Relative weights for frequency glide detection using narrowband noise^{a)}

Jinyu Qian^{b)} and Virginia M. Richards

Department of Psychology, 3401 Walnut Street, Suite 302C, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

(Received 18 July 2006; revised 8 February 2008; accepted 18 February 2008)

Narrowband noise stimuli were used to derive relative weights for detecting frequency glides in a yes/no procedure. One stimulus set was restricted to the duration of the glide. For the second stimulus set, there were fringe noise bands preceding and following the glide. For both sets, the center frequency of the linear glide was either fixed at 1000 Hz or randomly chosen on each trial from the range of 800–1200 Hz. Relative weights as a function of time were derived using a linear model and the linear classification method [A. Ahumada, *J. Vis.*, **2**, 121–131 (2002)]. Sensitivity was better for the fixed- than random-frequency conditions, and weight patterns from the random-frequency conditions were less reliable than those obtained from the fixed-frequency conditions. The magnitudes of the relative weights tended to be larger for the second half of the stimulus, suggesting that subjects paid more attention to the later than the earlier parts of stimuli. In the random-frequency conditions, the linear model failed to account for subjects' performance unless the stimuli were expressed in terms of relative changes in frequency rather than absolute frequency. © 2008 Acoustical Society of America. [DOI: 10.1121/1.2897030]

PACS number(s): 43.66.Fe [RAL]

Pages: 2692–2700

I. INTRODUCTION

Hartmann and Klein (1980) provided a signal-detection theoretical model to evaluate sensitivity to frequency modulation. As they point out, assuming the signal is known exactly, an ideal observer would cross correlate the stimulus with an internal template that is a copy of the signal to be detected. Noting that subjects can detect the modulation at signal levels smaller than those required in order to discriminate between forms of modulation (e.g., sinusoidal versus triangle frequency modulation), Hartmann and Klein suggested that the modulation form cannot be known exactly. Thus, they proposed that subjects sample the pitch of the stimulus, and difference the samples in order to detect the presence of modulation (see also Demany and Semal, 1989).

Here, the signal-detection theoretic approach noted by Hartmann and Klein is revisited with the goal of estimating the internal template for the detection of ascending frequency glides. The psychophysical estimate of the template is relative weights as a function of time, which will be referred to as the temporal weight pattern or TWP. The questions of interest include: What is the form of the TWP? And whether the TWPs can account for the human detection performance.

For the ideal observer model, the template for the detection of an ascending linear frequency glide is itself an ascending linear glide, with negative values for the first half and positive values for the second half of the template. In addition, the relative weights that form the template sum to

zero. For each yes/no trial, a decision variable would be the inner product of the template and the stimulus vectors. A decision can be made by comparing the resulting decision variable to a criterion. Therefore, for the ideal observer, the decision variable for a constant-frequency tone (no-glide tone) would be zero and that for an ascending tonal glide would be positive valued. Additionally, the sum-to-zero template would lead to decision variables independent of the central frequency (the frequency at the temporal midpoint of the tonal glide) variation across presentations. Therefore, the ideal template predicts no effect of randomizing the central frequency of the glide stimuli.

Human listeners might likewise make their decisions based on a comparison between an internal template and the stimulus, noting that the template may not be ideal. In addition to the work of Hartmann and Klein (1980), two additional psychophysical results suggest possible deviations of the listeners' actual internal templates from the ideal template. First, several investigators (Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999) have reported that the threshold for detecting a change in the frequency extent of tonal linear glides is higher when the central frequencies of the stimuli are random across presentations compared to fixed. Second, Moore and Sek (1998) have shown that the effects of central frequency randomization on thresholds are not always as large as expected if subjects depended on the frequencies solely at the stimulus onset or solely at the stimulus offset.

When the central frequency is drawn at random, neither the beginning nor the ending portion of the stimulus alone indicates whether or not a glide is present. The difference between these two portions, however, is invariant with respect to central frequency randomization. In terms of tem-

^{a)} Portions of this research were presented at the 151st meeting of the Acoustical Society of America [J. Acoust. Soc. Am. **119**, 3332 (2006)].

^{b)} Author to whom correspondence should be addressed. Electronic mail: qianj@sas.upenn.edu.

plates, “listening” only at the beginning of the stimulus implies non-zero relative weights only at the stimulus onset. As a result, the decision variable would reflect only the initial frequency. For the strategy of “differencing” the beginning and the ending frequencies, similar to the ideal template, the relative weights at the stimulus onset and offset would be of equal magnitude, but opposite sign. On one hand, the available psychophysical data suggest a form of differencing the beginning and ending portions of the stimuli because the listeners’ thresholds are superior to those expected if listeners were attending to, for example, just the initial frequency region of the stimuli (Moore and Sek, 1998). On the other hand, the differencing is not perfect; otherwise there would be no effect of frequency randomization at all. It is as though listeners fall between these two extremes, perhaps over weighting the initial or ending portions of the stimuli, but still differencing the frequencies of the initial and ending portions. Such a template would yield decision variables that are only somewhat influenced by central frequency randomization, consistent with the human literature.

Combining intuitions based on the ideal weight pattern and expectations based on psychophysical results, one might suggest the following expectation regarding the form of listeners’ internal templates for detecting ascending frequency glides. The estimated internal templates should follow a negative-to-positive form, and the magnitudes of the initial negative and the ending positive regions should not be equal.

This prediction regarding subjects’ internal templates suggests variations in the perceptual importance of different temporal segments of the glide stimuli. Since the beginning and ending frequencies are most informative for the detection of glides, they should be weighted heavily and the absolute values of the relative weights associated with these frequencies should be large. To emphasize the beginning and ending frequencies of glides, it is of interest to include conditions where two constant-frequency fringes were added at both ends of the glides to match the beginning and ending frequencies. The expected internal templates in these conditions should have large absolute values associated with the fringes.

In the current study, the internal templates for the detection of frequency glides were estimated when the central frequency of the stimuli were either fixed or randomly chosen across presentations. The stimuli were either frequency glides alone or glides with the beginning and the ending fringes depending on conditions.

The method used to estimate the template is based on Ahumada’s (2002) “classification image” approach, where the template is the pattern of relative weights as a function of time. The advantage of Ahumada’s approach for estimating relative weights, compared to other methods, is that the method can simply be applied to stimuli that vary continuously in time without requiring temporal “bins” (e.g., Berg and Green, 1990; Dai and Berg, 1992; Buus *et al.*, 1996; Turner *et al.*, 1998; Richards and Tang, 2006).

Ahumada’s (2002) approach is, roughly, as follows. In a Yes/No procedure the subject indicates either “yes signal” or “no signal.” Normally distributed random perturbations are added to the signal and no-signal stimuli. As an example, for

FIG. 1. Schematic spectrograms for the no-fringe-band (NFB, panel a) and fringe-band (FB, panel b) stimulus sets. The abscissa indicates time and the ordinate indicates frequency. Δf_c is the extent of the frequency glide. Solid lines show the carrier frequency of the narrowband of noise as a function of time and dashed lines delineate the 50 Hz noise bandwidth.

a two-dimensional visual stimulus, the perturbations are added on a pixel-by-pixel case. After data collection, the stimuli are categorized conditioned on listeners’ responses, yielding hit, false alarm, miss and correct rejection stimuli. Then, the perturbations for the trials that led to “yes” responses (hits and false alarms) are averaged (pixel-by-pixel) across trials, the perturbations for the trials that led to “no” responses (correct rejections and misses) are averaged across trials, and the latter image is subtracted from the former image. The resulting image provides an estimate of the subject’s internal template, or a “classification image.”

To apply this procedure to estimate the internal template for the detection of a frequency glide, perturbations in frequency as a function of time are required.¹ In the current experiment this was achieved by first applying a tonal frequency glide to a carrier tone and then multiplying the resultant glide by low-pass noise. The resulting narrow bands of noise have continuously changing frequencies. The difference, as a function of time, between the carrier frequency and the actual frequency of the narrow band of noise forms the perturbations in frequency as a function of time. Then, applying the method proposed by Ahumada (2002), the perturbations are used to estimate of the relative weights, which in turn provide a psychophysical estimate of the internal template.

II. METHODS

A. Stimuli and procedures

In a Yes/No paradigm, subjects indicated whether a 50-Hz-wide narrowband noise had a carrier with an ascending linear frequency glide (glide stimulus) or a constant-frequency carrier (no-glide stimulus). Two stimulus sets were tested, and both sets are shown in Fig. 1. The abscissa indicates time and the ordinate indicates frequency. The carrier frequency is shown using a solid line and the width of the narrow bands of noise are indicated using gray shading. The first stimulus set had a 120 ms duration, and the duration of the glide, when present, was also 120 ms. Figure 1(a) depicts these stimuli. For the second stimulus set, the glide-bearing 120 ms narrowband noise was preceded and followed by a 200 ms, 50 Hz band of noise [Fig. 1(b)]. The carrier frequencies of these “fringe bands” matched those of the initial and ending carriers of the glide-bearing portion of the band of noise. Thus, for a glide stimulus, the initial fringe band had a

carrier frequency lower than that of the ending fringe band, and for a no-glide stimulus, the initial and ending fringes had the same carriers. As a result, the listener could determine whether a glide was present using only the information in the fringe bands. This latter stimulus set will be referred to as the FB (fringe band) set, and the former stimulus set will be referred to as the NFB set (no fringe band). Note the NFB and FB stimulus sets parallel the “glide only” and “glide” conditions tested by Lyzenga *et al.* (2004), except that they used tonal glides.

The narrowband noise was generated digitally using the “quadrature multiplication” method (Bernstein and Trahiotis, 1992) in which two independent 25 Hz low-pass noises were multiplied by two tonal carriers, one in sine phase and the other in cosine phase, and the products summed to generate a 50-Hz-wide band of noise. For the NFB glide stimuli the 120 ms “sine” and “cosine” carriers varied linearly in frequency, whereas the no-glide stimuli had constant carrier frequencies. For the FB stimuli, stimulus generation was the same, except the carriers included two 200 ms constant carrier frequency fringes. All stimuli had 5 ms raised cosine rise/fall ramps at the stimulus onsets and offsets.

The central frequencies of the stimuli (the frequency at the temporal midpoint of the tonal carrier) were either fixed at 1000 Hz or randomly chosen from a uniform distribution ranging from 800 to 1200 Hz, yielding the *fixed-* or *random-frequency* conditions, respectively. In all, four conditions were tested: The 120 ms NFB stimuli with fixed (NFB-fixed) or random (NFB-random) central frequencies and the 520 ms FB stimuli with fixed (FB-fixed) or random (FB-random) central frequencies.

The extent of the tonal frequency glide, Δf_c , was the independent variable (see Fig. 1). The values of Δf_c ultimately tested were determined during two practice stages. In the first stage the threshold Δf_c was estimated using an adaptive tracking procedure in the FB-fixed condition (Levitt, 1971). This required approximately 1 h to complete. In the second stage, trials were blocked by condition and fixed Δf_c was used. The value of Δf_c was chosen with the goal of achieving performance of 75–80 percent correct. Depending on the subject and condition, this second practice stage required a total of 200–1300 trials.

Once the values of Δf_c were determined, data collection proceeded in blocks of 5 sets of 50 trials. Listeners rested between blocks. Listeners ran the FB conditions first. The FB-random and the FB-fixed conditions were intermixed such that blocks were tested using a random schedule. A total of five blocks for each condition were tested. Then the NFB conditions were tested using the same procedure. In all, a total of 1250 trials contributed to the final estimate of d' and to the formation of the TWPs.

Three young adults with normal hearing (<15 dB hearing level) served as subjects. The stimuli were presented diotically over Sennheiser HD410 SL headphones at an overall level of 57 dB sound pressure level with visual feedback following subjects' responses. Subjects were tested in a single-walled sound-attenuated booth in a quiet room. The stimuli were generated by TDT system II hardware controlled by MATLAB software.

FIG. 2. The procedure for deriving the temporal weight pattern is shown.

B. Computational methods

1. Temporal weight pattern

Figure 2 shows the primary steps used to derive the TWPs. The columns show the computational steps and the rows emphasize that the procedure was applied to each stimulus tested, grouped by listeners' responses (“yes glide” and “no glide”). First, for each stimulus, the instantaneous frequency and envelope were estimated using a Hilbert transform. For stimulus, $x(t)$, with a Hilbert transform $H[x(t)]$, the envelope, $e(t)$, is given by: $e(t) = \sqrt{x(t)^2 + H^2[x(t)]}$, and the instantaneous frequency, $f(t)$, is given by:

$$f(t) = \frac{1}{2\pi} \frac{\partial p(t)}{\partial t}, \quad (1)$$

where $p(t) = \tan^{-1}(H[x(t)]/x(t))$ (e.g., Green, 1998). Additionally, it is convenient to define the instantaneous carrier $f_c(t)$:

$$f_c(t) = f_0 - \frac{\Delta f_c}{2} g(t), \quad (2)$$

where f_0 is the central frequency of the stimulus, Δf_c is the extent of the frequency glide when present. For the no-glide stimuli, $g(t)$ has a constant value of zero over time. For the glide stimuli in NFB conditions, $g(t)$ increases linearly from -1.0 to $+1.0$. For the glide stimuli in FB conditions, $g(t)$ has the constant value of -1.0 for the initial fringe, linearly increased values from -1.0 to $+1.0$ for the glide portion and the constant value of $+1.0$ for the latter fringe.

Next, the *deviation*, $\Delta f(t)$, of the instantaneous frequency from the instantaneous carrier was calculated as:

$$\Delta f(t) = f(t) - f_c(t). \quad (3)$$

The deviations are depicted in the leftmost column of Fig. 2.

In the third step, the deviations were weighted using a variant of the squared-envelope-weighted average of instantaneous frequency (SEWAIF) of Anantharaman *et al.* (1993; see also Dai, 1993), a model developed to predict pitch estimates for dynamically changing stimuli. In parallel with the SEWAIF model, the deviations of the instantaneous frequency from the carrier were weighted by the squared instantaneous envelope and then divided by the mean value of the squared instantaneous envelope, $\bar{e^2(t)}$:

$$\Delta f_w(t) = \frac{e(t)^2 \Delta f(t)}{e(t)^2}, \quad (4)$$

The resulting *weighted deviations* are depicted in the second column of Fig. 2. The weighted deviations, $\Delta f_w(t)$, as a function of time formed the “perturbations” from the carrier frequency and allowed the estimation of the internal template or TWP using Ahumada’s (2002) approach.

To form the TWP, i.e., relative weights as a function of time, the weighted deviations ($\Delta f_w(t)$) were categorized in accordance with subjects’ responses (Ahumada, 2002). A “Yes” pattern was obtained by averaging the weighted deviations over all trials where the subject responded “Yes” the glide is present. Likewise, a “No” pattern was obtained by averaging the weighted deviations over the trials where the subject responded “No” glide. The ultimate TWP was subsequently formed by subtracting the “No” pattern from the “Yes” pattern, where the “Yes” and “No” patterns were weighted in proportion to the number of “Yes” and “No” responses² (final column of Fig. 2). Keep in mind that the estimated TWPs did not depend on the presence or absence of the glide because it was the deviations from the carrier frequency that were being averaged. However, the TWPs did depend on the responses, and therefore reflected subjects’ decision processes.

This procedure for estimating the relative weights (Ahumada, 2002) assumes that the perturbations ($\Delta f_w(t)$ in this application) added to the stimulus (the carrier frequency) are normally distributed and statistically independent. The assumption of normal distribution was reasonably met,³ but the assumption of independence was not. Evaluations of the autocorrelation functions of the $\Delta f_w(t)$ for several of the stimuli suggested consistent non-zero correlations for delays less than 40–50 ms. This represents a significant limitation in the current approach—changes in the estimated TWPs should be viewed with a conservative eye appreciating that there may be correlations in the TWPs that last 40 ms or so.

2. Model evaluation

The TWPs were evaluated from three perspectives. First, were the derived TWPs statistically discernable from zero? Second, was there sufficient information in the TWPs to account for subjects’ performance? Third, were the magnitudes of relative weights corresponding to the initial and ending portions of the stimuli approximately the same? The procedures used to address these questions will be addressed in turn.

a. Bootstrap estimate: Are the TWPs statistically different from zero? A bootstrap procedure (Efron and Tibshirani, 1993) was used to evaluate whether the obtained TWPs were statistically different from those patterns based on random responses. For each subject and condition 1250 weighted deviations ($\Delta f_w(t)$) were drawn, with replacement, from the 1250 trials encountered by the subject. Responses of “Yes” and “No” were randomly assigned to each stimulus using proportions consistent with the “Yes” and “No” responses of the subject. Then a TWP was generated. This simulated TWP provided an estimated function for the null hypothesis that the subject-based TWPs contained no infor-

mation concerning the internal template. The formation of a meaningful TWP requires that subjects’ responses be systematically influenced by the perturbations, a relationship not present when responses were randomly assigned. The bootstrap simulation was repeated 100 times. Across the 100 repetitions, the average TWP were near zero. For each point in time, 96% confidence intervals about the average were estimated. The confidence intervals for all points were then averaged across time to form a single 96% confidence interval for each TWP. To the degree the psychophysically derived TWPs exceed the 96% confidence intervals, it would indicate that the obtained TWPs reflected properties of the subjects’ responses (and so decision processes), not properties of the stimuli.

b. d' based on estimated TWPs. Values of d' based on TWPs (d'_{wgt}) were compared with the psychophysically obtained values of d' (d'_{obs}). The d'_{wgt} represents the best possible d'_{obs} subjects can reach if the derived TWPs were their exact internal templates on every trial, and there was no added “internal noise”. Should the d'_{obs} exceed the d'_{wgt} , the linear model (specified below by Eq. 5) would fail to account for the psychophysical data. An opposite pattern would indicate that the TWPs contained enough information to account for subjects’ performance. The approach used was similar to that used by Berg (1990).

The procedure used to estimate d'_{wgt} was as follows, and was repeated for each subject and condition. First, a decision variable β was computed for each trial. The decision variable was the inner product of the TWP (indicated as $W(t)$ below) and the weighted instantaneous frequency (i.e., $f_w(t) = \Delta f_w(t) + f_c(t)$) of the stimulus on that trial (trial i):

$$\beta_i = W(t) \cdot f_w(t)_i. \quad (5)$$

Equation (5) defines a linear detection model, in which the values of the decision variable β are a linear combination of the stimulus and the TWP. For a zero-mean TWP that resembled the form of the tonal carrier of the glide stimulus [see Fig. 1, solid lines] the β s for the glide stimuli would have a positive mean and the β s for the no-glide stimuli would have a zero mean.

Next, the β s were categorized depending on whether the stimulus was a glide or no-glide stimulus (β_g and β_{ng} , respectively). The distribution of the resulting β s was close to a normal distribution. Then, the two distributions of β_g and β_{ng} were fitted using a Gaussian distribution, $\hat{\beta}_g \sim N(\mu_g, \sigma_g^2)$ and $\hat{\beta}_{ng} \sim N(\mu_{ng}, \sigma_{ng}^2)$, where μ and σ^2 represent the mean and the variance of the estimated distributions. Finally, the d'_{wgt} was estimated as

$$d'_{wgt} = \frac{\mu_g - \mu_{ng}}{\sqrt{\frac{\sigma_g^2 + \sigma_{ng}^2}{2}}}. \quad (6)$$

A second model was also applied to the results of the random-frequency conditions. It was assumed that subjects formed representations of the frequency glides independent of the central frequency. For this model, the stimuli were translated to have the same central frequency, and the model was applied to the resulting stimuli. This model will be referred to as the *frequency-translation model*. The resulting values of d'_{wgt} will be indicated as $\tilde{d}'_{wgt}$ for this model. This model allows the determination of whether the predicted sensitivity improved when frequency randomization

TABLE I. For each subject and each condition, values of Δf_c (in Hz) and d' are listed.

Conditions		NFB			FB		
		S1	S2	S3	S1	S2	S3
Fixed frequency	Δf_c	46	44	42	16	18	20
	d'_{obs}	1.52	1.38	1.51	1.76	1.72	1.62
Random frequency	Δf_c	46	64	60	24	18	26
	d'_{obs}	1.60	1.42	1.72	1.76	1.49	1.52

was removed. The procedures used to estimate $\tilde{d}'_{wgt}$ were the same as those described above, except the central frequencies were translated to 1000 Hz prior to the formation of the decision variables, the β_s .

c. Comparisons of the magnitudes of the initial and ending portions of TWP. As will be presented in the Results section, the TWPs derived from the current study were consistent with the expectations based on previous psychophysical results (Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999): The TWPs shared a negative-to-positive form, and unlike the ideal template, the magnitudes of the negative and positive weights tended to be different. That result led to two questions: (1) Whether subjects relied relatively more on the initial or the ending portion of the stimuli in order to make their detection decisions, and (2) was there a relation between the subjects' relative reliance on the initial/ending portion and their sensitivity? To address these questions, the ratio of the magnitudes of the relative weights from the ending *re.* initial portions of the TWPs was computed. If the ratio was approximately unity it would suggest equal weighting of the initial and ending portions of the stimuli. Additionally, consistent with the ideal template, should the TWP follow the negative-to-positive form, the initial negative and the latter positive portions would "cancel" the influence of the central frequency on decision variable β . This TWP would predict little or no effect of central frequency randomization. In contrast, a ratio either larger or smaller than unity would suggest subjects relied more on either the ending or initial portion of the stimuli, respectively. This result would predict a detrimental effect of central frequency randomization because the initial negative and the latter positive portions could not fully "cancel" the influence of the central frequency randomization.

These ratios of the magnitudes of the ending *re.* initial relative weights were estimated using different methods depending on the stimulus set. For the NFB conditions, the values were calculated by taking the ratio of the absolute values of the maximal (which occurred at toward the end of the TWP) *re.* the minimal (which occurred toward the initial part of the TWP) relative weights. For the FB conditions the relative weights from the initial and ending 200 ms fringe regions of the stimuli were separately averaged, and the ratio of the magnitudes of the averaged ending *re.* initial relative weights was computed.

To provide a statistical evaluation of the magnitude ratios, additional simulations were run. In each condition, the assumed TWP had the same form as the carrier frequency of the glide stimuli (see Fig. 1, solid lines), but the relative weight of the temporal middle of the TWP was zero. The stimuli from subject 2 (S2) were used for the simulation. First the 1250 stimuli were drawn, with replacement, from among the 1250 stimuli tested. Then, the inner product of the assumed TWP and $f_w(t)$ for each stimulus was obtained [Eq.

(5)] and a zero-mean normal deviate was added to form the decision variable. The standard deviation of the zero-mean deviate was adjusted so as to approximate d'_{obs} for S2 in each condition. Simulated responses of "Yes" or "No" were assigned depending on whether the resulting decision variable was above or below the criterion, respectively. The criterion was the average decision variable across all trials. Once simulated responses were obtained, simulated TWPs were generated using the same methods as for the human subjects. This process was repeated 40 times, yielding 40 simulated TWPs for each of the four conditions. The ratio of the magnitudes of the ending and initial relative weights was calculated for each of the 40 replicate simulations. In each condition, the higher and lower boundaries of the 96% confidence interval of the ratios for the simulated TWPs were determined by adding or subtracting 2.054 times the standard deviations from the mean ratio across the 40 replicate indices.

III. RESULT AND DISCUSSION

A. Psychoacoustic results

Table I shows the frequency extents (Δf_c) used for each subject in each condition. The measured values of d'_{obs} are also listed. Given that the values were obtained from a Yes/No procedure, relative criteria (response bias) were also estimated (Macmillan and Creelman, 1991). The relative criteria were near zero, ranging from -0.1 to 0.2 across subjects and conditions, with S1 producing the largest relative criteria. For approximately equal values of d'_{obs} , Δf_c s were larger in the NFB than the FB conditions, and typically larger for random- than fixed-frequency conditions. These results were consistent with past work demonstrating the advantage of the fringe portions (e.g., Lyzenga *et al.*, 2004) and the disadvantages of frequency randomization (e.g., Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999), although in those studies the frequency extents at threshold were substantially smaller because the stimuli were tones rather than narrow bands of noises.

B. Estimated temporal weight patterns

The bold curves in Fig. 3 show the TWPs (relative weights as a function of time) for S1, whose results are as representative as any other subject. These results are shown separately to provide an example of the derived TWPs and the 96% confidence intervals associated with random responses. The left panels show the results for the NFB conditions and the right panels are for the FB conditions. Upper panels are for the fixed-frequency conditions and lower panels are for the random-frequency conditions. The shaded areas show the 96% of confidence intervals for TWPs derived

FIG. 3. Temporal weight patterns, i.e., relative weights as a function of time, derived from S1's responses for (a) NFB-fixed, (b) FB-fixed, (c) NFB-random and (d) FB-random conditions. The shaded areas specify the 96% confidence interval about the pattern derived from random responses. The vertical dashed lines in the right panels indicate the temporal regions associated with the glide-bearing portions of the stimuli.

from unrelated to the presence/absence of the glide. The vertical dashed lines in the right panels indicate the temporal regions associated with the glide-bearing portions of the stimuli. The TWPs shifted from negative to positive values. This suggested that S1 contrasted the early and later frequencies rather than relying exclusively on just one or the other in making detection decisions. Moreover, there were substantial regions where the TWPs exceeded the confidence intervals, indicating that the obtained TWPs contained reliable information regarding S1's decision process. Note that the TWPs were constrained to start and end at zero.

The TWPs were scaled by the inverse of Δf_c in an effort to provide a direct comparison of the TWPs across conditions and subjects. For an ideal observer (assuming the linear model specified by Eq. (5) was correct and assuming the hypothetical observer maximized percent correct), such scaling would provide TWPs independent of Δf_c . In the current experiment the results were unlikely to be ideal, and so a slight error was assumed for this scaling procedure. Even so, this approach allowed a means of comparing the magnitudes of the relative weights across conditions and subjects.

Figure 4 shows subjects' individual TWPs scaled by the inverse of Δf_c for all conditions. The panels are arranged as in Fig. 3. The vertical dashed lines in the right panels indicate the temporal regions associated with the frequency glide. The results for different subjects are plotted using different line types. The results from Fig. 3, after scaling, are reproduced using the solid gray lines.

1. Comparisons of TWPs across conditions

Two common features were apparent in the obtained TWPs: (1) There was a clear negative-to-positive pattern in all TWPs, and (2) Most of the negative weights were in the first half of the TWPs and most of the positive weights were

FIG. 4. Temporal weight patterns scaled by the inverse of Δf_c in (a) NFB-fixed, (b) FB-fixed, (c) NFB-random and (d) FB-random conditions. Data from different subjects (S1, S2, S3) are plotted using different line types. Weights of zero are specified by the horizontal dashed lines. The vertical dashed lines in the right panels indicate the temporal regions associated with the glide-bearing portions of the stimuli.

in the second half. Subjects' processing strategies were consistent with a linear model in which the first and second halves of the stimuli were differenced.

A third feature of the TWPs, not apparent in Fig. 4, was that for all but the FB-random conditions, the 96% confidence intervals were exceeded in both the beginning and ending portions of the TWPs. In the FB-random condition, the TWPs generally exceeded the confidence intervals only in the latter portion. This result might indicate a failure of the linear model in providing a reliable description of subjects' strategies in the FB-random condition, a possibility that will be considered more fully below (Sec. III B 2). Overall, however, the approach appears to have been successful in providing reliable and meaningful patterns of relative weights across time.

Across conditions, differences among TWPs were apparent. First consider the differences between the TWPs derived in the fixed- and random-frequency conditions. The magnitudes of the scaled weights were larger for the fixed than the random conditions. Moreover, the magnitudes of the weights were larger in the latter portion than the beginning portion of the TWPs. These greater magnitudes of the weights in the latter portion were clearer in the TWPs derived from the fixed- than random-frequency conditions. This suggested that, for the fixed-frequency conditions subjects relied more heavily on the ending frequency than the starting frequencies in decisions regarding the presence of a frequency glide. The relative magnitudes of the beginning and ending portions of the TWPs are considered more fully below (Sec. III B 3).

Next, consider the TWPs derived in the NFB and FB conditions. Restricting the comparison to just the glide portion of the TWPs, the extent of the scaled weights were roughly similar in the FB and NFB conditions, but the proportion of the TWP that fell outside the 96% confidence interval was much larger for the NFB condition (see Fig. 3 for an example). This suggests that subjects may have at-

TABLE II. Values of d'_{obs} , d'_{wgt} , and $\tilde{d}'_{wgt}$ are listed for each subject in fixed- and random-frequency conditions.

Conditions/model		NFB			FB		
		S1	S2	S3	S1	S2	S3
Fixed frequency	d'_{obs}	1.52	1.38	1.51	1.76	1.72	1.62
	d'_{wgt}	1.73	1.72	2.02	2.53	2.43	2.76
Random frequency	d'_{obs}	1.60	1.42	1.72	1.76	1.49	1.52
	d'_{wgt}	1.48	2.70	1.40	0.58	0.47	0.76
Random frequency - translated	$\tilde{d}'_{wgt}$	2.44	2.94	2.95	3.22	2.44	3.37

tended less to the glide portion in the FB conditions compared to the NFB conditions because there was additional information available in the fringe bands (Lyzena *et al.*, 2004).

2. Comparison of d' from observers and TWPs

The first four rows of Table II show the d'_{obs} and the d'_{wgt} for the four experimental conditions. The final row indicates the values of $\tilde{d}'_{wgt}$ for the frequency-translation model applied in the random conditions. The values of d'_{obs} are reproduced from Table I. For the NFB- and FB-fixed conditions, the values of d'_{wgt} were larger than the corresponding values of d'_{obs} , and this difference was magnified in the FB-fixed condition compared to the NFB-fixed condition. Because the TWP predicted better-than-obtained performance, the information in the TWPs was sufficient to describe subjects' performance in the fixed conditions.

Next, consider the random conditions. Except for S2 in the NFB-random condition, values of d'_{wgt} were smaller than the corresponding d'_{obs} . This indicated a failure of the linear model without frequency translation [i.e., Eq. (5)] because the model predicted lower-than-obtained values of d' . For the FB-random condition, the values of d'_{wgt} were even poorer. When Eq. (5) was altered to include frequency translation, the resulting predictions, $\tilde{d}'_{wgt}$, were much larger than the corresponding d'_{obs} . This result has two implications. First, the failure of the linear model in the random-frequency conditions does not reflect poorly estimated TWPs: $\tilde{d}'_{wgt}$ is much larger than d'_{obs} meaning that there was sufficient information in the TWPs to account for performance when frequency translation was applied. Second, there might be a psychophysical process corresponding to frequency translation but not captured by the linear detection model [in Eq. (5)], which allows the d'_{obs} to be higher than d'_{wgt} . By one potential process, the subject would form an internal representation of the stimulus distinct from the central frequency. This is simply the frequency translation model and predicts no effect of central frequency randomization, which is inconsistent with the data shown in Table I. If however, subjects failed to *completely* extract the stimuli away from the central frequencies, there would be on the one hand an effect of central frequency randomization, and on the other hand, a TWP that was successful in accounting for subjects' performance only after frequency translation.

3. Comparisons of the magnitudes of the initial and ending portions of TWPs

For TWPs in the fixed conditions, visual inspection suggested a tendency for the magnitudes of the weights to be smaller at the beginning than the end of the TWPs. This is consistent with the expectations of TWPs based on the previous psychophysical data (Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999): The unequal magnitudes of the initial negative and ending positive weights. Table III shows the ratio of the magnitudes of the ending *re.* initial portions of the TWPs. For the NFB conditions the ratio was based on the absolute values of the maximal and minimal weights and for the FB conditions the ratio was based on the absolute averaged values of the TWP associated with the ending *re.* initial fringe bands. The ratios in the first and second rows of Table III are from the fixed- and random-frequency conditions, respectively, and the columns are for the individual subjects in the NFB and FB conditions, respectively.

For the fixed-frequency conditions (first row) computer simulations indicated 96% confidence intervals surrounding unity of [0.58, 1.44] for the NFB-fixed condition and [0.75, 1.76] for the FB-fixed condition. In five of the six comparisons, the obtained ratios exceeded the 96% confidence intervals indicating that magnitudes of relative weights were larger at the ending than the initial portions of the TWPs. Subjects appeared to rely more heavily on the ending than the beginning portions of the stimuli in making their detection decisions. For the subjects with higher (lower) values of Δf_c the ratios were greater than (nearer) unity, although the comparisons are based on relatively few subjects and small differences.

For the random-frequency conditions (second row) computer simulations indicated the following 96% confidence intervals about a ratio of unity: [0.13, 1.74] for the NFB-random condition and [0.62, 1.55] for the FB-random condition. For the NFB condition, none of the ratios exceeded the

TABLE III. For each subject and each condition, the ratio of the magnitudes of the ending *re.* initial portions of the TWP are listed. Values in bold exceed the 96% confidence intervals about a ratio of unity.

Conditions	NFB			FB		
	S1	S2	S3	S1	S2	S3
Fixed frequency	3.07	2.74	2.26	0.82	2.90	3.45
Random frequency	1.40	1.43	1.59	2.29	1.09	1.94

96% confidence intervals and for the FB condition, the ratios for S1 and S3 exceeded the 96% confidence intervals. As in the fixed-frequency conditions, for the random-FB condition the two subjects with the larger ratios also had the poorer performance (greater values of Δf_c), but this reflects only three comparisons.

Subjects' reliance on the ending portion of the stimuli rather than the initial portion was stronger for the fixed- than random-frequency conditions. Potentially, in an effort to counteract the effect of central frequency randomization, subjects attempted to more equally rely on the initial and ending portions of the stimuli when the central frequency was random. However, to the degree subjects altered their strategies under the pressures of frequency randomization, they were not wholly successful. Such a strategy might, for example, introduce instability in their internal templates, which would potentially cause a failure of the linear model [outlined in Eq. (5)] to accurately estimate the TWP in the random-frequency conditions.

C. Summary of computational results

The obtained TWPs were significantly different from the patterns associated with random responses for all three subjects in all four conditions, indicating that the patterns contained information regarding subjects' decision strategies. In the fixed-frequency conditions the TWPs indicated that subjects relied on both the starting and ending portions of the stimuli, although more so on the ending portion. In the random-frequency conditions, the stronger dependence on the ending portion was less apparent. Moreover, for the random-frequency conditions the linear detection model indicated by Eq. (5) failed to account for subjects' sensitivity. When frequency translation was applied to the stimuli, the model could at least account for subjects' sensitivity. Together, these results suggested that subjects altered their decision processes depending on whether the central frequency was randomized so as to more effectively difference the initial and ending portions of the stimuli.

IV. SUMMARY AND CONCLUSIONS

A linear classification model (Ahumada, 2002) was successfully used to derive relative weights as a function of time for the detection of an ascending frequency glide applied to 50-Hz-wide narrowband noise. Four conditions were tested, depending on whether the glide had onset and offset fringes or not (FB versus NFB) and depending on whether the central frequencies of the stimuli were fixed or random. Sensitivity to changes in frequency were superior in the FB relative to the NFB conditions, and for both the NFB and FB conditions sensitivity was superior for the fixed- than the random-frequency conditions. Although the current psychophysical data were obtained using narrow bands of noise, the pattern of results was consistent with previous results using tonal stimuli (Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999; Lyzenga et al., 2004).

The pattern of relative weights as a function of time (TWPs), were derived for all four conditions. The estimated TWPs went from negative to positive, consistent with the

ideal template. This negative-to-positive pattern also suggested that subjects did not rely on just the beginning or ending portion of the stimuli in detecting the glide. In most cases, the relative weights associated with the TWPs had larger magnitudes in the latter portion of the stimuli, especially for the fixed-frequency conditions. An analysis of d' based on the obtained TWPs indicated that the linear model could at least account for subjects' sensitivity in the fixed-frequency conditions. For the random-frequency conditions, although the estimated TWPs have the forms consistent with the expectations from the previous psychophysical results (Madden and Fire, 1996; Moore and Sek, 1998; Sek and Moore, 1999), the linear detection model [specified by Eq. (5)] nevertheless failed. An additional model representing the stimuli as changes in frequency (i.e., not actual frequencies) could account for subjects' sensitivity.

ACKNOWLEDGMENTS

This work was supported by Grant No. R01 DC002012 from the National Institutes of Health. We acknowledge the very helpful comments provided by two reviewers on earlier versions of this manuscript. Dr. Daniel Shub and Rong Huang also provided very helpful comments on earlier versions of the manuscript.

¹Note that the frequency perturbations introduce variations in the decision variable, which results in more overlap between the two distributions of the decision variable for the glide and no-glide stimuli. This additional overlap due to the perturbations would make the glide more difficult to detect, and might cause the no-glide stimulus sound as though there is a glide.

²Ahumada (2002) derived a scaling procedure by which hit rates and false alarm rates are used to minimize the error in the weight patterns. Here, a somewhat less optimal scaling scheme, the proportion of "Yes" and "No" trials, is used. This approach ensures that the resulting weight patterns are in similar units (akin to deviations in Hz) across conditions.

³The distribution of the weighted deviations ($\Delta f_w(t)$) is roughly normal, with a mean near zero and a standard deviation of approximately 10. The kurtosis and skewness of the distribution are approximately the same as for a normal distribution.

- Ahumada, Jr. A. J. (2002). "Classification image weights and internal noise level estimation," *J. Vision* **2**, 121–131.
- Anantharaman, J. N., Krishnamurthy, A. K., and Feth, L. L. (1993). "Intensity-weighted average of instantaneous frequency as a model for frequency discrimination," *J. Acoust. Soc. Am.* **94**, 723–729.
- Berg, B. G. (1990). "Observer efficiency and weights in a multiple observation task," *J. Acoust. Soc. Am.* **88**, 149–158.
- Berg, B. G., and Green, D. M. (1990). "Spectral weights in profile listening," *J. Acoust. Soc. Am.* **88**, 758–766.
- Bernstein, L. R., and Trahiotis, C. (1992). "Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies," *J. Acoust. Soc. Am.* **91**, 306–316.
- Buus, S., Zhang, L., and Florentine, M. (1996). "Stimulus-driven, time-varying weights for comodulation masking release," *J. Acoust. Soc. Am.* **99**, 2288–2297.
- Dai, H. (1993). "On the pitch of two-tone complexes," *J. Acoust. Soc. Am.* **94**, 730–734.
- Dai, H., and Berg, B. G. (1992). "Spectral and temporal weights in spectral-shape discrimination," *J. Acoust. Soc. Am.* **92**, 1346–1355.
- Demany, L., and Semal, C. (1989). "Detection thresholds for sinusoidal frequency modulation," *J. Acoust. Soc. Am.* **85**, 1295–1301.
- Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap* (Chapman and Hall/CRC, New York).
- Green, D. M. (1988). *Profile Analysis: Auditory Intensity Discrimination* (Oxford U. P., New York).
- Hartmann, W. M., and Klein, M. A. (1980). "Theory of frequency modula-

- tion detection for low modulation frequencies," J. Acoust. Soc. Am. **67**, 935–946.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 167–177.
- Lyzenga, J., Carlyon, R. P., and Moore, B. C. (2004). "The effects of real and illusory glides on pure-tone frequency discrimination," J. Acoust. Soc. Am. **116**, 491–501.
- Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide* (Cambridge U. P., Cambridge).
- Madden, J. P., and Fire, K. M. (1996). "Detection and discrimination of gliding tones as a function of frequency transition and center frequency," J. Acoust. Soc. Am. **100**, 3754–3760.
- Moore, B. C., and Sek, A. (1998). "Discrimination of frequency glides with superimposed random glides in level," J. Acoust. Soc. Am. **104**, 411–421.
- Richards, V. M., and Tang, Z. (2006). "Estimates of effective frequency selectivity based on the detection of a tone added to complex maskers," J. Acoust. Soc. Am. **119**, 1574–1584.
- Sek, A., and Moore, B. C. (1999). "Discrimination of frequency steps linked by glides of various durations," J. Acoust. Soc. Am. **106**, 351–359.
- Turner, C. W., Kwon, B. J., Tanaka, C., Knapp, J., Hubbart, J. L., and Doherty, K. A. (1998). "Frequency-weighting functions for broadband speech as estimated by a correlational method," J. Acoust. Soc. Am. **104**, 1580–1585.