tone was presented. Perception of the illusion (Hartmann & Rakerd, 1989; Yost *et al.*, 1997), which has also been demonstrated in monkeys (Populin, 2006), cats (Dent *et al.*, 2004) and birds (Dent *et al.*, 2007), may last tens of seconds (Berkley, 1983).

It has been proposed that a subcortical mechanism, which discards implausible interaural cues that result from single tones without sharp onset/offset presented in an echoic environment in favor of plausible cues generated by the transient component of the Franssen stimuli (FS; Hartmann & Rakerd, 1989; Rakerd & Hartmann, 1986), is responsible for perception of the FI. However, mismatch negativity data suggest that such a mechanism resides in the auditory cortex (Yagcioglu & Ungan, 2006).

Perception of the FI, regardless of the locus of the underlying mechanism, involves the erroneous processing of auditory cues, interaural differences of time and level, used for horizontal sound localization (Stevens & Newman, 1936). Psychophysical evidence obtained under dichotic and free-field conditions, without acoustic reflections, indicates that humans do lateralize (Haftner *et al.*, 1979) and localize (Stevens & Newman, 1934, 1936) single tones without abrupt onsets. Furthermore, physiological recordings from single units of anesthetized and awake animals have revealed that such cues are encoded at the level of the inferior colliculus (IC), with a strong bias towards sound sources located contralaterally (Kuwada & Yin, 1983; Kuwada *et al.*, 1987; Semple & Kitzes, 1987; Irvine & Gago, 1990; Batra *et al.*, 1993).

This study first demonstrates that monkeys and humans orient to FS similarly, indicating that they both perceive the FI. Subsequently, we tested the hypothesis that the effectiveness of spatial cues in the transient component, but not the sustained component, of the FS determines the perception of the FI by recording from single IC units while the monkeys indicated perceived sound location with their gaze. Receiver operating characteristic (ROC) analyses of discharge rate based on the subjects' reports support this hypothesis, revealing that the transient component encodes the perception of sound location. Furthermore, they reveal that the persistent erroneous perception of the sustained stimulus results from continued excitation of the same neurons, first activated by the transient, by the sustained

stimulus lacking spatial information. Thus, a neural mechanism able to account for the behavioral responses to the FS by both humans and monkeys is found in the IC.

Materials and methods

Both human and monkey subjects were tested in this study. The experiments were carried out in different but similarly configured experimental rooms. Despite our efforts, however, some differences between the experimental setups may have contributed to differences in the extent to which the FI was perceived by both species; for example, monkeys sat in a plexiglass primate chair during experiments, which constitutes a large reflective surface, while the humans did not.

Because the FI is more reliably perceived in reverberant environments (Hartmann & Rakerd, 1989; Yost *et al.*, 1997; Dent *et al.*, 2004), we purposefully arranged the acoustic conditions of the experimental rooms to create semi-reverberant conditions, which resulted in reports of FI perception in 50–75% of trials. In this manner we were able to compare the responses of IC neurons between trials in which monkeys reported hearing and not hearing the illusion. We chose this approach because changing the acoustic conditions of the experimental room from reverberant to non-reverberant (or vice versa) while maintaining single unit recordings in the head-unrestrained animal would have been impractical. More details of the methods are given in Populin (2006, 2008).

Human experiments

Subjects—Two females, 29 and 33 years old, who participated in a psychophysical study of sound localization that used gaze to indicate perceived sound location (Populin, 2008) were tested for the perception of the FI. Before each experimental session, for which they were compensated monetarily, subjects were asked to read information about the experiment and the risks associated with wearing a scleral search coil, and were asked to provide written consent. Upon obtaining written consent, an ophthalmologist applied a drop of topical ophthalmic anesthetic (Novesine; oxybuprocaine, 0.4%) to the right eye and inserted the coil in the eye 5 min later. While waiting for the anesthetic to take effect, verbal instructions were provided about the experimental procedures, including the requirements of the memory-guided saccade. Specifically, subjects were instructed to ‘look at the sources of sounds’. All experimental procedures were approved by the University of Wisconsin Institutional Review Board (IRB), and conformed with The Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the British Medical Journal (18 July, 1964).

Experimental setup and eye and head movement measurements—The experiments were carried out in a ($3 \times 3 \times 2$ m) double-walled soundproof chamber (Acoustic systems, Austin, TX, USA). Gaze and head position were measured with the magnetic search coil technique (Robinson, 1963) using a phase angle system (CNC Engineering, Seattle, WA, USA). Fine wire coils embedded in silicon annuli (Skalar Instruments, Delft, the Netherlands) were used to measure gaze position. Measurements of head position, which were obtained with a similar coil attached to the head with a lightweight (104 g) plastic holder (Populin, 2008), were used to confirm that the head of the

subjects was aligned with the straight ahead position during the presentation of the stimuli in all trials. Signals were low-pass filtered at 250 Hz (Krohn-Hite, Brockton, MA, USA), digitally sampled at 500 Hz with an analog to digital converter (System 2; Tucker Davis Technologies, Alachua, FL, USA), and stored for off line analysis.

Experimental sessions and tasks—Each session consisted of a random mixture of trials using the fixation, standard- and memory-guided saccade tasks (Populin, 2008) using broadband stimuli, FS, and its single tone components. The memory-guided saccade task was chosen to expose the subjects to the entire duration of the FS (Fig. 1) and its components before responding. A trial started with the presentation of a visual stimulus, a red light-emitting diode (LED), at the straight ahead position, which the subjects were expected to fixate. After a variable period of fixation (300–600 ms), the auditory stimuli were presented, during which time the subjects were expected to maintain fixation on the LED straight ahead and not move their heads for 500 ms. At the end of this period the fixation LED was turned off, requiring the subject to orient to the perceived location of the sound. For the FS and the sustained component presented alone, both of which lasted 500 ms, the offset of the sound and fixation LED coincided. For the transient component presented alone, which lasted only 50 ms, the subjects were required to maintain fixation for the entire 500 ms, i.e. no auditory stimulus was presented during the last 450 ms. Trials with the FS were presented in ~20–25% of the total number of trials in a given session and no feedback was provided after each trial, unlike in the monkey experiments, or after the experimental sessions. As recommended by the manufacturer of the coils, the experimental sessions lasted 30 min. Accordingly, the data were collected in 3–5 experimental sessions per subject.

Monkey experiments

Subjects and surgery—Two adult male macaque monkeys (*Macaca mullata*) were used in this study. Each underwent at least two sterile surgical procedures. In the first, scleral search coils (Judge *et al.*, 1980) and a titanium head post, used to restrain the animal's head to care for the implant area and to hold a spout in front of the monkeys' lips to deliver water as reward during the experimental sessions, were implanted. The coils were constructed of stainless steel wire (SA632; Cooner Wire Co., Chatsworth, CA, USA). In the second surgery, a 19-mm-diameter craniotomy was performed and a recording cylinder implanted to reach the tectum with recording microelectrodes. The cylinder was mounted on the midline, tilted posteriorly 38° relative to the coronal plane. All surgical procedures were approved by the University of Wisconsin Animal Care Committee and were in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*.

Eye and head movement recordings—Gaze and head position were recorded with the magnetic search coil technique (Robinson, 1963) using a phase angle system (CNC Engineering). The horizontal and vertical channels representing eye and head position were low-pass filtered at 250 Hz (Krohn-Hite), digitally sampled at 500 Hz with an analog to digital converter (System 2) and stored for off line analysis.

Experimental sessions and tasks—Physiological recording sessions consisted of 50–60% of FI trials, compared with 20–25% in behavioral sessions, in order to maximize the number of trials comprising the perception of the FI given the difficulties of maintaining single unit recordings from the IC under the head-unrestrained condition, which is required for accurate sound localization in both monkeys (Populin, 2006) and cats (Sparks, 2005; Tollin *et al.*, 2005). Upon isolating an IC neuron and determining its characteristic frequency (see Fig. 3), a mixture of trials with FS and its individual components alone were presented from two transducers at $\pm (30^\circ, 0^\circ)$. In addition, trials with visual targets were presented from 11 locations within the oculomotor range using the standard saccade task (Populin, 2008; Fig. 1A) and were randomly intermixed among the auditory trials to maintain the subjects' interest in participating. The standard saccade task consists of the presentation of a fixation point straight ahead, followed by the presentation of a target, acoustic or visual, elsewhere at the time of fixation offset. A $(3^\circ, 3^\circ)$ error window was set around the fixation point in all trials. For single source stimuli, an error window of $(8^\circ, 8^\circ)$ was used for the acoustic targets and a $(4^\circ, 4^\circ)$ window for the visual targets.

Subjects were rewarded with a small amount of water after accurate orienting to visual targets and the sources of the individual components of the FS presented alone, and after all trials with FS, regardless of the subjects' responses, as correct/incorrect criteria could not be defined for such trials. If no rewards had been presented during FS trials, the animals would have stopped participating in the experiment given that a large percentage of trials in the session were of this type. Trials in which the subjects responded before the fixation LED was turned off were immediately terminated.

Visual and acoustic stimuli (human and monkey experiments)

The visual fixation point, presented straight ahead to stabilize gaze and head position before the presentation of acoustic stimuli, and the visual targets used in the monkey experiments consisted of red LEDs that subtended a 0.2° visual angle. Acoustic stimuli were generated with Tucker Davis Technologies System 3 hardware and presented with Morel MDT-20 28-mm soft dome tweeters (Morel America, Elmont, NY, USA) using separate hardware channels under the control of a custom switch box designed to present acoustic stimuli with microsecond accuracy without introducing unwanted transient noise at the time of speaker selection.

The FS consisted of pairs of pure tones presented from speakers located on the horizontal plane to the left and to the right of the subject (Fig. 1), at $\pm (60^\circ, 0^\circ)$ for the humans and at $\pm (30^\circ, 0^\circ)$ for the monkeys. For the initial experiments, which were carried out with the human subjects, FS were presented from more eccentric positions to facilitate localization of the actual sources of sounds and thus minimize the likelihood of eliciting the perception of the FI. We learned, however, that perception of the FI was very strong, even at such eccentric positions, and therefore in the monkey experiments transducers located at less eccentric positions, to which they could orient more easily, were used. This difference in transducer position between the human and monkey experiments may, nevertheless, have contributed to the overall differences in the proportion of trials in which the FI was perceived (Fig. 1G and K).

The transient tone consisted of a 50-ms single frequency waveform (1–11 kHz) with an abrupt start, i.e. a rise of 0 ms, and a linear fall of 50 ms, and the sustained tone consisted of a 500-ms single frequency waveform with a 50-ms linear rise and a 100-ms linear fall. For the human experiments acoustic stimuli were presented at 50 dB SPL and for the monkey experiments at 55–65 dB SPL, depending on the sensitivity of the single unit under study. Although we did not systematically study the difference between neural and behavioral thresholds, we found that in general neural thresholds in the IC were lower than behavioral thresholds, and stimuli presented at neural thresholds caused the monkeys to respond ineffectively and inconsistently, or not at all. Therefore, we had to present stimuli at a level high enough to engage our monkey subjects. The background sound levels in both experimental rooms was 27–30 dB SPL (Extech 407740; Extech, Waltham, MA, USA).

Single unit recordings

Tungsten microelectrodes (Fine Science Tools, Foster City, CA, USA, or Microprobe, Potomac, MD, USA) were used to record extracellularly from single IC units. The electrodes were inserted into the brain using custom made stainless steel guide tubes (23 gauge) held in place by a custom made acrylic grid, and a micromanipulator (MO-95; Narishige, Tokyo, Japan) attached to the recording cylinder. This procedure was carried out under head-restrained conditions. The IC was identified physiologically by the low-threshold responses to broadband acoustic stimuli, the tonotopic pattern of responses to single-frequency tones encountered as the electrode was lowered, and the location of the auditory evoked discharges relative to those of saccade burst neurons in the adjacent superior colliculus (Zwiers *et al.*, 2004).

The neural signals were amplified (Bak Electronics, Mount Airy, MD, USA) and band pass filtered (0.6–3.0 kHz; Krohn-Hite). The occurrence of action potentials was documented with microsecond accuracy by recording the timing of TTL pulses produced by a window discriminator (Tucker Davis Technologies). All data presented were collected from unequivocally well-isolated single units.

Search for single units and characterization of best frequency—Physiological recordings started by placing the tip of the microelectrode in the vicinity of the IC, the auditory evoked activity of which could be heard on the audio monitor, then releasing the head of the animal to start the recording session, which consisted of a mixture of auditory and visual trials presented in random order.

The search for single units was carried out using broadband stimuli presented from up to 11 transducers in the frontal hemifield using the auditory saccade task (Populin, 2006). Upon isolating a single unit, the experimental task was changed to determine the characteristic frequency of the unit under study. A fixation LED was turned on straight ahead and single-tone stimuli ranging in frequency between 1 and 12 kHz in steps of 1 kHz were presented from a speaker located contralateral to the recording side ($\pm 30^\circ, 0^\circ$). Each 500-ms tone pip was presented with the same rise/fall profile used for the sustained component of the FS (50-ms linear rise and 100-ms linear fall) with an interstimulus interval of 500 ms. The animals were rewarded for maintaining fixation during the presentation of the tones. This

was repeated 10–15 times and a tuning function was constructed online by computing the average discharge rate over the 500-ms time of presentation of the stimulus. The characteristic frequency of a neuron determined in this manner was then used for the FS and single source components during testing.

Data analysis

Behavioral data (humans and monkeys)—Final gaze position was used to determine the percentage of trials in which the subject oriented to the transient component of the FS, indicating that they perceived the illusion. It was also used to determine the number of correct and incorrect gaze shifts directed to the single components of the FS presented alone. Final gaze position was measured with custom software graphics written in Matlab (Mathworks, Natick, MA, USA) using a velocity criterion (Populin & Yin, 1998; Populin, 2006). Briefly, the start of a gaze shift was defined as the time when the magnitude of the gaze velocity signal exceeded two standard deviations (SD) of the mean velocity computed for an epoch starting 100 ms before the onset of the stimuli to 30 ms after its onset, during which time gaze was expected to be stationary. The end of a gaze shift, on the other hand, was marked by the time at which gaze velocity returned to within 2 SD of the mean velocity. For FS trials, gaze shifts greater than 5° away from the midline were considered consistent with perception of the FI if they were directed toward the source of the transient component of the stimuli, and inconsistent with the perception of the FI if they were directed toward the source of the sustained component.

Single unit data—The dependent variables for the neural data were first spike latency and average spike rate. First spike latency was computed for each trial directly from the spike trains using a window that started 5 ms after the time of stimulus presentation and ended at 50 ms. No trials were found with latency greater than 50 ms. Average rate was computed for each neuron using a window of 5–500 ms for the sustained and FS, and 5–50 ms for the transient stimuli.

ROC analysis—Receiver operating characteristic analyses were performed to determine the probability that an ideal observer could discriminate between the neural discharges of a population of IC neurons under different conditions using their discharge rate. To determine the time course of response selectivity, a sliding ROC analysis, in which the ROC curve was compared between two conditions, was performed for 50-ms epochs (centered around the time point) and repeated in 1-ms increments over the 500 ms of the stimulus (Everling & DeSouza, 2005). An ROC time course, plotting the area of the ROC curve at each time point, was computed for each neuron and then averaged across all neurons. The arrangement for each of the analyses, specifying what was compared in each and what ROC values > 0.5 represent, is stated in the text. To determine if the average ROC values for each comparison differed significantly from chance at any time point, a bootstrap analysis was performed. For each neuron, the two conditions being compared were randomly assigned (50% probability of being exchanged). This analysis procedure was performed 1000 times resulting in a single average ROC function. The resulting distribution of the 1000 average ROC values at each time point was used to compute the 5th and 95th percentile values, which represent the 5% significance criterion (Johnston *et al.*, 2009).

Results

Behavioral responses to FS

The presentation of two separate sounds of the same frequency, one abrupt and transient, the other slowly rising and sustained, on each side of the subject may result in perception of the FI. The experimental arrangement used to study responses of both humans and monkeys to FS is illustrated in Fig. 1A–C. In this example, showing an FS trial with the transient component presented from a speaker to the left of the subject (Fig. 1B), perception of the FI would have been indicated by a gaze shift to the left (Fig. 1C). Rhesus monkeys have been shown to orient to the source of the transient component of the FS, suggesting that they hear the sustained tone in the hemifield opposite to its presentation (Populin, 2006) and thus perceive the FI. To determine if this behavioral response to the FS could be indicative of perception of the FI, two humans were tested under experimental conditions that closely resembled those in the monkey experiments.

Human—The human subjects were verbally instructed to maintain fixation on an LED at the straight ahead position while lit and to ‘look to the sources of sounds’ after the fixation LED was turned off (Populin, 2008). To avoid introducing biases, subjects were explicitly not informed that two different types of acoustic stimuli, single-source broadband noise (Populin, 2008) and FS, would be used in the experiments.

Figure 1D and E show the configuration of the experimental task and the oculomotor behavior of one of the subjects. The offset of the fixation LED (vertical broken line), the signal for the subjects to respond, coincided with the end of the 500-ms sustained tone. This task configuration, which was used with broadband and FS in randomly selected trials, compelled the subjects to listen to the entire stimulus before making a response. Perception of the illusion was inferred from the subjects orienting to the source of the transient component of the FS (Dent *et al.*, 2004), which is illustrated by the gaze traces in magenta. Gaze traces directed to the sustained component of the stimuli, denoting trials in which the FI was not perceived, are plotted in black.

The human results are consistent with those from monkeys tested previously with identical stimuli (Populin, 2006). The subjects oriented to the perceived location of the sound sources with short latency and purposive gaze shifts, which, regardless of the component of the stimuli to which they were directed, suggest that a single sound source was perceived. The latency of these gaze shifts was surprisingly shorter than those of gaze shifts to 50-ms broadband acoustic targets presented from a single location in the context of a memory-guided saccade task, providing further support that one source was heard in the FS condition (Fig. 1F). Gaze shifts with changes in direction, which would have indicated that two separate sounds were perceived (Tollin & Yin, 2003), were not observed and at the end of the study subjects reported hearing only one sound in all acoustic trials.

A summary of the subjects’ perception of the FI is shown in Fig. 1G. The perception of the FI was weakest with the 500-Hz stimulus, approximately 20% of trials, whereas at the other frequencies tested the FI was perceived in approximately 60% of trials. The fact that the subjects oriented to the transient component of the FS in the 500-Hz condition in a

significantly fewer number of trials compared with the other frequencies tested indicates that perception of the FI was frequency dependent, as shown by Yost *et al.* (1997). However, these results, which showed that the FI was perceived similarly across the frequencies of 1–6 kHz tested, differ from those of Yost *et al.* (1997), who showed that the FI was strongest with the 1.5-kHz stimulus. This discrepancy could be due to differences in the experimental tasks used to assess the perception of the FI and the acoustic properties of the experimental rooms.

Monkey—The sound localization behavior exhibited by the two monkeys in response to the FS (Fig. 1H and I) is very similar to the behavior exhibited by the humans presented with identical stimuli (Fig. 1D and E) and consistent with our previous observations (Populin, 2006). Overall, the monkeys responded to FS stimuli with short-latency gaze shifts, similar to those of gaze shifts to single source broadband targets (Fig. 1J), directed to the transient component in 74% of all trials (Fig. 1K). Unlike for the human experiments, the frequency of the stimuli used for testing was determined by the tuning characteristics of the single IC units isolated for recordings (see Fig. 3A).

Gaze shifts directed to either the transient component of the FS, indicating that the FI was perceived (magenta traces in Fig. 1H and I), or gaze shifts directed to the sustained component of the stimuli (black traces in Fig. 1G and I), indicating the FI was not perceived, had similar short latencies and no changes in direction, both of which are characteristics indicative of responses executed without hesitation to a clearly perceived single target (Tollin & Yin, 2003). This assertion, given that we cannot query the monkeys directly about their perceptions, is therefore based on the fact that their behavioral responses to the FS were similar to those of the humans who reported always hearing one sound, a method that has been used extensively in the study of visual illusions (e.g. Logothetis, 1998; Blake & Logothetis, 2002).

Because our analysis of the physiological mechanisms underlying the FI involved single-unit responses to the components of the FS, we also examined behavioral responses to each component presented separately. Localization of the sustained stimulus without a sharp onset was very poor (Fig. 2A). The average error rate across all frequencies tested was 47%, which, in light of the fact that during recording sessions most acoustic stimuli were presented from one of two transducers at ($\pm 30^\circ, 0^\circ$), indicates that the subjects were essentially responding at chance level. By comparison, localization of the transient component of the FS presented alone, despite the fact that the subjects had to wait for 450 ms before reporting the perceived location of the sound, was significantly better, with an error rate of 17% (Fig. 2A; *t* test, $P = 0.0039$). The latency of the gaze shifts to the perceived sources of the sustained stimulus was significantly longer than those of the transient (Fig. 2B), therefore revealing greater difficulty in selecting the responses in this condition (Posner, 1978).

Physiological recordings

After having established that humans and monkeys localized FS similarly, and thus confirming that both perceive the FI, its underlying neural mechanisms were studied by

recording from the IC of the monkeys working with the head unrestrained. To date, there are no physiological data regarding the mechanisms underlying the perception of the FI from a preparation that indicates, on a trial-by-trial basis, when it is or is not perceived, a fundamental requirement to correlate perception with its underlying physiological mechanism.

The IC was chosen because the discharges of its neurons (i) exhibit sensitivity to sound location (Irvine, 1986; Zwiers *et al.*, 2004) resulting from numerous converging inputs from lower brainstem nuclei (Adams, 1979) that process interaural differences of time and level (Yin, 2002), and (ii) are dominated by the leading sound of a pair of stimuli presented in close succession, with the responses to the lagging sound attenuated relative to the magnitude of responses evoked by the same stimuli presented alone (Fitzpatrick *et al.*, 1995; Tollin *et al.*, 2004). Recordings were taken from both the left and the right IC. The data were analysed and are presented according to the location of the transducers, i.e. contralateral or ipsilateral, relative to the recording side. For ease of presentation, the illustrations in all figures show the left as the recording side.

General physiological findings

To dissect the physiological mechanisms underlying the behavioral responses evoked by the presentation of the FS (Fig. 1), the relationship between the physiological responses of IC neurons to the components of the FS presented separately and the associated sound localization behavior will be examined first.

Consistent with previous recordings from the IC of awake monkeys (e.g. Zwiers *et al.*, 2004; Versnel *et al.*, 2009), the present study revealed single units characterized by narrow tuning functions that were arranged tonotopically with characteristic frequency increasing dorso-ventrally with depth along the path of the electrode penetrations. Examples of such tuning functions are shown in Fig. 3A; a schematic of the arrangement used to determine each unit's characteristic frequency is shown in the upper left panel. The animal was required to fixate straight ahead for about 7–8 s for the presentation of the same sustained, 500-ms component of the FS in increasing frequency steps of 1 kHz from the field contralateral to the recording electrode. We had observed previously that head-unrestrained monkeys presented with an LED straight ahead align both their eyes and head with it during periods of fixation of several seconds (Populin, 2006), thereby ensuring consistent head position for multiple presentations of the single-tone stimulus within and across trials. The frequency that evoked the largest response over the entire 500-ms stimulus was used for both the FS and the single-source components for subsequent testing. Responses to the three stimuli were very consistent across the population, as demonstrated by the normalized firing rates shown in Fig. 3B–D.

The poor localization of the sustained stimuli documented above (Fig. 2A), albeit consistent with previous psychophysical findings in humans (Rakerd & Hartmann, 1986), is nevertheless puzzling because single neuronal responses evoked by the sustained stimulus presented contralateral to the recording side had both significantly shorter latencies (Fig. 4A) and higher discharge rates (Fig. 4B) than the responses evoked by the stimulus presented ipsilaterally. No significant differences were found in the first spike latency of

responses evoked by the transient stimulus presented contralateral and ipsilateral to the recording side (Fig. 4A), but the average discharge rate of the responses evoked by contralateral transient stimuli was significantly higher (Fig. 4B). Lastly, and pertinent to the mechanisms underlying the perception of the FI, the transient stimulus, regardless of the side from which it was presented, always produced responses with shorter latencies (Fig. 4C and D) and larger average discharge rates (Fig. 4E and F) than the sustained stimulus presented on the opposite side.

Single-unit responses to the sustained component of FS presented alone

The experimental arrangement for trials in which the sustained component of the FS was presented alone from transducers at ($\pm 30^\circ, 0^\circ$) in randomly selected trials is illustrated in Fig. 5A. As indicated, localization of the sources of this stimulus was at chance level (Fig. 2A). Figure 5B and C show examples of gaze shifts and single-unit recordings from correct trials, in which the subject oriented to the source of the stimulus, either contralateral (B) or ipsilateral (C) to the recording (dark and light blue traces), and from incorrect trials, in which the subject oriented away from the source of the stimuli (red and orange traces); spike density functions, computed by convolving each spike with a postsynaptic activation function with a rise time of 1 ms and a decay time of 10 ms (Hanes *et al.*, 1995), are plotted in the colors corresponding to the gaze shifts.

Receiver operating characteristic analyses were used to compare evoked responses in the sample of IC neurons based on the location of the source of the stimuli (Fig. 5D), and based on the subjects' reports of perceived sound location (Fig. 5E and F). The analyses were carried out across the entire 500 ms of the stimuli (Fig. 1). Consistent with the differences in discharge rate documented above (Fig. 4B), the neuronal responses exhibited selectivity for stimuli presented on the side contralateral to the recording (Fig. 5A and D), illustrated by the significant ROC values highlighted in blue, throughout most of the 500 ms of the stimulus (ROC value > 0.5 , contra $>$ ipsi). Most interesting, however, is the fact that the neurons did not show selectivity for correct vs. incorrect responses, i.e. trials in which the monkeys oriented to the source of the sound or away from it, respectively, when the stimulus was presented contralateral (Fig. 5B and E) or ipsilateral (Fig. 5C and F) to the recording side (ROC value > 0.5 , correct $>$ incorrect). In both configurations, the ROC value of the comparison between correct and incorrect trials did not differ significantly along the entire duration of the stimulus (Fig. 5E and F). Thus, the lack of selectivity of the IC neurons for correct vs. incorrect trials is consistent with the behavior of the subjects and indicates that contralateral vs. ipsilateral differences in first spike latency and discharge rate for slowly rising single-frequency tones are not sufficient to specify sound location at the level of the IC.

Single-unit responses to the transient component of FS presented alone

The same neurons exhibited a distinct pattern of selectivity in response to the transient component of the FS presented alone. The experimental arrangement for trials in which the transient component of the FS was presented alone from transducers at ($\pm 30^\circ, 0^\circ$) is illustrated in Fig. 6A. Because this stimulus was also presented in the context of the memory-guided saccade task, which required the monkeys to wait for an additional 450 ms

after the presentation of the 50-ms stimulus in order to respond, the analysis was carried out over the same time frame used for the sustained stimulus (Fig. 5). As expected, the neurons responded more strongly to stimuli presented in the contralateral hemifield, as shown by the significant ROC values highlighted in blue (Fig. 6D; ROC values > 0.5 , contra $>$ ipsi). Note that the stronger discharges associated with the contralateral stimulus exceeded its 50-ms duration.

Unlike responses to the sustained stimulus, the neurons selectively distinguished successful from unsuccessful trials when the transient stimulus was presented alone in the field contralateral to the recording (Fig. 6E; ROC values > 0.5 , correct $>$ incorrect). Note that in this example the subject made some incorrect responses (plotted in red) when the stimulus was presented on the contralateral side (Fig. 6B). Conversely, when the transient stimulus was presented on the side ipsilateral to the recording (note: the monkey did not make any errors in this example, Fig. 6C; but see Fig. 2A), the neurons did not show selectivity between correct and incorrect trials during the early part of the response (Fig. 6F; ROC values > 0.5 , correct $>$ incorrect). Significant ROC values were consistently observed after approximately 250 ms, while the monkey waited in the absence of acoustic stimulation for the offset of the fixation LED, which commanded a behavioral response (Fig. 6F), but relevance to the behavior is unclear.

These results, taken together, reveal that at the level of the IC, responses to the sustained and transient stimuli correlate well with the corresponding sound localization behavior of the monkeys. We hypothesize that the difference between localization performance to the transient vs. the sustained (Fig. 5) results from the spectral splatter, which activates more than the single frequency channel of the pure tone stimulus, produced when playing a signal with an abrupt onset (see Fig. 9D).

The perception of the FI, however, results from the simultaneous presentation of both sustained and transient stimuli (Fig. 1A) from transducers located on each side of the subject. The analysis of responses to the presentation of the FS, recorded from the same neurons, and its relation to the subjects' perception of the FI is presented below.

Responses to FS and mechanism of localization

The experimental arrangement used to test for the perception of the FI is shown in Fig. 6A and B. Again, the subjects were trained to orient to the perceived location of sound sources (Populin, 2006), and only short-latency, deliberate gaze shifts to the source of one stimulus were observed, suggesting that, regardless of whether the FI was perceived, a single stimulus was heard (Figs 1F and J, and 2B). Behavioral responses plotted in magenta illustrate trials in which the FI was perceived and those plotted in black illustrate trials in which it was not; the corresponding spike density functions are plotted in the same colors (Fig. 7C and D). The physiological recordings in Fig. 7 were obtained from the same neuron used to illustrate responses to the single-tone stimuli in Figs 5 and 6, which had a characteristic frequency of 8 kHz. In this example, the subject perceived the FI in 60% of the trials when the transient stimulus was presented in the contralateral field (Fig. 7C), and in 100% of the trials when the transient stimulus was presented in the ipsilateral field.

The data in Fig. 4C–F, showing that responses to the transient stimuli had shorter first spike latencies than those of responses to the sustained stimuli on the opposite side, are consistent with the hypothesis that the localization aspect of hearing the FI is dominated by the sound that produced the response with the shortest latency (Hartmann & Rakerd, 1989). As expected, and consistent with this hypothesis, no differences were found between the first spike latency of responses evoked by the transient stimulus alone and the FS (Wilcoxon signed-ranked test, $P = 0.6672$).

Thus, it is imperative to determine if physiological responses in the IC differ with perception of the illusion. If no differences were found at this level of processing in the auditory pathways, it would suggest that the mechanism underlying the perception of the FI is found above the midbrain, as proposed by Yagcioglu & Ungan (2006). Accordingly, the same ROC analysis described above for the single-source stimuli (Figs 5 and 6) was performed, comparing the responses of the sample of IC neurons from trials in which the subjects reported perceiving the FI vs. trials in which they did not (ROC values > 0.5 , FI perceived $>$ FI not perceived). The subjects' behavioral response is indicated by the arrows with the colors corresponding to the stimuli. For the two configurations of stimuli, the neuronal discharges were significantly larger in trials in which the subjects reported perceiving the illusion (Fig. 7E and F), as illustrated by the significant ROC values highlighted in magenta. Despite some differences in the responses between the two configurations, the initial portion of the response, which is attributed to the transient component of the FS based on its shorter latency and higher discharge rate (Fig. 4A and B), was larger in trials in which the FI was perceived (see arrows in Fig. 7E and F), therefore confirming that the localization aspect of the FI is due to the transient component of the FS. Furthermore, it suggests, by extension, that failure to perceive the FI results from ineffective encoding of the transient stimuli.

Comparison of responses to FS vs. single-source sustained: parsing the underlying mechanism of the FI

Perception of the FI is characterized by hearing the sustained component of the FS in the hemifield opposite from where it is presented (Fig. 1). Psychophysical studies of the FI (e.g. Yost *et al.*, 1997; Dent *et al.*, 2004, 2007; Populin, 2006) have used the sum of the components of the FS presented from a single location for comparison because that is what the subjects are assumed to perceive. However, as indicated at the outset, we sought to compare neural responses evoked by identical stimuli that led to different behavioral responses and those evoked by different stimuli that led to the same behavior, i.e. the sustained component presented by itself compared with the FS. Having access to both the behavioral responses and the neural activity evoked by each of these stimuli allowed us to compare neural responses not only based on where stimuli were presented but also on where they were heard by the subjects. Accordingly, we compared the responses evoked by the FS to those evoked by the sustained stimulus presented alone and correctly localized using the sliding ROC analysis (Fig. 8; ROC values > 0.5 , sustained $>$ FS).

Comparisons with the transient component of the FS located contralateral to the recording side are shown in Fig. 8A–D. In Fig. 8A, the sustained component, either alone or as part of the FS, was presented ipsilateral to the recording. The subjects' behavioral response is

indicated by the arrows with the colors corresponding to the stimuli. Figure 8A illustrates the comparison between trials in which the subjects reported hearing the FI vs. those in which they successfully localized the source of the sustained stimulus presented alone, i.e. the same sustained stimulus led to different perception, whereas in Fig. 8B, responses to the same sustained stimulus are compared for trials in which the subjects did not perceive the FI, i.e. the same sustained stimulus was heard in the same location. The first ~25 ms of the sliding ROC comparison shows significant values attributed to neural activity driven by the transient component of the FS (ROC values < 0.5 , FS $>$ sustained; Fig. 8A, area in magenta with arrow), which we hypothesize defines sound location in this context. This was followed by a period of significant ROC values in the opposite direction representing larger activation produced by the sustained stimulus presented alone compared with the FS (area in yellow with double-headed arrow). This difference indicates that the presentation of the transient component of the FS contralateral to the recording effectively inhibited, albeit for a brief period, the activation resulting from the sustained component of the FS presented ipsilaterally. The timing of this inhibition coincides with the time in which the sustained component increases in amplitude, suppressing any potential spatial cue that could arise from its onset, however small. In the remaining ~450 ms of the stimulus there were essentially no differences between the FS and the sustained stimulus. At this point, based on the subjects' behavioral responses, the sustained component of the FS had to have been heard in the opposite hemifield, at the location defined by the transient stimulus. Thus, sound location is defined by the transient, and inhibition does play a role in suppressing potential onset cues from the sustained stimulus, although it cannot account for the persistent, erroneous perception of sound location of the sustained component of the FS. Figure 8B illustrates the comparison of responses evoked by an identical set of stimuli but considering only trials in which the subjects did not perceive the FI, and reveals a pattern of responses in which the transient component of the FS was not salient enough relative to the responses evoked by the sustained component. Despite the fact that the subjects did not localize the transient in these trials, the effective encoding of which is essential for the perception of the FI, its presence in the contralateral field attenuated the response to the sustained compared with the sustained presented alone (area in yellow).

In Fig. 8C and D the sustained component was presented ipsilateral to the recording in the FS and contralateral when presented alone. Figure 8C illustrates the comparison between trials in which the subjects reported hearing the FI vs. those in which they successfully localized the source of the sustained, i.e. different sustained stimulus, same perception, whereas in Fig. 8D, responses to the same stimuli are compared for trials in which the subjects did not perceive the FI, i.e. different sustained stimulus, different perception. In Fig. 8C, as expected given that the subjects perceived the FI and the sustained stimulus was contralateral to the recording, no significant differences were observed. We attribute the lack of significance of the responses evoked by the transient component of the FS in Fig. 8C to the fact that it is compared with a strong response evoked by the sustained stimulus alone presented contralaterally. Furthermore, it reveals that the activity of these neurons is not different between trials in which the sustained stimulus was presented from the contralateral field or simply perceived there. In Fig. 8D, comparison of the neural responses evoked by the same stimuli, but considering trials without perception of the FI, reveals that the neural

response evoked by the sustained stimulus was significantly larger during the 25–50-ms epoch. This difference reflects the fact that a contralateral sustained stimulus evokes a larger response than an ipsilateral one.

Similar comparisons based on the reversed configuration of the FS relative to the recording side are shown in Fig. 8E–H, and also illustrate the dominant role of the response evoked by the transient component of the FS in leading to the perception of sound location in the FI. In addition, they illustrate activity that may underlie the sustained erroneous perception that characterize the FI.

In Fig. 8E and F the sustained component, either alone or as part of the FS, was presented contralateral to the recording. Figure 8E illustrates the comparison between trials in which the subjects reported hearing the FI, i.e. the same sustained stimulus resulted in different perception of sound location, whereas in Fig. 8F, responses to the same stimuli are compared for trials in which the subjects did not perceive the FI, i.e. the same sustained stimulus resulted in the same perception. Figure 8E shows significant ROC values at the onset of the stimulus, during the first ~30 ms (smaller than the 5th percentile, magenta area with arrow), but in contrast to the opposite configuration (Fig. 8A), significant ROC values below the 5th percentile were also observed for the sustained portion of the stimuli. Thus, the activity evoked by the sustained component of the FS was enhanced relative to the activity evoked by the presentation of the sustained stimulus alone from the same location (Fig. 8E). With the same configuration of stimuli but considering trials in which the FI was not perceived (Fig. 8F), the ROC values are strikingly different compared with those obtained in Fig. 8E. The significant ROC values within the first 50 ms lie above the 95th percentile, therefore indicating that the response evoked by the sustained stimulus alone in the contralateral field was greater. Thus, when the FI was perceived, the transient component of the response was detected, specifying sound location, and the response to the sustained portion of the FS, which lacks spatial information, was enhanced and could be responsible for the sustained perception of the FI (Fig. 8E). When the FI was not perceived, the transient component of the FS was not detected, and the response to the sustained component of the FS was attenuated (Fig. 8F).

Lastly, in Fig. 8G and H the sustained component was presented contralateral to the recording in the FS and ipsilateral when presented alone. Figure 8G illustrates the comparison between trials in which the subjects reported hearing the FI, i.e. different sustained stimulus, same perception, whereas in Fig. 8H, responses to the same stimuli are compared for trials in which the subjects did not perceive the FI, i.e. different sustained stimulus, different perception. Unlike the complementary comparison in Fig. 8C, in which different stimuli produced the same perception and the subject perceived the FI, the comparison in Fig. 8G reveals significant ROC values in the first 30 ms (marked with arrow), which we attribute to effective encoding of the transient component of the FS compared with a weaker response evoked by the sustained stimulus presented ipsilaterally. Significant ROC values were also observed during the last ~425 ms of the comparison resulting from the contralateral sustained component of the FS and, as in Fig. 7E, this could be responsible for the sustained perception of the FI. For the same configuration of stimuli but considering trials in which the FI was not perceived (Fig. 8H), the ROC comparison

reveals a different pattern, without the large significant peak during the first 25 ms, which we attribute to ineffective encoding of the transient. The remaining duration of the comparison reveals a similar pattern of significant ROC values resulting from a stronger response to the contralateral vs. ipsilateral sustained stimulus.

These data therefore indicate that the responses of IC neurons were correlated with the subjects' perception of sound location and not strictly with the actual location of the stimuli. Specifically, they demonstrate that perception of the FI is correlated with effective encoding of the transient component of the FS and persistent activation of the same neurons by the sustained component at the level of the IC.

Discussion

The data show that (i) humans and monkeys respond to the FS similarly, (ii) the transient component of the FS, which produces a shorter first spike latency and higher discharge rate than the sustained, determines the perception of sound location, and (iii) the persistent, erroneous perception of the sustained stimulus location is due to continued excitation of the same neurons, first activated by the transient, by the sustained stimulus without location information. The physiological basis to create the perception of the FI is therefore found at the level of the IC. In addition, and contrary to well-established principles, the data also show that ipsilateral vs. contralateral differences in first spike latency and discharge rate in the responses of IC neurons are not sufficient to specify sound location for slowly rising single-frequency tones. This is the first study to date demonstrating, on a trial-by-trial basis, a correlation between perception of the FI and an underlying physiological substrate.

Behavioral and physiological responses to components of FS

Central to the perception of the FI is the inability to localize the source of single-frequency stimuli without a sharp onset (Hartmann & Rakerd, 1989). Yost *et al.* (1997), for instance, demonstrated an inverse relationship between localization accuracy and perception of the FI for mid-frequency signals. Hartmann & Rakerd (1989) suggested that under echoic conditions such stimuli produce implausible localization cues that the system cannot use for accurate localization. Although our data are agnostic about the nature of the mechanism that discards such implausible cues, they are nevertheless consistent with this hypothesis, because the monkeys were unable to localize the source of the sustained component of the FS presented alone (Figs 2B, and 4A and B).

Interestingly, despite the implausible cues generated by the sustained stimulus under echoic conditions, significant contralateral vs. ipsilateral differences in first spike latency and average discharge rate were documented in the responses of single IC units. In theory, such differences should have sufficed to determine the laterality of the stimulus source. Not surprisingly, however, in light of the psychophysical data (Fig. 2B), the discharge of the IC neurons was indistinguishable between trials in which the monkeys correctly determined the location of the sound sources and trials in which they failed (Fig. 5E and F). The transient stimulus, in contrast, presented little difficulty to the subjects (Fig. 2A), who were able to localize its source despite the lack of significant differences between contralateral and ipsilateral first spike latencies. The average discharge rate was significantly higher for

stimuli presented contralaterally (Fig. 6A) and, in agreement with the localization behavior, the discharge rate of IC neurons was higher in correct compared with incorrect trials when the transient was located contralateral to the recording (Fig. 6E and F).

The difference in localizability between the two stimuli of the same frequency may stem from spectral splatter generated when playing a signal with an abrupt onset, such as the transient component of the FS, through a transducer (Fig. 9D). This assertion implies that activation of a single-frequency channel may not be sufficient to determine sound location.

Behavioral and physiological responses to FS: perceiving vs. not perceiving the illusion

Monkeys and humans responded similarly to the FS with deliberate, short-latency gaze shifts directed more often to the source of the transient component, suggesting that they perceived the FI, than to the source of the sustained component, suggesting that they did not. No corrective responses were observed in which the subjects first looked in the direction of one of the sources and then to the other, which would indicate that more than one sound was heard (Tollin & Yin, 2003). The use of the memory-guided saccade task (Fig. 1) may have contributed to the elimination of corrective responses by compelling the subjects to listen to the entire 500-ms stimulus before responding, resulting in a clean trial-by-trial measure of whether the FI was perceived or not. Monkeys oriented to the transient component of the FS in a larger proportion of trials than humans (Fig. 1G and K), which could have resulted from differences in the acoustics of the two nearly identical rooms used for testing and differences in the location of the transducers used to present the stimuli.

The comparison of the responses evoked by the FS in IC neurons between trials in which the subjects perceived and did not perceive the FI revealed that the responses were larger in trials in which they did, regardless of the spatial configuration of the stimuli (Fig. 6E and F). Importantly, the difference was particularly salient during the first 50 ms of the stimuli, which corresponds to the duration of the transient component, therefore suggesting that its effective encoding is essential for perceiving the FI. Thus, ineffective encoding of the transient leads to lack of perception of the FI, the cause of which could reside either in the IC or in earlier stages of auditory processing. While the localization aspect of the FI is clearly dependent on the effectiveness with which the transient component is encoded and on the spatially undefined nature of the sustained, the question remains as to the neural mechanism underlying the erroneous, prolonged perception of the sustained component's location. Two potential explanations are considered.

The perception of the ongoing sound at the wrong location may result from the action of an inhibitory mechanism that affects its binaural processing. The opened-ended time course of the illusion argues against this possibility, however. Comparison of the responses evoked by the sustained stimulus presented alone to the responses evoked by the FS (with both sustained stimuli presented ipsilaterally), from trials in which the subjects correctly localized the sustained stimulus and perceived the FI (Fig. 8A), reveals that the magnitude of the response attributed to the sustained component of the FS is indeed diminished by inhibition produced by the transient stimulus presented contralaterally. Importantly, the timing of this inhibition, the duration of which matches the 50 ms of the transient, coincides with the initial, long-latency response to the sustained but does not last for its entire duration. Thus,

any localization information that could result from the onset of the sustained component of the FS is suppressed but the duration of the response is not. Accordingly, inhibition cannot account for the long-lasting misperception of the sustained stimulus.

Alternatively, the persistent misperception of location of the sustained component of the FS may result from a mechanism that assigns location based on spatial information from the transient stimulus of the same frequency. As shown in Figs 2 and 5, the monkeys' localization of the slowly rising, single-tone stimuli was essentially at chance level. Such a stimulus is expected to activate neurons of a particular frequency channel in the IC, without spatial information (Fig. 9A). This mode of neuronal activation is essential for the perception of the FI.

In contrast, the activity of the same IC neurons in the same frequency channel evoked by the transient component of the FS presented alone, which has both shorter latency and higher discharge rate than the responses evoked by the sustained stimulus presented in the opposite side (Fig. 4C–F), do encode the location of the sound. The schematic representation of this type of evoked activity illustrated with the transient presented contralaterally in Fig. 9B includes activation of frequency channels beyond the center frequency of the stimulus because of the spectral splatter that results from its sharp onset. This is illustrated in Fig. 9D, which shows the comparison of the frequency spectrum of 4-kHz sustained and transient signals recorded from an acoustic transducer used in this study. Clearly, the frequency content of the transient stimulus was broader than the frequency content of the sustained, which was narrowly centered at 4 kHz.

Thus, when the two stimuli, the sustained and transient tones, are presented simultaneously from transducers on each side of the subject to create the FS (Fig. 9C), IC neurons are first activated by the transient component, defining the location of the sound. Illustrated in the time domain (Fig. 9E), as the amplitude of the transient stimulus and therefore its effectiveness for driving the neuronal responses decreases during its 50-ms duration, the amplitude of the sustained stimulus, and its effectiveness in driving the same neurons, slowly increases but without spatial information. This mechanism links the localization and the erroneous prolonged perception of the sustained component of the FS. In this manner, the FI can be perceived for various lengths of times depending solely on the duration of the sustained stimulus. Therefore, a physiological correlate of the FI is present at the level of the IC.

Acknowledgments

This work was supported by grants from the National Science Foundation (IOB-0517458) and the National Institutes of Health (DC003693 and T32GM007507). We thank K. Johnston and K. Shen for software to carry out the ROC analysis, J. K. Harting, P. H. Smith and A. E. Anderson for comments on an earlier version of this manuscript, to Jane Sekulski for help with the data acquisition software, and to Joel Weinstein for ophthalmological help in the human experiments.

Abbreviations

FI Franssen illusion

FS	Franssen stimuli
IC	inferior colliculus
LED	light-emitting diode
ROC	receiver operating characteristic

References

- Adams JC. Ascending projections to the inferior colliculus. *J Comp Neurol.* 1979; 183:519–538. [PubMed: 759446]
- Batra RR, Kuwada SS, Stanford TRT. High-frequency neurons in the inferior colliculus that are sensitive to interaural delays of amplitude-modulated tones: evidence for dual binaural influences. *J Neurophysiol.* 1993; 70:64–80. [PubMed: 8395589]
- Berkley DA. Room acoustics and listening. *J Acoust Soc Am.* 1983; 73:S17.
- Blake RR, Logothetis NKN. Visual competition. *Nat Rev Neurosci.* 2002; 3:13–21. [PubMed: 11823801]
- Dent M, Tollin D, Yin T. Cats exhibit the Franssen effect illusion. *J Acoust Soc Am.* 2004; 116:3070. [PubMed: 15603152]
- Dent ML, McClaine EM, Welch TE. The Franssen effect illusion in budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). *J Acoust Soc Am.* 2007; 122:3609–3614. [PubMed: 18247768]
- Everling S, DeSouza JFX. Rule-dependent activity for prosaccades and antisaccades in the primate prefrontal cortex. *J Cognitive Neurosci.* 2005; 17:1483–1496.
- Fitzpatrick DC, Kuwada S, Batra R, Trahiotis C. Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol.* 1995; 74:2469–2486. [PubMed: 8747207]
- Franssen, NV. Stereophony. Phillips Technical Library; Eindhoven, the Netherlands: 1962.
- Hafta E, Dye R Jr, Gilkey R. Lateralization of tonal signals which have neither onsets nor offsets. *J Acoust Soc Am.* 1979; 65:471. [PubMed: 489816]
- Hanes DP, Thompson KG, Schall JD. Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. *Exp Brain Res.* 1995; 103:85–96. [PubMed: 7615040]
- Hartmann WM, Rakerd B. Localization of sound in rooms. IV: the Franssen effect. *J Acoust Soc Am.* 1989; 86:1366–1373. [PubMed: 2808910]
- Irvine, DRF. The auditory brainstem: a review of the structure and function of auditory brainstem processing mechanisms. In: Ottoson, D., editor. *Progress in Sensory Physiology*. Springer; Berlin: 1986. p. 165–177.
- Irvine DR, Gago G. Binaural interaction in high-frequency neurons in inferior colliculus of the cat: effects of variations in sound pressure level on sensitivity to interaural intensity differences. *J Neurophysiol.* 1990; 63:570–591. [PubMed: 2329362]
- Johnston K, DeSouza JFX, Everling S. Monkey prefrontal cortical pyramidal and putative interneurons exhibit differential patterns of activity between prosaccade and antisaccade tasks. *J Neurosci.* 2009; 29:5516–5524. [PubMed: 19403819]
- Judge SJ, Richmond BJ, Chu C. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res.* 1980; 20:535–538. [PubMed: 6776685]
- Kuwada S, Yin TC. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol.* 1983; 50:981–999. [PubMed: 6631473]
- Kuwada S, Stanford TR, Batra R. Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit – effects of changing frequency. *J Neurophysiol.* 1987; 57:1338–1360. [PubMed: 3585471]

- Logothetis NK. Single units and conscious vision. *Philos T Roy Soc B*. 1998; 353:1801–1818.
- Populin LC. Monkey sound localization: head-restrained versus head-unrestrained orienting. *J Neurosci*. 2006; 26:9820–9832. [PubMed: 16988053]
- Populin LC. Human sound localization: measurements in untrained, head-unrestrained subjects using gaze as a pointer. *Exp Brain Res*. 2008; 190:11–30. [PubMed: 18575853]
- Populin LC, Yin TC. Behavioral studies of sound localization in the cat. *J Neurosci*. 1998; 18:2147–2160. [PubMed: 9482800]
- Posner, MI. *Chronometric Explorations of Mind*. Erlbaum; Hillsdale, NJ: 1978.
- Rakerd B, Hartmann WM. Localization of sound in rooms, III: onset and duration effects. *J Acoust Soc Am*. 1986; 80:1695–1706. [PubMed: 3794076]
- Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE T Biomed Eng*. 1963; 10:137–145.
- Semple MN, Kitzes LM. Binaural processing of sound pressure level in the inferior colliculus. *J Neurophysiol*. 1987; 57:1130–1147. [PubMed: 3585457]
- Sparks DL. An argument for using ethologically ‘natural’ behaviors as estimates of unobservable sensory processes. Focus on ‘Sound localization performance in the cat: the effect of restraining the head’. *J Neurophysiol*. 2005; 93:1136–1137. [PubMed: 15738271]
- Stevens S, Newman E. The localization of pure tones. *Proc Natl Acad Sci USA*. 1934; 20:593–596. [PubMed: 16577646]
- Stevens SS, Newman EB. The localization of actual sources of sound. *Am J Psychol*. 1936; 48:297–306.
- Tollin DJ, Yin TCT. Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. *J Neurophysiol*. 2003; 90:2149–2162. [PubMed: 12801901]
- Tollin DJ, Populin LC, Yin TCT. Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J Neurophysiol*. 2004; 92:3286–3297. [PubMed: 15295015]
- Tollin DJ, Populin LC, Moore JM, Ruhland JL, Yin TCT. Sound-localization performance in the cat: the effect of restraining the head. *J Neurophysiol*. 2005; 93:1223–1234. [PubMed: 15483065]
- Versnel H, Zwiers MP, van Opstal AJ. Spectrotemporal response properties of inferior colliculus neurons in alert monkey. *J Neurosci*. 2009; 29:9725–9739. [PubMed: 19657026]
- Yagcioglu S, Urgan P. The ‘Franssen’ illusion for short duration tones is preattentive: a study using mismatch negativity. *Brain Res*. 2006; 1106:164–176. [PubMed: 16831407]
- Yin, TCT. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In: Oertel, D.; Fay, RR.; Popper, AN., editors. *Integrative Functions in the Mammalian Auditory Pathways*. Springer; Heidelberg: 2002. p. 99-159.
- Yost WA, Mapes-Riordan D, Guzman SJ. The relationship between localization and the Franssen effect. *J Acoust Soc Am*. 1997; 101:2994–2997. [PubMed: 9165744]
- Zwiers MP, Versnel H, van Opstal AJ. Involvement of monkey inferior colliculus in spatial hearing. *J Neurosci*. 2004; 24:4145–4156. [PubMed: 15115809]

Fig. 1. Demonstration of Franssen illusion (FI) in humans and monkeys. (A–C) Schematic representation of the experimental setup, task and Franssen stimuli (FS). (D, E) Horizontal gaze position traces from a human subject depicting trials in which the subject oriented to either the transient component of the FS, demonstrating perception of the FI (magenta), or the sustained component (black), indicating that the subject did not perceive the FI. The time course of the fixation LED is shown above each panel. The two components of the FS, which were presented from transducers at $(\pm 60^\circ, 0^\circ)$ in either of two configurations, are illustrated by the shape of the envelope. (F) Latency of gaze shifts from trials with FS and broadband stimuli. (G) Percentage of trials in which gaze shifts were directed to the transient component of the FS for the two subjects tested ($n = 441$ trials); the standard bars represent 95% confidence intervals. (H, K) Behavioral data from a monkey subject and summary as described for D and G, but with transducers located at $(\pm 30^\circ, 0^\circ)$. (J) Latency of gaze shifts from trials with FS and broadband stimuli. (K) Percentage of trials in which the gaze shifts were directed to the transient component of the FS demonstrating perception of the FI ($n = 969$ trials).

Fig. 2.

Monkey sound localization errors and gaze shift latency. (A) Percentage of right/left (left/right) localization errors of sustained and transient acoustic stimuli. The horizontal broken line represents the mean localization error of the sustained stimuli (47.3%) and solid line represents the mean localization error for the transient stimuli (17.5%). (B) Latency of the gaze shifts directed at the sources of sustained and transient stimuli presented alone.

Fig. 3.

Basic physiological properties of inferior colliculus neurons: frequency selectivity and normalized discharge rate. (A) Representative tuning functions of 11 units illustrating the range of frequencies studied and characteristic frequencies (the frequency that evoked the maximal discharge) encountered. The long duration component of the FS was presented alone from a speaker located contralateral to the recording side at either ($\pm 30^\circ, 0^\circ$), while the animal fixated on an LED straight ahead (diagram in the upper left). The characteristic frequency of each neuron was used to study the Franssen illusion. (B, C) Normalized

discharge rate of single-unit responses to sustained and transient stimuli presented contralaterally (left) and ipsilaterally (right) to the recording. (D) Normalized discharge rate of single-unit responses to FS for both stimulus configurations. Shaded area about the mean represents standard error of the sample mean.

Fig. 4.

First spike latency and average discharge rate of neuronal responses to sustained and transient stimuli. (A) First spike latency of responses evoked by the sustained (+) and transient (triangles) stimuli from all 38 neurons studied. For the sustained stimulus, contralaterally evoked responses had a significantly shorter latency, whereas no difference was found for the transient stimuli (Wilcoxon sign-ranked test). (B) The average discharge rate of responses evoked by both sustained and transient stimuli, computed over the duration of the stimuli, was significantly greater for stimuli presented contra-lateral to the recording side. (C, D) The first spike latency of the transient stimuli was significantly shorter than the sustained stimuli presented on the opposite side, regardless of stimulus configuration. (E, F) The average discharge rate of responses evoked by the transient stimuli was significantly higher than the average discharge rate of responses evoked by the sustained stimuli presented on the opposite side, regardless of configuration.

Fig. 5.

Behavioral and neural responses to the sustained component of the Franssen stimuli (FS) presented alone. (A) Schematic representation of the experimental arrangement. (B) Comparison between correct (blue) and incorrect (red) behavioral responses (horizontal gaze movement traces) to an 8-kHz tone stimulus, the unit's characteristic frequency, presented contralateral to the recording side. The unit's responses in each case are shown in raster format with their corresponding spike density function superimposed; scale bars also apply to C. (C) Comparison between correct (light blue) and incorrect (orange) responses to the same stimulus presented ipsilateral to the recording side and corresponding neuronal responses as in B. (D) Population ROC analysis comparing the magnitude of the responses evoked by the sustained stimuli presented contralaterally and ipsilaterally, with ROC values greater than 0.5 indicating greater activity in response to a contralateral stimulus compared with an ipsilateral stimulus. (E, F) Population ROC analysis comparing neural responses from correct and incorrect trials when the stimuli were presented contralateral and ipsilateral to the recording side, respectively; ROC values greater than 0.5 indicate greater activity in correct trials than in incorrect trials. Consistent with the behavior of the subjects, who could not accurately localize the sources of this type of stimuli (Fig. 2A), the neural activity was indistinguishable between correct and incorrect trials.

Fig. 6.

Behavioral and neural responses to the transient component of the Fransen stimuli (FS) presented alone; all other details as in Fig. 5. (A) Schematic representation of the experimental arrangement. (B) Comparison between correct (blue) and incorrect (red) behavioral responses (horizontal gaze movement traces) to an 8-kHz tone stimulus, the unit's characteristic frequency, presented contralateral to the recording side. The unit's responses in each case are shown in raster format with their corresponding spike density function superimposed; scale bars also apply to C. (C) Comparison between correct (light blue) and incorrect (orange) responses to the same stimulus presented ipsilaterally to the recording side and corresponding neuronal responses as in B. The subject made no errors in this example. (D) Population ROC analysis comparing the magnitude of the responses evoked by the transient stimuli presented contralaterally and ipsilaterally, assuming the response to the contralateral stimulus to be greater. Because the stimuli were presented in the context of the memory-guided saccade task, as in the other conditions (e.g. Fig. 5), the analyses were performed over a 500-ms window. The end of the stimulus is marked by a thin vertical dotted line at 50 ms. The response evoked by the presentation of the stimulus contralateral to the recording side was larger than the ipsilateral response. (E) Population ROC analysis comparing neural responses from correct and incorrect trials when the stimuli were presented contralateral to the recording side, respectively, with ROC values greater than 0.5 indicating greater activity in response to a contralateral compared with an ipsilateral stimulus. The discharge rate from correct trials was greater compared with incorrect trials. Interestingly, albeit by a smaller margin, the activity in the correct trials was greater throughout the entire 450-ms portion of the delay period following the presentation of the stimulus. (F) Population ROC analysis comparing neural responses from correct and

incorrect trials when the stimuli were presented ipsilateral to the recording side, with ROC values greater than 0.5 indicating greater activity in correct than in incorrect trials. In this condition, the activity evoked by the stimulus was not significantly different between correct and incorrect trials during the time of its presentation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 7.

Comparison of neuronal responses from trials in which the subjects perceived the Franssen illusion (FI), shown in magenta, vs. trials in which they did not, shown in black. (A, B) Schematic representation of the experimental arrangement. (C, D) Behavioral (left, horizontal gaze traces) and physiological (right, raster and corresponding spike density function) responses to 8-kHz (the unit's characteristic frequency) Franssen stimuli (FS), which were presented during the time period shown by the thin solid and broken vertical lines superimposed on the gaze traces. In D, with the transient component of the FS located ipsilateral to the recording, all responses were indicative of perception of the FI. (E, F) Population ROC analysis of neural responses to FS from all trials in which the subjects perceived the FI vs. all trials in which they did not. ROC values greater than 0.5 indicate greater response in trials when the subject perceived the FI compared with when they did not. For both configurations the responses to the FS from trials in which the subjects reported perceiving the illusion were significantly larger than the activity from trials in which they did not. Vertical dotted line indicates the end of the 50-ms transient stimulus.

Fig. 8.

Comparison of population neural responses evoked by Franssen stimuli (FS) to responses evoked by the sustained stimulus presented alone and correctly localized. The subjects' behavioral responses indicating their perception of sound location are illustrated by the direction of the color-coded arrows. ROC values greater than 0.5 indicate a greater response to the correctly localized sustained tone compared with the FS. Areas of the ROC curves highlighted in yellow represent values in neural activation resulting from the presentation of the sustained stimulus presented alone that were larger than the activation resulting from the presentation of the FS. Conversely, areas highlighted in magenta represent instances in which the activation resulting from the presentation of the FS was larger. (A) Franssen illusion (FI) perceived, with transient component of the FS presented contralateral to the recording side and sustained stimulus presented ipsilaterally. (B) FI not perceived, with the same stimuli configuration as in A. (C) FI perceived, with the transient component of the FS presented contralateral to the recording side and the sustained stimulus presented contralaterally. (D) FI not perceived, with the same stimuli configuration as in C. (E) FI perceived, with the transient component of the FS presented ipsilaterally to the recording side and the sustained stimulus presented contralaterally. (F) FI not perceived, with the same stimuli configuration as in E. (G) FI perceived, with the transient component of the FI presented ipsilaterally and the sustained stimulus presented ipsilaterally. (H) FI not perceived, with the same stimuli configuration as in G. See Results for further details.

Fig. 9.

Hypothesized mechanism underlying the perception of the Franssen illusion in the inferior colliculus (IC). (A) Activation produced by an ipsilateral, slowly rising, single-tone stimulus, which subjects cannot localize despite ipsi/contra differences in first spike latency and discharge rate. Frequency channels (fn) in the IC are depicted on the ipsilateral and contralateral side of the recording. (B) Activation produced by a transient, abrupt single-tone stimulus presented contralateral to the recording side. This stimulus is easily localized because it evokes discharges that encode sound location. It activates neurons outside its center frequency because of spectral splatter that results from its abrupt onset. (C) Activation produced by Franssen stimuli (FS) with the transient component contralateral to the recording electrode. The subject orients to the source of the transient stimulus, indicating that the FI was perceived. The transient component of the FS activates neurons within the specified frequency channel first, establishing in this manner the location of the target. Shortly thereafter, the sustained component of the stimulus, which does not encode source location, starts to activate the subset of neurons of the same frequency that have already encoded the location of the target. The transiently determined sound location, accordingly, is assigned to the slower occurring frequency-specific activation of the selected channel by the sustained stimulus. The sustained stimulus is therefore heard at the location of the transient stimuli. (D) Comparison of frequency spectrum of 4-kHz sustained (yellow) and transient (aquamarine) signals recorded at the output of an acoustic transducer used in this study. The

shaded region represents the spectral splatter that results from the abrupt onset of the transient signal. (E) Schematic representation of the first 100 ms of the FS; the sharp 50-ms tone is plotted in aquamarine, the slowly rising tone in yellow, and their sum in magenta (broken lines).