Detection and discrimination of frequency glides as a function of direction, duration, frequency span, and center frequency

John P. Madden and Kevin M. Fire
Department of Communication Sciences and Disorders, University of North Dakota, P.O. Box 8040, University Station, Grand Forks, North Dakota 58202

(Received 30 January 1997; revised 21 July 1997; accepted 21 July 1997)

The study investigated the ability to detect and discriminate frequency glides under a variety of experimental conditions. The subjects distinguished between a comparison signal that either was level in frequency or was swept across a fixed frequency span, and a target signal that changed more in frequency than the comparison signal. Tone durations were 50 and 400 ms. Nominal center frequencies were 0.5, 2, and 6 kHz; actual center frequencies were varied randomly, or roved, over a range equal to 0.1 times the nominal center frequency. Up- and down-glides were used. The transition span of the comparison signal was either 0, 0.5, 1, or 2 times the equivalent rectangular bandwidth of the auditory filter at the nominal center frequency. Discrimination thresholds were obtained for all combinations of center frequency, direction, and span. Overall, thresholds expressed as ΔHz/ERB varied little as a function of center frequency. Glide duration had no effect on discrimination. The 50-ms down-glides were more difficult to detect than the 50-ms up-glides; otherwise, the effect of direction was not significant. With the exception of the 50-ms down-glides, detection/discrimination thresholds were similar for the 0-, 0.5-, and 1-ERB transition spans, but increased significantly for the 2-ERB span. The absence of significant variation across frequency supports a place mechanism for the detection of frequency change in gliding tones, based on the detection of changes in the excitation pattern. An excitation pattern model cannot account for the asymmetry noted for glide detection, however. © 1997 Acoustical Society of America.

PACS numbers: 43.66.Ba, 43.66.Fe [WJ]

INTRODUCTION

Recently, Madden and Fire measured detection and discrimination thresholds for frequency glides (Madden and Fire, 1996). When expressed as a proportion of the equivalent rectangular bandwidth (ERB) of the auditory filter, thresholds varied little across center frequencies ranging from 0.5 to 6 kHz. This lack of significant variation in threshold across center frequency was interpreted as evidence for a place mechanism for coding frequency change that monitors the excitation pattern produced by the stimulus. Signal duration was fixed at 50 ms, however, and thus the study provided evidence supporting a place mechanism for signals of brief duration only. A “sluggish” temporal mechanism based on phase-locking might be used to track glides of greater duration and more gradual rate of change, as suggested by Sek and Moore (1995). Also, all signals were up-glides, and therefore the effect of direction was not investigated, nor has it been systematically investigated for glide discrimination in previous studies. Thus, the present study was undertaken to extend the initial study to determine the effect of duration and direction on glide detection and discrimination thresholds.

For the most part, previous studies have found that glide detection and discrimination thresholds decrease as signal duration increases, with most of the decrease having occurred when signal duration reaches 100 ms (Sergeant and Harris, 1962; Pollack, 1968; Tsumura et al., 1973; Dooley and Moore, 1988a). However, in a glide detection task, if the center frequencies of the tones being distinguished are the same, the subject potentially can compare the end points of the tones and select the tone with the end point that is higher or lower in pitch. If this is the case, the subject is performing a between-signal pitch comparison which is similar to a DLF determination, rather than estimating the pitch change within each glide. It is possible that the subjects in these studies were making such end-point comparisons. The utility of this cue could increase as signal duration increases because the subjects could obtain a longer-duration sample at the end point of the signal. In the present study, center frequencies were roved over a range of frequencies that was a fixed proportion of the nominal center frequency, to force the subjects to use only within-signal information.

Several studies have noted direction asymmetries for glide detection, but the results are contradictory. Dooley and Moore (1988a) found that at most durations and center frequencies tested, average glide detection thresholds for up-glides were larger than those for down-glides, although the differences were small except at 2 kHz. Tsumura et al. (1973) found an asymmetry in the same direction and of a similar magnitude, with the up-glide thresholds being about 20% greater than the down-glide thresholds. Arlinger et al. (1977) and Nabelek and Hirsh (1969) found no asymmetry. Finally, both Gardner and Wilson (1979) and Schouten (1985) found that detection thresholds for down-glides were greater than those for up-glides. The present study in-
vestigated the effect of glide direction in greater depth than these previous studies by measuring thresholds for up- and down-glides across a range of center frequencies in several glide discrimination conditions, as well as a glide detection condition.

The overall purpose of the study was to investigate further the mechanisms that underlie glide perception. The primary objective was to determine whether there is evidence for temporal coding of long-duration glides, by comparing detection/discrimination thresholds for 50- and 400-ms glides with fixed frequency spans. Based on the results from the authors' previous experiments (Madden and Fire, 1996), it was expected that thresholds for the 50-ms signals would be a constant proportion of ERB across center frequency in both the up and down conditions at all transition spans, thus supporting a place mechanism. In contrast, it was hypothesized that a temporal mechanism based on phase-locking to the fine structure of the signal might be involved in the detection and discrimination of the more slowly changing 400-ms signals. Evidence for this would be a shift from the “constant proportion of ERB” threshold pattern for the 50-ms signals to one that is similar to the pattern of the DLF for pulsed tones in the case of the 400-ms signals. The latter pattern generally is thought to be determined by the operation of a temporal mechanism in the low frequencies and a place mechanism in the high frequencies (Moore, 1973). DLF thresholds typically are relatively constant from 0.5 to 2 kHz, but increase by a factor of 3 or 4 at 6 kHz (see, e.g., Moore and Glasberg, 1989). It was thought that this pattern of threshold variation across frequency would be most evident for the glide detection conditions, in which the comparison signal does not change and the target signal changes very little, giving the “sluggish” temporal mechanism the greatest opportunity to operate. This pattern should gradually become less evident in the discrimination conditions as the frequency span of the standard signal increases, and with it the rate of change of the signals. A secondary objective was to determine whether direction affects glide detection and discrimination. Neither a place nor a temporal mechanism of frequency coding predicts that there should be an effect of direction.

I. METHOD

A. Subjects

Four subjects with normal auditory sensitivity participated in the experiment. All had hearing thresholds of 15 dB HL or less at the audiometric test frequencies and were paid for their participation.

B. Procedure

In a two-alternative, forced-choice task, the subjects were asked to identify a target signal which either increased or decreased in pitch (for the glide detection condition) or increased or decreased more in pitch than the comparison signal (for the glide discrimination conditions). Center frequencies were 0.5, 2, and 6 kHz, and the start and end points of the stimuli were linearly equidistant from their center frequencies. Stimulus durations were 50 and 400 ms, with additional rise–fall times of 5 ms. A cosine-squared windowing function was used. Thresholds were determined for each of the four combinations of duration and direction using comparison signals that changed in frequency by the following amounts: 0, 0.5, 1, and 2 ERB. Table I gives the transition spans used at the various center frequencies for the three gliding comparison signals. All signals were presented monotonically at 70 dB SPL, with an interstimulus interval of 500 ms.

The center frequencies of the signals were roved over a range equal to 0.1 times the nominal center frequency. Roved center frequencies were randomly drawn from a uniform distribution. The center frequencies of both the comparison signal and the target were roved independently both within and between trials. A more detailed description of this procedure can be found in Madden and Fire (1996).

Thresholds were measured using an adaptive procedure that estimated the 70.7% correct response point on the psychometric function (Levitt, 1971). After two consecutive correct responses, the change in frequency of the target signal was decreased; after one incorrect response it was increased. Step size was decreased after the first six reversals and varied depending on the center frequency and condition. At a center frequency of 0.5 kHz, for the 0-Hz condition, the initial step size was 10 Hz, at the higher center frequencies and for the 1-and 2-ERB conditions, where thresholds were greater, initial step sizes ranged up to 40 Hz. Final step sizes ranged in a similar fashion between 2 and 10 Hz. The frequency increase of the target signal was not permitted to fall below that of the comparison signal.

A run consisted of 80 trials, with a break after the first 40, to maintain alertness. The first six reversals were discarded, and threshold was computed as the arithmetic mean of the subsequent reversals (always at least eight). The subjects were practiced for at least 16 hours before data collection began. For each combination of center frequency, duration, and transition span, they then were tested until performance did not consistently improve over six runs. Of these six, the runs that varied by more than one standard deviation from the mean were discarded, and the threshold calculated from the remaining data (always at least four runs). Data collection was completed for each center frequency-duration-direction combination before the next combination of conditions was begun, to facilitate optimal learning of detection/discrimination cues. The order of the duration, direction, and center frequency conditions was counterbalanced across subjects.

Stimulus presentation and response acquisition were...
Controlled by a PC. Stimulus and response intervals were signaled on a computer monitor. The subjects received visual feedback after every trial, indicating the stimulus interval containing the target signal.

C. Stimulus generation

Digital representations of the stimuli were generated by an array processor (TDT-QAP2) running on a PC. A 16-bit D-to-A converter (TDT-DD1) operating at a 50-kHz sampling rate converted the digital representations to analog waveforms which then were low-pass filtered at 8 kHz (TDT-FLT3), attenuated (TDT-PA3), and sent through an output amplifier (TDT-HBUF3) to an earphone (Sennheiser HD-414SL). The frequency response of this earphone has been reported to be relatively smooth across the range of the test frequencies (see Moore and Sek, 1995). Signal spectra were verified using Hypersignal/Workstation spectrum analysis software.

II. RESULTS

Inspection of the data suggested that whereas absolute performance differed among the subjects, all demonstrated highly similar patterns of performance across the various experimental conditions. This supposition was tested statistically with a series of two-way ANOVAs (analyses of variance) which showed that the subject factor did not interact with the factors of duration, direction, center frequency, or transition span. Therefore, the data were collapsed across subject, and the average thresholds for the various experimental conditions are presented in Fig. 1. Thresholds are expressed as a proportion of the ERB at the nominal center frequency of the signal.

An inspection of the four graphs in Fig. 1 reveals that overall there is a tendency for thresholds to be lower at 2 than at 0.5 and 6 kHz, but that the differences in threshold magnitude between the center frequencies are not great. A one-way ANOVA indicated that there was not a significant main effect of center frequency $F(1,176) = 2.012$, $p = 0.137$. In addition, two-way ANOVAs with frequency and direction, frequency and duration, and frequency and transition span as factors revealed no significant interactions. Therefore, to get a clearer picture of the effects of direction, duration, and transition span, the data were collapsed across center frequency. The averaged thresholds are presented in Fig. 2.

It is evident from Fig. 2 that the 50- and 400-ms thresholds are very similar when compared within the same transition span/direction condition, except for the 0-ERB down-glides. In this case, the 50-ms threshold is 1.9 times as great as the 400-ms threshold. Down-glide thresholds are greater than up-glide thresholds in all cases, but the differences are small, except for the 50-ms, 0-ERB thresholds. Here, the down-glide threshold is 1.8 times greater than the up-glide threshold. With the obvious exception of the 0-ERB, 50-ms, down-glide condition, thresholds are similar for the 0- to 0.5-, and 1-ERB transition spans, but increase considerably for the 2-ERB transition.

To test these conclusions statistically, a three-way ANOVA was conducted with duration, direction, and transition span as factors. There were significant main effects of duration $F(1,176) = 6.19$, $p = 0.014$, direction $F(1,176) = 23.23$, $p < 0.001$, and transition $F(1,176) = 38.338$, $p < 0.001$. There also were significant interactions between duration and direction $F(3,176) = 4.02$, $p = 0.046$, direction and transition $F(3,176) = 2.63$, $p = 0.05$, and duration and transition $F(3,176) = 6.37$, $p < 0.001$.

To determine the sources of these effects, pairwise comparisons were conducted of the threshold values for all 16 combinations of the various levels of duration, direction, and transition span. The 50-ms, 0-ERB, down-glide threshold was significantly greater than the 50-ms, 0-ERB, up-glide threshold and the 400-ms, 0-ERB, down-glide threshold. There were no other significant differences associated with the effect of duration or direction. For the 400-ms signals, within both directions, the 0-, 0.5-, and 1-ERB thresholds did not differ from one another but differed from the 2-ERB thresholds. The same was essentially true for the 50-ms sig-

![FIG. 1. Average discrimination thresholds for the various combinations of duration, direction, and frequency span, plotted as a function of center frequency. Thresholds are expressed as a proportion of the ERB at the signal center frequency. Error bars indicate one standard deviation from the mean.](image1)

![FIG. 2. Detection and discrimination thresholds for each signal duration, averaged across center frequency and plotted as a function of transition span.](image2)
nals, with the exception of the 0-Hz, down-glide threshold, which differed from the corresponding 0.5- and 1-ERB thresholds, but did not differ from the corresponding 2-ERB threshold. Thus the statistical analysis supported the conclusions reached by inspection of the data.

Summarizing the main trends in the data, the effect of center frequency was not significant. There was no effect of duration or direction, except in the glide-detection condition, where there was an interaction between duration and direction. Here, the down-glide threshold was significantly greater than the up-glide threshold at 50 ms but not at 400 ms. Thresholds were similar for the 0-, 0.5-, and 1-ERB transition spans and increased for the 2-ERB span, with the exception of the 50-ms down-glides.

III. DISCUSSION

A. The effect of transition span

The overall average of the 2-ERB thresholds is slightly less than twice the overall average of the other transition spans, when the 50-ms, down-glide, 0-ERB threshold is not included. This was the pattern described for 50-ms up-glides by Madden and Fire in their 1996 paper, in which it was suggested that glide discrimination thresholds exhibit a constant Weber fraction. The results of the current study extend this finding to include glides of long duration and down-glides, providing further support for the idea that glide discrimination thresholds follow Weber’s law.

B. The effect of duration

The absence of a general effect of duration is surprising in view of the results from previous studies. In fact, the 400-ms, up-glide direction thresholds are considerably higher than those obtained by others using signals of similar duration. For example, Dooley and Moore (1988a) obtained average up-glide thresholds of about 4 Hz at 0.5 kHz and about 13 Hz at 2 kHz using 500-ms signals, whereas in the present experiment the corresponding 400-ms thresholds were 13 and 34 Hz. The most probable cause of this discrepancy is the use of signals roved in center frequency. Although a previous experiment (Madden and Fire, 1996) found that the effect of roving signals was not significant for the detection of 50-ms frequency glides, an ensuing pilot study indicated that the effect of roving the signals increases greatly when signal duration is increased. For example, an average detection threshold of 13 Hz was obtained for a 2000-Hz, 200-ms up-glide in the nonroved condition versus a roved threshold of 38 Hz for the same signal. It was speculated that the subjects used between-signal, end-point cues in the nonroved condition for these longer-duration signals, as was discussed in the Introduction. It thus appears that although glide detection thresholds decrease with increasing signal duration when center frequencies are constant, they do not change with duration (over the range tested here) when center frequencies are roved.

C. The effect of direction

Because an effect of direction was noted only for the 50-ms, glide detection portion of the experiment, it was decided to recheck these thresholds by testing a second set of subjects consisting of two of the original subjects and two new subjects. The results are presented in Fig. 3 and are similar to those of the main experiment. A two-way ANOVA with frequency and direction as factors revealed significant main effects of direction \(F(1,18) = 14.43, p = 0.001\) and center frequency \(F(2,18) = 3.61, p = 0.048\) but no interaction \(F(2,18) = 1.66, p = 0.219\). Pairwise comparisons indicated that the up and down thresholds differed significantly at 0.5 kHz, but not at 2 or 6 kHz. The greatest asymmetry was at 0.5 kHz in the main experiment, also. It was pointed out in the Introduction that other workers have reported results showing the opposite asymmetry, but including the present study, there now are at least three studies in which the same asymmetry has been observed for brief duration glides (also Gardner and Wilson, 1979 and Schouten, 1985). It is not clear what sort of underlying mechanism might be responsible for the asymmetry, or why its occurrence should be affected by duration and frequency.

D. Modeling glide detection and discrimination

The data yield little support for temporal coding of the 400-ms signals. As was suggested in the Introduction, if a temporal mechanism plays a role in tracking glides, one would expect an effect of center frequency, with an increase in threshold magnitude at 6 kHz relative to thresholds at the lower center frequencies, particularly in the glide detection conditions. Moore and Sek (1995) and Sek and Moore (1995), for example, observed that FM detection thresholds expressed as a proportion of ERB increased by a factor of almost 2 at a carrier frequency of 6 kHz, relative to the lower carrier frequencies, when the modulation rate was 2 Hz. In fact, the 400-ms thresholds are more consistent with the predictions of a model in which frequency changes are detected.
by a mechanism that monitors changes in the level of excitation at the edges of the excitation pattern generated by a stimulus (Zwicker, 1970; Moore and Sek, 1994). For such a model, the accuracy with which the start and end points of a glide can be estimated will be determined by the steepness of the excitation pattern evoked by the signal, which in turn is determined by the bandwidth of the auditory filters centered at or just below the start and end frequencies. An excitation pattern model thus predicts that detection and discrimination thresholds should be a fixed proportion of auditory filter bandwidth across frequency. However, an excitation pattern model cannot account for the direction asymmetry that was observed for glide detection. The absence of a duration effect also is problematic. The 400-ms signals remain near their start and end-point frequencies for a longer time period than the 50-ms signals. This extra time should allow more accurate sampling by whatever mechanism is involved, and result in better discrimination at longer durations.

IV. CONCLUSIONS

(1) Glide discrimination was not affected by center frequency, duration, or direction.

(2) There was an effect of frequency transition span. Except for the 50-ms down-glides, thresholds remained relatively constant as the comparison signal span increased from 0 to 1 ERB, but almost doubled when the comparison signal span increased to 2 ERB.

(3) Glide detection thresholds were affected by an interaction between direction, duration, and transition span: 50-ms down-glides were more difficult to detect than 50-ms up-glides or 400-ms down-glides and up-glides. This effect was most evident at 0.5 kHz.

(4) The results are better accounted for by a place mechanism of frequency coding based on the detection of changes in the excitation pattern generated by the signal than by a temporal mechanism based on phase-locking to signal fine structure. However, the absence of a duration effect and the presence of a direction asymmetry are not explained by an excitation pattern model.

ACKNOWLEDGMENT

This research was supported by research Grant No. 1 R15 DC 02662-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

The frequency response of the HD-414SL earphone exhibits the greatest amount of variation in the higher frequencies. A difference in change in level between signals to be discriminated is, of course, a potential discrimination cue, if it is large enough. For example, the average change in level for a 2-ERB, 6-kHz standard signal (transition span = 1344 Hz) is about 4.4 dB. For a corresponding target signal with a transition span of 1744 Hz, which is about 100 Hz greater than the average target span at discrimination threshold for the 2-ERB condition in the previous study (Madden and Fire, 1996), the average change in level is about 5.7 dB. Thus the change in level of the target in this case would be about 1.3 dB greater than that of the standard. Dooley and Moore (1988b) found that subjects could detect a difference in change in level as small as 0.2 dB for 750-ms signals at 2 kHz. This would indicate that the detection of level changes is a potential cue for frequency glide detection and discrimination at the 6-kHz center frequency. However, the signals used by Dooley and Moore either did not change in frequency or changed by only a relatively small amount in comparison with the signals used in the glide discrimination tasks in the current study. Thus it is difficult to know the extent to which the findings of Dooley and Moore are relevant to the present study. Also, when the subjects in Madden and Fire (1996) were informally asked to comment on discrimination cues, they did not mention level changes and appeared unaware that any occurred.

The authors are indebted to Brian C. J. Moore for suggesting this interpretation of the data reported in Madden and Fire (1996).


