Predicting the path of a changing sound: Velocity tracking and auditory continuity

Poppy A. C. Crum and Ervin R. Hafter
University of California Berkeley, Berkeley, CA 94720

(Received 6 August 2007; revised 6 May 2008; accepted 16 May 2008)

Three studies demonstrate listeners' ability to use the rate of a sound's frequency change (velocity) to predict how the spectral path of the sound is likely to evolve, even in the event of an occlusion. Experiments 1 and 2 use a modified probe-signal method to measure attentional filters and demonstrate increased detection to sounds falling along implied paths of constant-linear velocity. Experiment 3 shows listeners perceive a suprathreshold tone as falling along a trajectory of constant velocity when the frequency is near to the region of greatest detection as measured in Experiments 1 and 2. Further, results show greater accuracy and decreased bias in the use of velocity information with increased exposure to a constant-velocity sound. As the duration of occlusion lengthens, results also show a downward shift (relative to a trajectory of constant velocity) in the frequency at which listeners' detection and experience of a continuous trajectory are greatest. A preliminary model of velocity processing is proposed to account for this downward shift. Results show listeners' use of velocity in extrapolating sounds with dynamically changing spectral and temporal properties and provide evidence for its role in perceptual auditory continuity within a noisy acoustic environment. © 2008 Acoustical Society of America. [DOI: 10.1121/1.2945117]

PACS number(s): 43.66.Fe, 43.66.Cb, 43.66.Lj [RYL]

Pages: 1116–1129

I. INTRODUCTION

The perception of a visual source moving through space is influenced by predictions of the path of motion (Orban de Xivry *et al.*, 2006; Churchland *et al.*, 2003; Rushton and Wann, 1999; Assad and Maunsell, 1995; Yantis, 1995; Pavel and Stone, 1992). As a source travels along a path, it generates an expectation for that source's location at a future moment in time (Pavel and Stone, 1992). For instance, when a fixated observer sees a ball roll behind a screen, properties of the ball's motion such as trajectory and velocity, initially represented as a pattern of activity moving across the retina, provide the observer with a reasonably accurate estimate of when and where it will reappear. Here, we ask whether the auditory system uses properties such as the rate of frequency change, corresponding to a pattern of activity moving along the basilar membrane, to make similar predictions about the path of a sound changing over time in frequency space, i.e., when a sound disappears behind an auditory screen, are there expectations in the sensory system of where it should re-emerge?

It is well known that interruptions in a sound, due to masking and occlusion by other more intense sounds, often go perceptually undetected (Miller and Licklider, 1950; Warren, 1999; Dannenbring, 1976). This is true even when a sound's frequency is changing across time [as during a frequency-modulated sweep (FM-sweep), Fig. 1(a)] and is truly discontinuous [Fig. 1(b)] but contains higher intensity noise during the discontinuity [Fig. 1(c)]. In this instance, energy present during the interruption that is sufficient to mask the absent components creates an illusory condition where the sound appears continuous [Fig. 1(d)] despite discontinuity of the original signal (Houtgast, 1972; Warren *et al.*, 1972). In a noisy environment, this ability is crucial. It

enables a listener to successfully encode and interpret complex sounds where masking from more intense sounds has degraded the content of the original source and perceptual restoration of information is necessary to achieve a more useful experience of the sound's content. For hearing-impaired listeners this phenomenon may be even more critical. In conditions where the effective signal-to-noise ratio is lower the perceptual restoration of occluded content is likely more heavily relied upon to maintain a source's stability and contextual relevance in the acoustic scene.

Referred to as auditory continuity or auditory induction, this phenomenon has been demonstrated throughout studies of auditory processing (Miller and Licklider, 1950; Bregman, 1997; Warren, 1999), and in many conditions has led to a relevant perceptual representation of the missing or occluded information despite dynamic shifts in the sound's spectral content across time. In these instances, properties of the sound prior to an occlusion may be used to induce the probable path of change during an interruption.

This tendency of sounds interrupted by noise to be perceptually continuous (Warren *et al.*, 1972; Warren 1999; Ciocca and Bregman, 1987; Kluender and Jenison, 1992; Bregman, 1994) has historically been regarded as a problem resolved by interpolation between audible regions of a sound across an occlusion (Warren, 1999; Grossberg *et al.*, 2003; Husain *et al.*, 2005). In this scenario, perceptual restoration is accomplished through a best-fit approximation between information occurring before and after the interruption. In contrast, perceptual restoration during an interruption could occur through extrapolation as previously described in the visual analogy. This scenario would, instead, depend on prediction of the spectral and temporal path of the changing sound as indicated from the portion of the sound that preceded the interruption. Use of information such as the rate of

FIG. 1. A cartoon representation of perceptual continuity for sounds changing in frequency across time. (a) A continuous stimulus. (b) Two discrete sections interrupted by a silence. (c) Two discrete sections interrupted by an N-occluder. (d) The perceptual restoration that occurs during conditions of (c).

frequency change to extrapolate the path of a dynamically changing sound would suggest a computationally more efficient system for resolving occluded information than previously discussed. In an extrapolative scenario, listeners' narrow and position their listening region in expectation of the continuation of an occluded sound rather than monitoring multiple frequency regions (necessary in an interpolative scenario) for a continuation and interpolating across the interruption only after energy has been detected in a subsequent band. There are similarities between the predictive processes which occur while observing an object moving in time through visual space (Orban de Xivry *et al.*, 2006; Churchland *et al.*, 2003; Pavel and Stone, 1992; Assad and Maunsell 1995; Yantis, 1995) and a sound changing in time through frequency space. Understanding these similarities should elevate the role of extrapolation within current views of auditory processing.

To address the influence of continuously changing spectral and temporal information on listeners' perception, we measured whether listeners had an expectation of where an auditory trajectory should reappear when it disappears behind an occluder. Previously, Kluender and Jenison (1993) allowed listeners to adjust the starting frequency of the second of two disjoint FM-sweeps across a noise-occluder (N-occluder) to produce the strongest perception of continuity. Listeners' most often heard two sweeps separated by a N-occluder as optimally continuous when the sweeps were aligned along a trajectory of constant rate of change. While Kluender and Jenison (1992) discussed both the influence of extrapolation and interpolation on the perception of auditory continuity, their method did not allow extraction of the independent contribution of extrapolation without the influence of interpolation. Both FM-sweeps were audible, and, conse-

quently, the possibility that listeners were interpolating between audible portions of each sweep cannot be discounted. In order to isolate the role of extrapolation, we measured changes in listeners' ability to detect frequencies falling on and in close proximity to a path predicted by a constant rate of frequency change induced by an FM-sweep/N-occluder combination. (A constant rate of frequency change will hereafter be referred to as frequency velocity with the implication that a sound increasing or decreasing in frequency across time invokes a spatially shifting pattern of response across the basilar membrane and subsequent areas of the auditory pathway that can be interpreted as having a velocity.) A modified probe-signal method (Schlauch and Hafter, 1991) was used to measure changes in listeners' detection of low-level pure-tone signals of variable frequencies following an FM-sweep/N-occluder cue. The probe-signal method measured change in listeners' threshold detectability to frequencies falling at, and surrounding, frequencies indicated by a preceding auditory stimulus. This enabled use of simple detection tasks to observe both the frequency locations and breadth (typically bandpassed in shape) of change in listeners' thresholds induced by the cue.

The indirect nature of the probe-signal method was crucial. It allowed measurement of which frequency region a listener was attending to along an implied path without relying on a listener's answer regarding where an estimated signal should appear. Rather, it implied a sensory change in the expectation for the location of the signal without direct input from the listener (Schlauch and Hafter, 1991; Hafter *et al.*, 1993; Wright and Dai, 1993). Therefore, unlike studies of interpolation between two disjoint FM-sweeps (Kluender and Jenison, 1992; Ciocca and Bregman, 1987), a listener here would have had to rely solely on extrapolation from a single sweep. While these two listening strategies need not be mutually exclusive, the purpose of the present work was to determine whether, in the frequency domain, listeners used frequency velocity to track how sounds change and whether this information could be used to extrapolate predictions of a sound's projected path, even in the presence of an occluder. It is well known that higher-order processes such as a listener's familiarity with the semantics of speech can influence the perceptual restoration of occluded content (Bashford and Warren, 1987b; Sivenon *et al.*, 2006). However, the aim of our studies was to see whether listeners showed a relative increase in detection of a low-level stimulus (hereafter referred to as sensitivity) along a projected path. This would imply the use of velocity tracking and emphasize the importance of a lower-level process where, potentially, changes in lower sensory processing contributed to the restoration of occluded content.

Three experiments were completed. Experiments 1 and 2 used a modified probe-signal method to measure detection of sounds falling along an implied path of constant-linear-velocity. Experiment 1 asked whether listeners were able to use velocity information to predict the frequency location of a sound following an N-occluder and further whether the duration of constant-velocity information preceding the interruption affected listeners' ability to do this task. Experiment 2 asked whether the duration across which listeners were

required to extrapolate affected their prediction, and, finally, Experiment 3 used suprathreshold tones to determine listeners' perceptions of optimally continuous trajectories in order to address the influence of extrapolation on the perceptual phenomenon of auditory continuity.

II. METHODS OF EXPERIMENTS 1 AND 2

A modified probe-signal method (Schlauch and Hafter, 1991) was used to measure listeners' ability to detect a low-level (90% detection) probe tone of varying frequency following an above-threshold FM-sweep/N-occluder combination. Studies using a modified probe-signal method have measured the detection of low-level pure-tones in background noise following cues of the same frequency or frequencies in close relational proximity (i.e., $\pm 10\%$ – 20%) (MacMillan and Schwartz, 1975; Wright and Dai, 1993; Schlauch and Hafter, 1991). Previously, a majority of experiments using probe-signal methods used pure-tone "cues" that were either the same as or bore a constant ratio to the expected frequency (Scharf *et al.*, 1987; Schlauch and Hafter, 1991; Hafter *et al.*, 1993; Wright and Dai, 1993). Results from these studies have demonstrated a significant increase in listeners' performance for frequencies near to the implied frequency of the cue relative to frequencies that were not. A paucity of experiments used cues that changed dynamically across time (Ebata *et al.*, 2001; Wright and Dai, 1998). However, the spectra of cues in these conditions also contained energy that directly stimulated the frequency region of the probe signal. In contrast, the present studies used cues that only contained spectral energy outside the frequency region of the probe signal and, instead, implied the frequency region of the probe solely through extrapolation of dynamic properties of the cue.

A. Stimulus and experimental design

1. General procedure

On each trial, listeners heard an FM-sweep followed by an N-occluder (exemplar conditions used in Experiment 1 are shown in Fig. 2) where the duration and starting frequency of the FM-sweep varied from trial to trial. This combination will be referred to as the FM-sweep/N-occluder cue. On half of the trials, the cue was followed immediately by a low-level 100-ms pure-tone. This combination will be referred to as a signal trial. The pure-tone was absent for the remaining 50% of the trials. This combination will be referred to as a noise trial. All signal and noise stimuli were presented in a continuous low-level background noise masker that was bandpass filtered between 125 and 5125 Hz with a spectrum level of 20 dB/Hz with regard to 20 μ Pa. Sixty-percent of the time the signal fell at the expected frequency, as implied by the combination of the velocity of the FM-sweep and the duration of the N-occluder (Fig. 2). For the remaining 40% of the trials, the signal fell half of the time at either $+10\%$ or -10% of the expected frequency and half of the time at either $+20\%$ or -20% of the expected frequency. The level of the signal varied as a function of frequency and was determined by the listener's 90% equal-detectability function (EDF, see Appendix A for description

FIG. 2. Stimulus design for Experiment 1. Solid lines represent cartoon exemplar stimuli used in Experiment 1. Stimuli 1–4 demonstrate both a rove on start frequency and initial-sweep duration.

of methods used to establish listeners' EDFs) (Green *et al.*, 1959) for detection of a pure-tone in low-level background noise (having spectral characteristics as previously described).

In a single-interval yes-no task, listeners responded "yes" if they detected a signal following the N-occluder and "no" if they did not detect a signal. Response was indicated through a mouse click on a graphical user interface. After making a response, listeners received immediate visual feedback indicating whether they were correct or incorrect. Listeners were given unlimited time to respond, and the subsequent trial began approximately 1 s following a response.

2. Trial presentation

A testing block consisted of 60 trials, and a listener heard between 40 and 80 blocks. For each stimulus, the cue duration and starting frequency were unique. Cue durations were roved in equal intervals between 300 and 600 ms so that the final distribution the listener heard across all blocks was uniform in linear space. Starting frequency was roved in logarithmically spaced equal intervals between 300 and 1200 Hz so that the final distribution the listener heard across all blocks was uniform in log space. Thirty rove levels on each parameter were used in a single block so that every trial contained a unique stimulus (each parameter value was presented as both a signal and noise stimulus creating 60 unique trials per block). However, for each block a different set of randomly sampled parameter values was drawn from uniform distributions spanning the rove ranges. For every

listener, at least 2400 unique stimuli were used in testing that, when analyzed across multiple blocks, enabled a dense sampling of each parameter across the total rove range and prohibited the listener from learning any properties of the cue other than the intended velocity.

3. Specifications of the FM sweep

Haft^{er} *et al.* (1993) demonstrated that listeners could use the relationship between a cue and a signal of a musical fifth to increase detection of randomly presented tones in noise. While many varieties of cues have been shown to enhance detection, this study was one of few that measured a change in detection to a frequency region outside the range of the cue. For this reason, FM-sweeps used in the current study were linear rather than logarithmic to avoid listeners' use of a musical relationship between starting frequencies of the sweep and the frequency of the signal. While attention was taken to limit the potential correlation between starting frequency and signal frequency by roving the starting frequency, the duration of the FM-sweep, and the duration of the N-occluder (Experiment 2), the decision to use only upward FM-sweeps (rather than mixing upward and downward sweeps) necessitated some degree of correlation between the two variables. Nonetheless, use of a linear sweep prohibited listeners from using a musical relationship to make direct estimates of the signal frequency from the starting frequency. The overall range for the initial-sweep duration of 300–600 ms was chosen to be predominantly outside the temporal limits of energy integration (Green *et al.*, 1957; Stephens, 1973). This was done to avoid change in listeners' performance as a function of duration occurring as a result of absolute stimulus detection rather than as a change in the perception of velocity. The velocity of the FM-sweep ($\Delta f/\Delta t$, where f =frequency and t =duration) was always 3000 Hz/s. Furthermore, linear sweeps were judged to produce a strong experience of perceptual continuity in pilot listening by the first author and members of the laboratory.

4. Specifications of the noise occluder

The duration of the N-occluder in Experiment 1 was fixed at 125 ms. In Experiment 2, the duration of the N-occluder was roved across a range of 50–200 ms with occluder duration densely sampled on a trial-by-trial basis from a linear uniform distribution across this range. Figure 3 demonstrates exemplar conditions used in Experiment 2. For all experiments, Gaussian noise was filtered to have a flat spectrum between 200 Hz and 5 kHz, and the start and end of the N-occluder overlapped the cue and signal by 3.34 ms, respectively, to maximize perceptual continuity (Warren, 1999).

5. Establishment of stimulus levels

As previously described, the level of the signal was derived from the listener's 90% EDF (typically ranging between 36 and 39 dB SPL). The level of the FM-sweep cue was set to be 10 dB higher. The level of the N-occluder was held constant for each listener during a specified experiment. This level was set to be 15 dB higher than the level of the

FIG. 3. Stimulus design for Experiment 2. Solid lines represent cartoon exemplar stimuli used in Experiment 2. Stimuli 1–4 demonstrate a rove on start frequency, initial-sweep duration, and the duration of the N-occluder.

signal for the median expected frequency in the range of frequency roves (approximately 5 dB higher than the FM-sweep). In Experiment 1, the lowest potential expected frequency was for a starting frequency of 300 Hz and a sweep duration of 300 ms, and the highest potential expected frequency was for a starting frequency of 1200 Hz and a sweep duration of 600 ms. Thus, the minimum and maximum expected frequencies were 1745 and 3665, respectively, and the median expected frequency was 2705 Hz. Therefore, the level of the N-occluder was 15 dB higher than the 90% threshold level for a pure-tone of 2705 Hz for each listener. Listeners' EDFs were computed (due to the use of the continuous masker) to account for a well-known frequency dependence of signal-to-noise ratios under conditions of simultaneous masking. In contrast, while some degree of forward masking was anticipated, the relative effects of frequency-dependent forward masking across stimuli were expected to be small.

After establishing the relative levels of the cue, N-occluder, and signal for each stimulus, RMS averages of the stimuli were computed before presentation. The 90% threshold levels used to establish the relative stimulus levels were derived from a listener's detection of a pure-tone signal in background noise. Clearly, the N-occluder introduced some amount of forward masking to the detection of the signal so that listeners' performance was not expected to be 90%. In testing, it was acceptable if listeners' overall performance in the experimental blocks was near to 80%. If the listener consistently performed well below 80%, the overall

stimulus level was increased by 1 dB. This increase meant that all ratios of the FM-sweep, N-occluder, and signal were maintained, but the entire rms averaged stimulus was increased.

6. General apparatus

All stimuli were generated digitally with a sampling rate of 50 kHz. Stimulus presentation was through a locally made 16 bit D/A converter, Crown D-75 amplifier, and Stax (SR5) electrostatic headphones. The spectrum of the headphones was flat across the frequency range used in the described experiments. The background noise was produced through a locally made random-noise generator. Its spectrum was bandpass filtered through a Kemo Dual Variable Filter Type VBF8.

B. Analysis methods

1. Establishment of comparison parameters

To establish a quantitative measure for estimating the location and width of the frequency region where listeners were optimally sensitive, all data from Experiments 1 and 2 were fitted with a form of rounded-exponential function. Previously, [Dai et al. \(1991\)](#) suggested that changes in listeners' sensitivity to signal frequency locations during a probe-signal task could be represented as changes in the relative attenuation of the probe frequencies to the expected frequencies. This assumption has been used and extended to interpret shifts in detectability in terms of a listener's attentional filter or measured "listening band" ([Schlauch and Hafter, 1991](#); [Hafter et al., 1993](#); [Moore et al., 1996](#)). Relative change in the shape of the best-fitting function to performance data collected through the probe-signal method was converted to decibels and approximated with the following equation:

$$W(x) = A \left(1 + \frac{p(x-g)}{g} \right) e^{p(x-g)/g}, \quad (1)$$

where W was the transfer function of the listening band, p was the bandwidth, g was the relative center of the band, and A was the non-normalized performance gain of the listener. Historically, studies of the modified probe-signal method have had reason to assume that the listening band would be centered on a single specified frequency and assumed that maximum sensitivity would be to signals at, and closely surrounding, the cued frequency. However, in the current studies, no assumptions were made regarding the predicted center of the listening band, and, instead, the estimated center frequency of the band was derived from the measured probe-signal data. In the visual system, an underestimation in the predicted path of an object's motion has been shown following increasingly longer periods of extrapolation ([Peterken et al., 1991](#)). If a similar underestimation exists in prediction of a sound's frequency trajectory across time, the location of the estimated center frequency cannot be assumed. Thus, the parameter g in the symmetric exponential function was treated as a free parameter. A third free parameter was included in the model to allow a measure of the relative gain of the system (performance level) across conditions. Therefore,

TABLE I. The range and resolution used to estimate the values of parameters p , g , and A in the least-squares optimization that was performed are shown in Table I. An initial coarse search spanning a broader range was computed to determine the preliminary location of the local optimum. A second cost function using the values in Table I was then computed at a high sampling rate to span the location of the preliminary estimation of the local optimum.

	Lower bound	Upper bound	Sampling unit
p	3	15	0.1
g	0.8	1.2	0.01
A	1	5	0.01

Eq. (1) was used in the present analyses to fit a least-squares approximation to a symmetrical rounded-exponential function in the following manner:

For each listener, values of d' ([Green and Swets, 1966](#)) measured in the single-interval probe-signal task were converted to decibels using the listener's psychometric functions for detecting a pure-tone in background noise. The psychometric functions were approximated from data collected during a two-interval-forced-choice (2IFC) task (described in Appendix A). In order to compare the d' values from the single-interval task to the corresponding decibel values of the psychometric function, all scores recorded in the probe-signal conditions were converted to percent correct using the following equations. Equation (2) converts the d' value obtained in the single-interval task used in the probe-signal studies to a comparable d' value for a 2IFC task ([Wickens, 2002](#), p. 104).

$$d'_{2IFC} = d' \sqrt{2}. \quad (2)$$

Equation (3) converts the rescaled d' value into a comparable percent-correct score ([Wickens, 2002](#), p. 97)

$$P_C = \Phi(d'_{2IFC} / \sqrt{2}), \quad (3)$$

where Φ represents the cumulative distribution function of a Gaussian normal $N(0, 1)$.

Once percent-correct scores for detection were obtained, listeners' psychometric functions were used to convert from percent correct into decibel values relative to a listener's 80% detection level. This was done for each value of d' for each listener prior to any averaging. Individual scores in decibels were then averaged across all listeners.

For each condition, the averaged data were fitted with Eq. (1) to estimate the three free parameters: p , g , and A , and a least-squares optimization was performed to compute a cost function on each parameter for values within the following ranges and resolutions (Table I).

The rounded-exponential function was used only as a rough estimate of the listening band that enabled measurement of changes in center frequency and width of the listening region across conditions. Further, introduction of roving parameters in the current studies was expected to increase the variance in listeners' estimation ([Drennan and Watson, 2001](#)). This inclusion was absolutely necessary to exclude listeners' potential use of cues besides velocity to optimally detect the location of the signal; however, an increase in the

variance of a listener's listening region was expected that, in turn, would have produced an overestimation of the parameter p .

2. Significance testing for individual parameters

A Monte Carlo randomization test (Manley, 1997) was performed to establish a measure of relative change between testing conditions and to test the null hypothesis that observed condition differences in the estimated parameters: p , g , and A were due to the sampling error rather than true differences in the underlying populations. Details of the randomization analysis used are described in Appendix B.

III. EXPERIMENT 1

Experiment 1 measured listeners' ability to use velocity information implied by a constant-velocity FM-sweep to extrapolate the frequency of a signal following an interrupting N-occluder. In Experiment 1, both the starting frequency and the duration of the FM-sweep were roved (as described in Sec. II). The duration of the N-occluder was always a constant 125 ms.

A. Subjects

Four normal-hearing listeners participated in the experiment. All were undergraduate university students between the ages of 18 and 24 who received monetary compensation for their listening time.

B. Results

Following an N-occluder, listeners show increased sensitivity to frequencies that fall along an implied path of constant-linear velocity. Further, listeners more accurately use velocity information to extrapolate the trajectory, showing decreased variance and bias in the width and position of the listening band, when presented with longer duration FM cues (corresponding to longer samples of constant velocity).

Figure 4(a) shows averaged d' values of four listeners for two ranges of cue duration. Figure 4(b) shows least-squares approximations to the averaged data of the four listeners. Each listener's d' values are transformed into the corresponding decibel values through their psychometric functions for detecting a pure-tone in noise. The estimated center (g) for the longer durations (450–600 ms, $g=0.97$) is significantly closer to the expected frequency value (1) than the estimated center (g) for the shorter durations (300–450 ms, $g=0.93$), $p<0.001$. The width of the estimated function is significantly narrower for the longer duration sweep range than for the shorter duration sweep range, $p=9.5$ and $p=6.8$, respectively, $p<0.05$ (where a higher value of p corresponds to a narrower estimated band). A narrower width corresponds to decreased variance and increased accuracy in use of velocity information to predict the frequency location of the signal. Additionally, listeners show a shift toward a more conservative criterion value in conditions with longer initial-sweep durations [$\lambda=1.1119$ (300–450 ms) and $\lambda=0.8455$ (450–600 ms)].

(a)

Average d' Values – Experiment 1

Normalized Deviation from Expected Frequency

(b)

Exponential Function Estimates – Experiment 1

Normalized Deviation from Expected Frequency

FIG. 4. (a) d' values for averaged data of four listeners as a function of normalized deviation from the expected frequency. The dotted line and dashed lines represent data for initial-sweep durations ranging between 300–450 ms and 450–600 ms, respectively. (b) Fits to a rounded-exponential function for the same initial-sweep durations. Actual data used to generate the fit are indicated for the 300–450 ms condition by open circles, and for the 450–600 ms condition by crosses. All dB levels are across listener averages taken from each listener's 80% EDF as determined by their psychometric function, and all between subject error is accounted for in the statistical analyses that were performed.

IV. EXPERIMENT 2

Results of Experiment 1 show that velocity information can be used to alter sensitivity along a projected path. Accuracy in using this information increased as the duration of the cue increased. However, it is unclear whether accuracy in use of velocity information is dependent on the duration across which a prediction is made. As discussed earlier, it is possible that listeners' use of velocity information may not produce a linear extrapolation of frequency across an indefinite time interval (previously observed in studies of visual and auditory temporal predictions, Peterken *et al.*, 1991; Szelag *et al.*, 2002). Experiment 2 measured listeners' ability to use velocity information across increasingly longer periods of extrapolation. In addition to roving the initial start frequency

and FM-sweep duration (450–600 ms), the duration of the N-occluder also was varied (50–200 ms) per trial to introduce variability in the period listeners were required to extrapolate. (Exemplary conditions are shown in Fig. 3.) Normalized center frequency, bandwidth, and non-normalized performance gain of the measured listening band were assessed to determine dependence of the listening region on the duration of extrapolation.

A. Subjects

Three normal-hearing listeners participated in the experiment. All three participated in Experiment 1.

B. Results

Across increased occluder duration (and corresponding periods of extrapolation), listeners continue to use velocity information to predict the frequency of a sound following the occluder. However, listeners show a downward shift in the location of the listening region relative to a trajectory of constant-linear velocity as the duration of the occluder increases. In contrast to results of Experiment 1, this shift is not coupled with an increase in the width of the listening band suggesting that the effective use of velocity information does not change but, instead, points listeners to a less accurate position.

Figure 5(a) shows averaged d' values of three listeners. Figure 5(b) shows least-squares approximations to the averaged data for each range. The estimated values for each parameter for each range of occluder duration are shown in Table II. The estimated-center-frequency of the listening band decreases as the duration of the occluder increases, $g = 1.01$, $g = 0.95$, and $g = 0.92$ for the 50–100 ms (S), 100–150 ms (M), and 150–200 ms (L) ranges, respectively. A comparison between the S and M conditions is significant with $p < 0.01$, whereas a comparison between the M and L conditions is not ($p = 0.08$). Nonetheless, a downward trend in the estimated-center-frequency with increased occluder duration is evident.

The estimated bandwidths of the optimal fits for all three ranges of occlusion are similar. The change in bandwidth is not significant for the S and M conditions, $p = (7.2)$ and $p = (6.4)$, respectively, $p > 0.05$. However, the change in bandwidth for the M and L conditions is marginally significant ($p = 8.7$ for L), $p < 0.05$. It is important to note that the measured listening band in the L condition is not significantly wider than either the M or S conditions (larger values of p corresponding to narrower estimated bandwidths).

Differences in measured detectability across the three (S, M, and L) conditions are marked. Performance is significantly higher for conditions containing longer occluder durations [L in comparison to M, $A_L = (4.31)$ and $A_M = (3.42)$, $p < 0.01$; M in comparison to S, $A_S = (2.84)$, $p < 0.01$]. This result is addressed in Sec. III.

V. EXPERIMENT 3

Experiments 1 and 2 measured changes in detection of pure-tone signals falling on or within 20% of a frequency implied by a trajectory of constant-linear velocity. The re-

FIG. 5. (a) Averaged d' values for the data of three listeners as a function of normalized deviation from the expected frequency. The dash-dot, dashed, and dotted lines represent data for N-occluder durations ranging between 50–100 ms, 100–150 ms, and 150–200 ms, respectively. (b) Fits to a rounded-exponential function for the same N-occluder durations. Actual data are indicated for the 150–200 ms range by open circles, for the 100–150 ms condition by crosses, and for the 50–100 ms range by asterisks. All dB levels are across listener averages taken from each listener's 80% EDF as determined by their psychometric function, and all between subject error is accounted for in the statistical analyses that were performed.

sults of both experiments suggested that listeners were able to use this information to alter sensitivity along a predicted path. Further, results of Experiment 2 showed a downward shift in the estimated-center-frequency as a function of whether listeners limited their listening range by avoiding listening

TABLE II. Estimated values of parameters p , g , and A for Experiment 2 as determined using the rounded exponential function defined by Eq. (1).

	S 50–100 ms	M 100–150 ms	L 150–200 ms
p	7.2	6.4	8.7
g	1.01	0.95	0.92
A	2.8	3.4	4.3

for frequencies near the upper edge of the testing range or truly experienced a perceptual downward shift as the duration of extrapolation increased. To address this, in Experiment 3, listeners were asked to judge whether suprathreshold tones were above or below the frequency implied by the linear trajectory. This study allowed for both the effects of extrapolation and interpolation and enabled comparison between estimates of the center of the listening bands found through the detection studies of Experiments 1 and 2 to listeners' perceptual experience of a constant trajectory across a given occluder duration. If the previously observed downward shift in relative frequency was also observed in a subjective, suprathreshold testing condition, it would imply that the downward shift seen in Experiment 2 was not a product of listeners' listening strategies biasing them to listen away from the upper edge of the range. Rather, it would suggest that the observed shift in performance indicated a change in listeners' sensitivity for certain frequencies along a trajectory that, in turn, affected the conditions under which listeners' perceived the trajectory to be continuous.

A. Subjects

Three normal-hearing listeners participated in the experiment. All three listeners participated in Experiment 1. Two of the three listeners participated in Experiment 2.

B. Methods

1. Stimulus design

All stimulus design was identical to the descriptions in Sec. II except for the following changes. To reduce any possibility that the observed downward shift was specific to the previously tested velocity, the velocity of the initial FM-sweep was changed to 3400 Hz/s. The duration of all initial sweeps was a constant 500 ms, and the level of the pure-tone signal following the N-occluder was set to be equal to that of the initial FM-sweep. The frequency of the pure-tone signal was set to occur at the following percentages above and below the expected frequency: -25, -20, -15, -10, -5, +5, +10, +15, +20, +25. Unlike Experiments 1 and 2 which used a distribution of signal frequencies weighted toward the expected frequency, the distribution used in Experiment 3 was uniform around the expected frequency. No signals were included at the expected frequency. The starting frequency of the FM-sweep was roved, as described in Experiment 2. Three durations of N-occluder were presented randomly within a block: 50, 125, and 200 ms.

2. Procedure

In a single-interval task, listeners heard an FM-sweep followed by an N-occluder and a pure-tone. All were told that the pure-tone would never fall directly on the trajectory, but, rather, would be slightly above or below it. On each trial, listeners were required to assign a label of high or low to the tone following the occluder, even if the tone appeared to fall along the trajectory. Listeners did not receive feedback informing them of the correctness of their response. Prior to beginning testing, listeners performed training sessions using the most extreme percentages above and below the expected

FIG. 6. [(a)–(c)] For individual listeners, the probability of responding that a tone was below a continuous trajectory as a function of normalized deviation from the expected frequency. The dotted, dashed, and dash-dot curves plot data for the 200 ms, 125 ms, and 50 ms N-occluder durations, respectively. (d) Least-squares estimates of Gaussian-cumulative-distribution functions for the average of the data of listeners A, B, and C. The smooth curves indicate least-squares fits to the averaged data. Solid thin lines show actual averaged data to which the functions were fitted.

frequency. Training continued until successful completion of a block of 20 trials with 100% correct identification.

C. Results

As the length of the occluder increases, listeners consistently perceive lower frequencies relative to a trajectory of constant-linear velocity as falling along a path of constant rate of frequency change. Figures 6(a)–6(c) plot individual responses for three listeners. Each curve demonstrates the probability a listener responded that the tone was lower than a trajectory of constant velocity as a function of normalized frequency. Figure 6(d) shows least-squares estimates of a psychometric function fit to the averaged data of the three listeners. A Gaussian-cumulative-distribution function was used to compute the estimates. The solid thin lines overlaid on each curve represent the data used to derive the optimal fit.

A downward shift in the point-of-subjective-equality, at which listeners respond “low” 50% of the time, is consistent across all listeners for increasing N-occluder durations. The crosses in Fig. 7 plot, for the averaged data, the estimated normalized frequency at which listeners respond low 50% of the time as a function of N-occluder duration. This point-of-subjective-equality is compared to the solid circles in Fig. 7 which represent the estimated-center-frequencies observed in Experiment 2. (Additionally, while listener C shows a clear upward bias [50 ms curve shown in Fig. 6(c)] with respect to listeners A and B, a downward trend with increased occluder duration is evident.) The similarity between the points-of-subjective-equality measured in Experiment 3 and the estimated-center-frequencies measured in Experiment 2 (Fig.

FIG. 7. Closed circles indicate points at which listeners respond 50% of the time that the signal tone falls below a perceptually continuous trajectory for levels of increasingly longer durations of N-occluder duration. Data are extracted from the estimated Gaussian-cumulative-distribution functions fit to the average data of three subjects, as shown in Fig. 6(d). The point-of-subjective-equality shifts downward from the expected frequency (defined as a trajectory of constant-linear increase in frequency equal to a normalized value of 1 and indicated by the solid line) as the duration of the N-occluder increases. Crosses indicate the frequency regions of maximal sensitivity (estimated-center-frequencies) as a function of the mean N-occluder duration for the ranges measured in Experiment 2 (i.e., 75 ms for an occluder range of 50–100 ms).

7) suggests that changes in listeners' sensitivity to frequencies along an implied path may be directly reflected in their perceptual experience of a continuous trajectory.

VI. GENERAL DISCUSSION

The present experiments demonstrate listeners' use of velocity information to increase sensitivity along an extrapolated path. Consistent with studies of visual perception (Petersen *et al.*, 1991), listeners show increased accuracy in the use of velocity information with extended exposure to a constant-velocity cue (Experiment 1). That is, as the duration of constant-velocity information increases, the center frequency of the listening band shifts closer to the expected frequency and the width of the listening band narrows.

Additionally, the accuracy of listeners' use of velocity information to extrapolate the path of the changing sound is dependent on the duration of extrapolation (Experiment 2). During conditions of shorter occlusion, listeners' extrapolated estimates closely match the expected frequency (as predicted by extrapolation of a trajectory of constant-linear-velocity). In comparison, during conditions of longer occlusion, listeners' normalized estimates of the center frequency of the listening region shift downward relative to the expected frequency. This downward shift across longer occluder durations is not coupled with an increase in the bandwidth of the listening region indicating that listeners are not simply listening across a larger frequency region as the duration, and corresponding range of frequency extrapolation (Δf), increases. Rather, it indicates that listeners are, instead, listening across a comparable frequency region centered at a lower frequency than the expected frequency would predict.

Finally, results of Experiment 3 show that the downward shift in estimated center frequency corresponds to listeners'

perception of the frequency at which they optimally experience a continuous trajectory rather than the result of biasing listening strategies to optimize performance. While previous studies have seen a downward shift (Kluender and Jenison, 1992; Pollack, 1977), results of Experiments 1 and 2 are the first to show that a downward shift in the perception of an optimally aligned trajectory does not depend on what follows the interruption, but, rather, arises from changes in threshold sensitivity of the system during the period of extrapolation. This change leads listeners to expect a sound to reappear at a lower frequency than extrapolation of constant-linear-velocity would predict, and it may represent a direct influence of changes in the sensory system on the perceptual representation of dynamic sounds across an interruption.

A. Discussion: Effects of noise-occluder duration on performance gain (Experiment 2)

Although not directly relevant to the primary conclusions and discussion of the present studies, the counterintuitive increase in performance for longer occluder durations seen in Experiment 2 [Fig. 5(b)] warrants discussion. While a direct answer for this difference in performance is not obvious, one explanation may be the result of effects of forward masking from the initial FM-sweep on information contained within a theoretical temporal integration window with respect to the signal. All signals were 100 ms long. The effects of forward masking due to the N-occluder are expected to increase as the duration of the occluder increases (Kidd and Feth, 1982). This did not happen, and, instead, listeners' performance improved in conditions with longer occluder durations, suggesting a property of the stimulus other than the noise occluder was influencing the loudness of the signal. Rather, it may be the case that signal detection in Experiment 2 is affected by forward masking introduced by the ending portion of the initial FM-sweep falling within the integrated window of the signal. In studies of the effects of the duration of a forward masker on listeners' increment detection, results of Schlauch *et al.* (1997) were also inconsistent with previous studies of forward masking. They found that shorter duration maskers produced more masking than longer duration maskers. Further, their results, as well as those seen by Oxenham and Plack (2000), suggest that the degree of masking is dependent on the similarity of the masker to the signal, with greater similarity producing conditions of greater masking. In the current experiments, trials with shorter occluder durations created conditions where both the N-occluder as well as the ending portion of the initial FM-sweep would likely be integrated within the same temporal window as the signal. This led the ending portion of the initial sweep to contribute to the overall masking of the signal in some conditions and not in others. For example, a trial containing the shortest N-occluder duration (50 ms) and beginning at the highest initial-sweep start frequency (1200 Hz) has an expected frequency of 3150 Hz. The estimated equal-rectangular-bandwidth [ERB, as computed using Eq. (4)] for this frequency is 365 Hz.

$$\text{ERB}_N = 24.7(4.37F + 1). \quad (4)$$

Therefore, for a condition with an FM-sweep rate of 3000 Hz/s, 72 ms of the initial FM-sweep, traversing a 215 Hz region, would fall within the ERB of the signal frequency and likely contribute to the effective loudness of the signal. This is in addition to any forward masking effects introduced by the 50-ms N-occluder. While for all shorter duration occluders it is true that some portion of the initial FM-sweep falls within the ERB of the expected frequency (and is assumed to contribute to its effective loudness), it would not be true for signals following longer duration occluders. Therefore, the amount the initial FM-sweep contributes to the effective loudness of any signal is a function of both the initial start frequency and the occluder duration, with the shortest occluder durations most affected—having both the longest portion of initial FM-sweep included in their estimated temporal window and the frequency content of the initial FM-sweep most near (and potentially most similar) to the expected signal frequency. Assuming a temporal integration window of approximately 200 ms, both the S and M conditions would have integrated effects of the end of the initial FM-sweep on the signal (with the greatest effects present in the S condition). In contrast, the entire signal duration in L conditions would be free of the influence of the initial FM-sweep, leaving only the effects of the N-occluder to alter the effective loudness of the signal. While this is only a proposed explanation for the change in performance as a function of N-occluder duration, it is an intrinsic property of the stimulus set that may have affected overall performance (A). However, all three parameters, p , g , and A , are estimated independently and the estimates of p and g would not be influenced by the observed change in performance.

B. Proposed explanation for downward shift in estimated center frequency

A simple process is offered to explain the observed downward shift in the estimated-center-frequency as a function of the N-occluder duration. A continuous representation of velocity information would require computation of the instantaneous time derivative of the frequency (df/dt) across the duration of the changing sound. Results of Experiment 1 demonstrate that use of velocity information continues to increase in accuracy (narrowing of the listening band and a decrease in error of the estimated-center-frequency away from the expected frequency) beyond a duration of 300 ms. This suggests that the instantaneous velocity estimate is a function of past velocities where information is integrated or averaged across a longer temporal window.

Figure 8 shows the estimated-center-frequency as a function of time for a trajectory of constant-linear velocity. Plotted with this is the predicted-center-frequency following application of a finite impulse response (FIR) exponentially-weighted moving average (windowed at 365 ms) to the instantaneous velocity. A 365-ms window was applied as an estimate of the window width across which velocity information is averaged based on the results of Experiment 1 and optimal prediction of the results of Experiment 2. (Further analysis of data from Experiment 1 using initial-sweep dura-

FIG. 8. The solid line indicates the expected center frequency for exact linear extrapolation of velocity information of a 500-ms FM-sweep with a start frequency of 400 Hz and a velocity of 3000 Hz/s through a broad-band N-occluder. The dotted line represents the predicted-center-frequency with an FIR exponentially-weighted moving average with a window width of 365 ms applied to the instantaneous values of df/dt . The dotted vertical line indicates the onset of predicted values, and the dashed vertical line indicates the onset of the N-occluder. Crosses indicate estimated-center-frequencies obtained in Experiment 2, and asterisks indicate points-of-subjective-equality obtained in Experiment 3.

tions ranging between 375 and 525 ms showed minimal change in convergence in comparison to 450–600 ms conditions, suggesting that the window width is approximately between 350 and 400 ms.) The FIR filter is defined by

$$\begin{aligned} 0, & \quad t < 0, \\ ce^{-0.35t}, & \quad 0 < t < 365 \text{ ms}, \\ 0, & \quad t > 365 \text{ ms}, \end{aligned} \quad (5)$$

where c is a constant of normalization. It has been applied under the assumption that the instantaneous velocity becomes zero at the onset of the N-occluder. While it is clear that the instantaneous df/dt for a broad-band noise is not zero, when averaged across a temporal window the average df/dt quickly converges to zero. Crosses shown in Fig. 8 indicate estimated-center-frequencies obtained in Experiment 2, and asterisks indicate points-of-subjective-equality obtained in Experiment 3.

The process used to produce the estimates shown in Fig. 8 is a simple representation of how the auditory system may be transforming information. Additional studies manipulating the local time derivative of the frequency around the onset of the N-occluder as well as further manipulations of the initial-sweep duration in order to observe a more accurate value of the proposed window size and weighting of information across the duration of the window are needed to enable a more accurate model of this process. Nonetheless, it is striking how closely the low-pass filtered estimates are to the measured data.

It is also necessary to consider that the auditory system may represent velocity logarithmically or linearly in units of ERBs (Bark scale). If this is true, an upward linear sweep would be interpreted as decelerating. A representation of a

decelerating signal would predict a lower estimated-center-frequency across longer duration noise occlusions. However, previous studies considering the optimal representation of a signal of changing frequency suggest that this is not the case. Results show that a logarithmic or Bark scale representation of frequency change is not a better predictor than a linear scale of the optimal frequency that continues a trajectory of constant change (Pollack, 1977; Kluender and Jenison, 1992). Rather, downward shifts in the optimal frequency were observed for signals changing at constant rates within all three scales (linear, logarithmic, and Bark). Pollack found a similar ratio for the downward shift in the optimal frequency relative to the estimated frequency for both linear and logarithmic trajectories. This would represent a considerably greater underestimation in absolute frequency for the logarithmic stimulus relative to the linear stimulus. The consistent shift across all scales implicates a more general process leading to the downward shift for increasing trajectories and would be predicted by the weighted moving average proposed above. Such a process would predict an underestimation of the expected frequency across longer periods of prediction regardless of the scale of transformation. In addition, while all frequency trajectories used in the present studies were increasing across time, the proposed model also predicts an underestimation in prediction for frequency trajectories decreasing across time. In such conditions, listeners would show an increase rather than a decrease in the measured optimal frequency of continuation. Data from Pollack (1977) using series of pulsed tones support this prediction.

Further, it is important to consider that the well-known logarithmic representation of frequency in the auditory system (Moore, 1997) does not require that velocity also be represented logarithmically. Carre (2004) demonstrated that first formant transitions of natural speech signals are more or less linear in an absolute-frequency scale rather than a log-frequency scale. The presence of roughly linear acoustic changes in prominent natural sounds such as speech implies an important role for linear representations of frequency change in the auditory pathway.

C. Prediction of frequency change in the absence of noise

Results from previous studies (Kluender and Jenison, 1992; Pollack, 1977) suggest that the observed downward shift in center frequency with increased duration of the N-occluder may occur in prediction regardless of whether masking energy is present during the period of interruption. Therefore, it is conceivable that the observed changes in sensitivity for tones falling along the projected path, as shown in Experiments 1 and 2, would still arise if an intervening silence replaced the occluding noise. If, within the auditory system, df/dt is averaged or integrated across a time window greater than 300 ms, one would predict similar trends in the estimated-center-frequency regardless of whether a silence or noise interval is present. However, it is difficult to ignore the marked difference in the perception of a sound with an interrupting silence in comparison to an interrupting N-occluder. One possibility is that prediction does not require stimulation, but when sufficient masking energy is

present, sensitivity to frequencies falling along the predicted path is elevated. A conceivable scenario could be that global stimulation in the auditory periphery introduced by the N-occluder is unevenly represented at higher auditory areas such that frequencies falling along the projected path are amplified relative to the background, in turn, enhancing and shaping the perception of continuity.

D. Neural correlates of velocity coding and temporal induction

Physiological studies of the visual system have observed cells in the primate parietal cortex that show greater response to inferred motion during temporary occlusion of a moving stimulus (Assad and Maunsell, 1995). In comparison, it is currently unclear whether auditory continuity for dynamically changing sounds is similarly encoded in the auditory cortex at the level of a single neuron. Studies of continuity (using a constant-frequency tone) in the primary auditory cortex (AI) have found cells showing similar response patterns to both a truly continuous sound and a perceptually continuous sound (Petkov *et al.*, 2007; Sugita, 1997). While these studies have independently demonstrated sensitivity to conditions of continuity and discontinuity, the neural correlates underlying the robust perception of continuity for a dynamically changing stimulus may be quite different. In fact, encoding of velocity for predictive use as suggested by the current studies may lead to a very different change at the level of a single cell independent of additional processing necessary for the experience of perceptual continuity. For example, study of the encoding of spatial velocity in the optic tectum of the barn owl finds, for cells responsive to interaural timing differences, context-dependent receptive field shifts suggesting predictive anticipation of future stimulus locations that are dependent on the speed and direction of a moving source (Witten *et al.*, 2006). It is probable that cells sensitive to rate-of-change in frequency space (presently referred to as velocity) may show a similar predictive change in receptive field at some level of the auditory system. Continued convergence and increased predictive accuracy of velocity information beyond a duration of 300 ms (as demonstrated in Experiment 1) imply that physiological circuits involved in this process would likely be located high in the auditory pathway (i.e., beyond AI).

E. Summary

The present experiments demonstrate that listeners do use velocity information to track the change of a dynamic sound, and further that increased sensitivity along an extrapolated path significantly influences listeners' predicted and perceived trajectories of constant change through an occluder. These findings suggest that use of velocity information in the extrapolation of the path of a changing sound may be a fundamental aspect of auditory processing in the presence of spectrally and temporally dynamic sounds such as speech and music. In particular, these findings offer strong implications for the underlying physiology that enables processing of both speech and music in a noisy environment. At

this time, further studies of both the psychophysical and physiological correlates of the phenomenon are warranted.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Thomas Wickens for discussion and guidance in appropriate quantitative techniques, and Dr. Frederic Theunissen, Dr. David Wessel, Dr. Pierre Divenyi, Dr. Anne-Marie Bonnel, and anonymous reviewers for thoughtful opinions and editorial comments of the present research. P.A.C.C. is now employed by the Johns Hopkins School of Medicine.

APPENDIX A: ESTABLISHING EQUAL DETECTABILITY

During the typical single-interval task used in testing, levels of all signals were set to yield equal percentages of correct detection as obtained from a 2IFC task for detecting a pure-tone in background noise. In order to equate the signal tones in detectability prior to probe-signal sessions, listeners' thresholds were measured for detecting a pure-tone of varying frequency in a low-level background noise as previously described. A 2IFC, three-down one-up, tracking method (Levitt, 1971) was used to establish listeners' 79.4% threshold for detecting a 100-ms signal of a specified frequency. Each listener completed five blocks of the tracking 2IFC task at 400, 800, 1000, 1500, 2000, 2500, 3000, and 4000 Hz.

For each subject at each frequency, the average level of the five blocks was considered their base line testing level. Following tracking sessions, listeners completed three to five blocks of 100 trials in a 2IFC detection task at each of three different levels. The levels for each listener were as follows: base line (as obtained from the average of their tracking block), +1.5 or +2 dB from their base line, and -1.5 or -2 dB from their base line. Whether they completed blocks at ± 1.5 or ± 2 dB depended on their initial performance at ± 2 dB. Desired performance was near 90% at +2 dB. If performance at this level was generally above 95%, the level was lowered by 0.5 dB in order to measure points where the slope of the listener's psychometric function was expected to be steeper. Similarly, the level was raised following performance below 65% at -2 dB. This set was completed for all eight frequencies.

Performance at these levels was then used to fit a least-squares optimization of a psychometric function (using a Gaussian-cumulative-distribution function) to a listener's performance as a function of stimulus intensity for each frequency. The slope of the psychometric function for each listener was assumed to be constant across frequency. All fits were made with a single free parameter for each listener's set, the decibel level of the function's mean (relative offset in decibels). Following this, a least-squares regression line was approximated for each listener as a function of decibel level versus linear frequency (Green *et al.*, 1959). This approximation will be referred to as the listener's EDF. Throughout all experiments, for each trial, the signal level of the expected frequency was imputed from the point on the regression line of the listener's 90% EDF corresponding to the specified testing frequency.

APPENDIX B: RANDOMIZATION TEST FOR SIGNIFICANCE

The following Monte Carlo randomization test was performed to establish a measure of relative change between testing conditions and to test the null hypothesis that observed condition differences in the estimated parameters: p , g , and A were due to sampling error rather than true differences in the underlying populations.

For each comparison, all raw data files for each listener were combined on a trial-by-trial basis across the assessed conditions and stratified by percentage away from the expected frequency. In this way, all data collected at a fixed percentage of the expected frequency were combined only with data from the same percentage. During a single iteration, rows of data from the combined file for each listener were randomly permuted and a sample equal to one-half the size of the combined conditions was drawn without replacement from the combined file. A d' value was calculated for each stratified percentage level relative to the expected frequency: -20%, -10%, 0%, +10%, and +20% and will be referred to as a d' set. This process was repeated 1000 times to create a matrix of size 1000 by 5 containing 1000 sets of d' values computed by randomly drawing samples from the combined file.

The new sets of d' values were used to estimate the exponential function [Eq. (1)] in the following manner. On a single iteration, a single set of d' values was drawn from a permutation of the listener's set of d' values. This was done for each listener without replacement across the series of iterations. For each listener, data were converted to decibels using an estimate of the listener's psychometric function for detecting a pure-tone in noise. The decibel values were averaged for the group of listeners at each percentage value from the expected frequency. The averaged values were used to approximate a least-squares estimate of the exponential function with three free parameters: width, center frequency, and gain. One iteration yielded a single estimate of each parameter. This process was completed 1000 times to generate a file of size 1000 by 3 containing 1000 estimates of each parameter.

Each file of parameter estimates was then randomly divided into two groups. For each parameter, an estimated difference value, referred to as $d\hat{v}$, was computed by randomly drawing an estimated parameter value from each group and taking the absolute value of the difference. For example, if two center frequency estimates were drawn, 0.97 and 0.985, $d\hat{v}$ would be 0.015. This process was repeated for each parameter 1000 times to create a file of size 1000 by 3 consisting of estimates of $d\hat{v}$ for each parameter. These sets were then used to establish a measure of variability in the experimental data. For each condition being compared, we computed the difference value of the uncombined (and unpermuted) measured parameter values following fitting to the exponential function (shown in Table II and now referred to as $d\bar{v}$). For each parameter, the number of times that $d\hat{v}$ exceeded $d\bar{v}$ was recorded. This number provided a probability measure of the chance that the observed differences in the actual uncombined, unpermuted estimates were due to the sampling error rather than representing a true difference in

TABLE III. Individual listeners estimated value of d' —Experiments 1 and 2.

Experiment 1: Normalized deviation from expected frequency					
	0.8	0.9	1	1.1	1.2
A_{300_450}	1.5874	2.2718	1.5921	1.2846	0.5230
A_{450_600}	1.5468	1.8819	1.8041	1.1709	0.4080
B_{300_450}	1.2632	1.7456	1.7510	1.5217	1.5434
B_{450_600}	0.5743	1.3789	1.6113	1.5417	1.2782
C_{300_450}	1.8314	1.9348	1.8229	1.8314	1.4049
C_{450_600}	1.7905	1.9984	2.1500	1.8683	1.2747
D_{300_450}	1.7978	1.9346	1.7848	1.6362	1.6206
D_{450_600}	1.3987	1.8345	1.6313	1.6994	1.7475
Experiment 2: Normalized deviation from expected frequency					
	0.8	0.9	1	1.1	1.2
A_{50_100}	1.2069	1.8347	2.0710	1.5256	1.1701
A_{100_150}	1.5135	2.0437	1.9843	1.6460	0.7799
A_{150_200}	1.5573	2.3066	1.8135	1.1228	0.4798
C_{50_100}	1.1229	1.6413	1.5221	1.4768	1.3891
C_{100_150}	1.7879	1.8193	2.0161	1.6421	1.7151
C_{150_200}	1.7671	2.4272	1.8570	1.5314	1.3536
D_{50_100}	1.2945	1.4815	1.4551	1.6832	1.5462
D_{100_150}	1.4097	1.6822	1.5655	1.4056	1.4602
D_{150_200}	1.8395	1.9900	1.7319	1.9730	1.8358

the population. For example, in Experiment 1, the $d\bar{v}$ value for center frequency as a function of initial-sweep duration was 0.04 (the difference between the actual values 0.97 and 0.93). After combining data across initial-sweep duration, we found that the probability of observing $d\hat{v}$ of 0.04 or greater when the conditions were combined was 9/1000. It may then be said that the observed difference in center frequency as a function of initial-sweep duration was significant, $p < 0.01$.

APPENDIX C: INDIVIDUAL DATA: EXPERIMENTS 1 AND 2

Table III shows the data of Experiments 1 and 2.

Assad, J. A., and Maunsell, J. H. R. (1995). "Neuronal correlates of inferred motion in primate posterior parietal cortex," *Nature (London)* **373**, 518–521.

Bashford, J. A., Jr., and Warren, R. M. (1987b). "Multiple phonemic restorations follow the rules for auditory induction," *Percept. Psychophys.* **42**, 114–121.

Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Carre, R. (2004). "From an acoustic tube to speech production," *Speech Commun.* **42**, 227–240.

Churchland, M. M., Chou, I., and Lisberger, S. G. (2003). "Evidence for object permanence in the smooth-pursuit eye movements of monkeys," *J. Neurophysiol.* **90**, 2205–2218.

Ciocca, V., and Bregman, A. S. (1987). "Perceived continuity of gliding and steady-state tones through interrupting noise," *Percept. Psychophys.* **46**, 39–48.

Dai, H., Scharf, B., and Buus, S. (1991). "Effective attenuation of signals in noise under focused attention," *J. Acoust. Soc. Am.* **89**, 2837–2842.

Dannenbring, G. L. (1976). "Perceived auditory continuity with alternately rising and falling frequency transitions," *Can. J. Psychol.* **30**, 99–114.

Drennan, W. R., and Watson, C. S. (2001). "Sources of variation in profile analysis. II. Component spacing, dynamic changes, and roving level," *J. Acoust. Soc. Am.* **110**, 2498–2504.

Ebata, M., Miyazono, H., Kumamaru, K., and Chisaki, Y. (2001). "The

formation of attentional filters for a missing-fundamental complex tone and frequency-gliding tones," *Acoust. Sci. & Tech.* **22**, 401–406.

Green, D. M., Birdsall, T. G., and Tanner, W. P. (1957). "Signal detection as a function of signal intensity and duration," *J. Acoust. Soc. Am.* **29**, 523–531.

Green, D. M., McKey, M. J., and Licklider, J. C. R. (1959). "Detection of a pulsed sinusoid in noise as a function of frequency," *J. Acoust. Soc. Am.* **31**, 1446–1452.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Haft, E. R., Schlauch, R. S., and Tang, J. (1993). "Attending to auditory filters that were not stimulated directly," *J. Acoust. Soc. Am.* **94**, 743–747.

Houtgast, T. (1972). "Psychophysical evidence for lateral inhibition in hearing," *J. Acoust. Soc. Am.* **51**, 1885–1894.

Husain, F. T., Lozito, T. P., Ulloa, A., and Horwitz, B. (2005). "Investigating the neural basis of the auditory continuity illusion," *J. Cogn. Neurosci.* **17**, 1275–1292.

Kluender, K. R., and Jenison, R. L. (1992). "Effects of glide slope, noise intensity, and noise duration on the extrapolation of FM glides through noise," *Percept. Psychophys.* **51**, 231–238.

Levitt, H. (1971). "Transformed up-down methods in psychophysics," *J. Acoust. Soc. Am.* **49**(2), 467–477.

Macmillan, N. A., and Schwartz, M. (1975). "A probe-signal investigation of uncertain-frequency detection," *J. Acoust. Soc. Am.* **58**, 1051–1075.

Manley, B. F. J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology*, 2nd ed. (Chapman & Hall, New York/CRC, Boca Raton, FL).

Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.

Moore, B. C. J., Haft, E. R., and Glasberg, B. R. (1996). "The probe-signal method and auditory-filter shape: Results from normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **99**, 542–552.

Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (Academic, London).

Orban de Xivry, J., Bennet, S. J., Lefevre, P., and Barnes, G. R. (2006). "Evidence for synergy between saccades and smooth pursuit during transient target disappearance," *J. Neurophysiol.* **95**, 418–427.

Oxenham, J., and Plack, C. J. (2000). "Effects of masker frequency and duration in forward masking: further evidence for the influence of peripheral nonlinearity," *Hear. Res.* **150**, 258–266.

Pavel, M. C. H. and Stone, V. (1992). "Extrapolation of linear motion," *Vision Res.* **32**, 2177–2186.

Peterken, C., Brown, B., and Bowman, K. (1991). "Predicting the future position of a moving target," *Percept. Psychophys.* **20**, 5–16.

Petkov, C. I., O'Connor, K. N., and Sutter, M. L. (2007). "Encoding of illusory continuity in primary auditory cortex," *Neuron* **54**, 153–165.

Pollack, I. (1977). "Continuation of auditory frequency gradients across temporal breaks: The auditory Poggendorff," *Percept. Psychophys.* **21**, 563–568.

Rushton, S. K., and Wann, J. P. (1999). "Weighted combination of size and disparity: A computational model for timing a ball catch," *Nat. Neurosci.* **2**, 186–190.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., and Reeves, A. (1987). "Focused auditory attention and frequency selectivity," *Percept. Psychophys.* **42**, 215–223.

Schlauch, R. S., and Haft, E. R. (1991). "Listening bandwidths and frequency uncertainty in pure-tone signal detection," *J. Acoust. Soc. Am.* **90**, 1332–1339.

Sivonen, P., Maess, B., and Friederici, A. D. (2006). "Semantic retrieval of spoken words with an obliterated initial phoneme in a sentence context," *Neurosci. Lett.* **408**, 220–225.

Stephens, S. D. G. (1973). "Auditory temporal integration as a function of intensity," *J. Sound Vib.* **37**, 235–246.

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

Szelag, E., Kowalska, J., Rymarczyk, K., and Poppel, E. (2002). "Duration processing in children as determined by time reproduction: implications for a few seconds temporal window," *Acta Psychol.* **110**, 1–19.

Warren, R. M., Obusek, C. J., and Ackroff, J. M. (1972). "Auditory induction: Perceptual synthesis of absent sounds," *Science* **176**, 1149–1151.

Warren, R. M. (1999). *Auditory Perception: A New Analysis and Synthesis*, 1st ed. (Cambridge University Press, Cambridge).

Wickens, T. D. (2002). *Elementary Signal Detection Theory* (Oxford University Press, New York).

- Witten, I. B., Bergan, J. F., and Knudsen, E. I. (2006). "Dynamic shifts of the owl's auditory space map predict moving sound localization," *Nat. Neurosci.* **9**, 1439–1445.
- Wright, B. A., and Dai, H. (1993). "Detection of unexpected tones with short and long durations," *J. Acoust. Soc. Am.* **95**, 931–938.
- Wright, B. A., and Dai, H. (1998). "Detection of sinusoidal amplitude modulation at unexpected frequencies," *J. Acoust. Soc. Am.* **104**, 2991–2996.
- Yantis, S. (1995). "Perceived continuity of occluded visual objects," *Psychol. Sci.* **6**, 182–186.