INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Auditory discrimination of frequency transitions by human listeners and a computational model

Hsu, Chien-Yeh, Ph.D.

The Ohio State University, 1993
AUDITORY DISCRIMINATION OF FREQUENCY TRANSITIONS BY HUMAN LISTENERS AND A COMPUTATIONAL MODEL

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University by Chien-Yeh Hsu, B.S., M.S.
The Ohio State University 1993

Dissertation Committee:
Lawrence L. Feth
Robert A. Fox
Ashok K. Krishnamurthy

Approved by
Adviser
Department of Speech and Hearing Science
To my parents and wife
ACKNOWLEDGMENTS

I would like to express my sincere thanks to my advisor, Professor Lawrence Feth, for his encouragement, guidance and useful advises throughout my graduate work. I would like to thank to Professor Ashok Krishnamurthy and Professor Robert Fox for their helpful suggestions and comments.

This work was supported, in part, by a grant from the AirForce Office of Scientific Research and a grant from the National Institute on Aging (#5R01 AG08353). Thanks also go to the Graduate Student Alumni Research Award program sponsored by the Graduate School of the Ohio State University for the awarded fund as part of the expense of this study.

I would like to thank to my wife, Chen-jui Chao, for her assistance and moral supports that made this research possible.
March 10, 1963 ............. Born - Taiwan, R.O.C.
June, 1985 ............... B.S.E.E., National Cheng Kung University, Tainan, Taiwan, R.O.C.
Feb., 1986 - May, 1987 ... Computer system and program designer, Chinese Army, R.O.C.
June, 1988 - March, 1989 .. Graduate Research Associate, Nisonger Center, The Ohio State University, Columbus, Ohio
May, 1989 - Sept., 1989 ... Graduate Research Associate, Department of Electrical Engineering, The Ohio State University, Columbus, Ohio
1989 ..................... M.S.E.E., The Ohio State University, Columbus, Ohio
Oct., 1989 - 1993 ......... Graduate Research Associate, Division of Speech and Hearing Science, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Speech and Hearing Science
Other studies: Digital Signal Processing, Electrical and Computer Engineering.
# TABLE OF CONTENTS

DEDICATION ............................................... ii
ACKNOWLEDGMENTS ......................................... iii
VITA ................................................... iv
LIST OF TABLES .......................................... viii
LIST OF FIGURES ......................................... ix

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Experiments of Discrimination in Frequency Transition</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Roving Starting Frequency Procedure</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Intensity Weighted Average of Instantaneous Frequency</td>
<td>10</td>
</tr>
<tr>
<td>1.4 The ASP’s Excitation Pattern</td>
<td>12</td>
</tr>
<tr>
<td>1.5 Computational Prediction Model for Discrimination of Frequency Transitions</td>
<td>13</td>
</tr>
<tr>
<td>1.6 Statement of Problem</td>
<td>15</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>18</td>
</tr>
<tr>
<td>2.1 Discrimination of Frequency Transitions</td>
<td>18</td>
</tr>
<tr>
<td>2.2 Randomization of Stimulus Parameters</td>
<td>26</td>
</tr>
<tr>
<td>2.3 The Patterson’s ASP Model</td>
<td>27</td>
</tr>
<tr>
<td>2.4 The IWAIF and EWAIF Model</td>
<td>30</td>
</tr>
<tr>
<td>2.5 The Independent Channels Model used for Discrimination predictions</td>
<td>33</td>
</tr>
<tr>
<td>III. METHODS</td>
<td>40</td>
</tr>
<tr>
<td>3.1 IWAIF Calculation</td>
<td>40</td>
</tr>
<tr>
<td>3.1.1 Computation of IWAIF using the Fourier transform</td>
<td>41</td>
</tr>
<tr>
<td>3.1.2 Computation of IWAIF using excitation pattern</td>
<td>42</td>
</tr>
<tr>
<td>3.2 Signals and Parameters</td>
<td>45</td>
</tr>
</tbody>
</table>
3.2.1 FM Sinusoids with Linear Frequency
Modulation ........................................ 45
3.2.2 Frequency Transition Signals
Generated by Linear Moving Filters .......... 47
3.2.3 Summary of signals ......................... 50
3.3 A Computational Prediction Model ........... 51
3.3.1 Slope Discrimination for Fixed
Starting Frequency ............................... 54
3.3.2 Slope Discrimination for Roved
Starting Frequency ............................... 56
3.3.3 Model Evaluation and Computer
Realization ........................................ 58
3.4 Experiment using human listener ........... 59
3.4.1 Subjects ...................................... 59
3.4.2 Signals and Equipment ...................... 61
3.4.3 Detection Threshold for Signals
Presented in Quiet ................................ 63
3.4.4 Frequency Difference Limen ............... 64
3.4.5 Transition Slope Discrimination ........... 65

IV. RESULTS ........................................ 69
4.1 Psychoacoustic Data for Listening
Experiment .......................................... 69
4.1.1 Detection Threshold ......................... 69
4.1.2 Frequency Difference Limen ............... 70
4.1.3 Frequency transition slope
discrimination ....................................... 74
4.1.3.1 Linear tone glides ....................... 74
4.1.3.2 Moving filter signals ...................... 83
4.1.3.3 Comparison between the
discrimination of pure tone glide and
moving filter signal ............................... 91
4.2 Performance of the Computational Model ..... 95
4.2.1 Mathematical Model Performance ......... 96
4.2.2 Model Performance for Tone Glides ...... 113
4.2.3 Model Performance for Moving
Filters .............................................. 122
4.2.4 Using the ASP Model as a
Preprocessor ....................................... 124

V. DISCUSSION ...................................... 129
5.1 Frequency Discrimination for
Steady-state tones and Transition Stimuli .. 129
5.2 Effect of Roving Starting Frequency ...... 132
5.3 Duration Effect for Transition
Discrimination ..................................... 136
5.4 The Asymmetric Discrimination for
Different Adaptive Directions ................. 137
5.5 Discrimination Difference between
Tone-glides and Moving Filter Signals ....... 140

vi
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Summary of signals.</td>
<td>51</td>
</tr>
<tr>
<td>2. Detection Threshold (in dB SPL) for Signals Presented in Quiet.</td>
<td>70</td>
</tr>
<tr>
<td>3. Listener Frequency DLs (in Hz).</td>
<td>71</td>
</tr>
<tr>
<td>4. TSDT (in Hz) for the tone glide slope discrimination.</td>
<td>75</td>
</tr>
<tr>
<td>5. Pairwise comparison matrix for different levels (tone glides)</td>
<td>81</td>
</tr>
<tr>
<td>6. TSDT (in Hz) for the moving filter slope discrimination.</td>
<td>84</td>
</tr>
<tr>
<td>7. Pairwise comparison matrix for different levels (moving filter signals)</td>
<td>90</td>
</tr>
<tr>
<td>8. Listener’s EL values for the three subjects. (100 ms tone glide, ideal IWAIF)</td>
<td>106</td>
</tr>
<tr>
<td>9. Listener’s EL values for the three subjects. (100 ms tone glide)</td>
<td>121</td>
</tr>
<tr>
<td>10. Listener’s EL values for the three subjects. (100 ms moving filter signal)</td>
<td>123</td>
</tr>
<tr>
<td>11. Listener’s EL values for the model using EPN data. (100 ms tone glide)</td>
<td>126</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Schematic sound spectrograms showing the formant transitions that are appropriate for the voiced stop consonants /b/, /d/, and /g/ before different vowels.</td>
<td>2</td>
</tr>
<tr>
<td>2. Spectrographic patterns with different transition directions and slopes.</td>
<td>5</td>
</tr>
<tr>
<td>3. Block diagram of the computation of d' for two signals.</td>
<td>16</td>
</tr>
<tr>
<td>4. Block diagram of generating real-time moving filter signals.</td>
<td>48</td>
</tr>
<tr>
<td>5. Independent channels model using IWAIF.</td>
<td>52</td>
</tr>
<tr>
<td>6. Diagram of the Procedures of the prediction model and how the model’s prediction is compared to the listener's performance.</td>
<td>60</td>
</tr>
<tr>
<td>7. Block diagram of equipment for frequency slope discrimination task and frequency DL task.</td>
<td>62</td>
</tr>
<tr>
<td>8. Two adaptive directions in the up-down adaptive procedure.</td>
<td>67</td>
</tr>
<tr>
<td>9. Results of frequency DLs measured at 79.4% correct performance (d' = 1.16).</td>
<td>72</td>
</tr>
<tr>
<td>10. Transition slope discrimination threshold (TSDT) for 100 ms tone glides.</td>
<td>76</td>
</tr>
<tr>
<td>11. Transition slope discrimination threshold (TSDT) for 50 ms tone glides.</td>
<td>77</td>
</tr>
<tr>
<td>12. Transition slope discrimination threshold (TSDT) for 25 ms tone glides.</td>
<td>78</td>
</tr>
<tr>
<td>13. Transition slope discrimination threshold for tone glides for three subjects.</td>
<td>80</td>
</tr>
</tbody>
</table>

15. Transition slope discrimination threshold (TSDT) for 100 ms moving filter signals.

16. Transition slope discrimination threshold (TSDT) for 50 ms moving filter signals.

17. Transition slope discrimination threshold (TSDT) for 25 ms moving filter signals.

18. Transition slope discrimination threshold for moving filter signals. Data are plotted separately for three subjects.

19. Transition slope discrimination threshold for moving filter signals. Data are plotted separately for three signal durations.

20. TSDT score averaged across signal durations for tone glides and moving filter signals. Data are plotted separately for three subjects.

21. TSDT score averaged across subjects for tone glides and moving filter signals. Data are plotted separately for three durations.

22. The d' prediction for subject AM with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

23. The d' prediction for subject LS with a frequency DL equals 8.09 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

24. The d' prediction for subject FL with a frequency DL equals 7.03 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

25. The comparison of the discrimination (TSDT) between the prediction model and subject AM. Ideal IWAIF.

26. The comparison of the discrimination (TSDT) between the prediction model and subject LS. Ideal IWAIF.

27. The comparison of the discrimination (TSDT) between the prediction model and subject FL. Ideal IWAIF.

28. The d' prediction for subject AM with a frequency DL equals 5.81 Hz. Signals are 50 ms tone glides. Ideal IWAIF.
29. The d’ prediction for subject AM with a frequency DL equals 7.64 Hz. Signals are 25 ms tone glides. Ideal IWAIF.

30. The comparison of the discrimination (TSDT) between the prediction model and subject AM for 50 ms signals. Ideal IWAIF.

31. The comparison of the discrimination (TSDT) between the prediction model and subject AM for 25 ms signals. Ideal IWAIF.

32. The d’ prediction for subject AM with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides.

33. The d’ prediction for subject LS with a frequency DL equals 8.09 Hz. Signals are 100 ms tone glides.

34. The d’ prediction for subject FL with a frequency DL equals 7.03 Hz. Signals are 100 ms tone glides.

35. The comparison of the discrimination (TSDT) between the prediction model and subject AM.

36. The comparison of the discrimination (TSDT) between the prediction model and subject LS.

37. The comparison of the discrimination (TSDT) between the prediction model and subject FL.

38. The d’ prediction for subject AM with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides. EPN data are used in the model.

39. The comparison of the discrimination (TSDT) between the prediction model and subject AM. EPN data is used in the model.

40. Neural excitation patterns corresponding to the signal offsets for different adaptive directions and the reference.

41. The d’ prediction for subject AM with TSDT equals 39.4 Hz at fixed condition. Signals are 100 ms tone glides. Ideal IWAIF.

42. The comparison of the discrimination (TSDT) between the prediction model and subject AM. Ideal IWAIF.
43. The d' prediction for subject AM with TSDT equals 39.4 Hz at fixed condition. Signals are 100 ms tone glides. .......................... 148

44. The comparison of the discrimination (TSDT) between the prediction model and subject AM. ............ 149

45. The Neural Excitation Pattern for a 100 ms tone glide starting at 1000 Hz and ending at 1600 Hz. 154
Speech information is conveyed by acoustic waveforms that change in frequency and intensity over time. Speech sounds are produced by the vocal mechanism. The resonances in sound transmission through the vocal tract are called formants. The formant frequency locations are affected by the length of the pharyngeal-oral tract, the location of constrictions in the tract, and the degree of narrowness of the constriction. The formants of a speech sound are numbered from low to high frequencies and are called the first formant (F1), second formant (F2), third formant (F3), and so on. In speech, formant transitions are resonance shifts leading into or out of quasi-stable vocal tract configurations. For example, in the conjunctions between vowels and consonants, or vowels and other vowels, we can see short duration frequency transitions.

Some early research showed that the second formant transitions carry important information about the place of production of most consonants (Delattre, 1958; Liberman, 1957; Liberman, Delattre, Copper, & Gerstman, 1954). Liberman et al. (1967) state that the second-formant transition is the most important for the perception of consonants because it carried
valuable linguistic information in speech signals. Information contained in formant transitions includes the direction and slope of frequency transitions. The formant transitions contained in speech sounds play a significant role in the speech perception. Therefore, the human auditory system must be able to detect, process, identify and categorize fast frequency transitions.

![Figure 1](image.png)

**Figure 1:** Schematized sound spectrograms showing the formant transitions that are appropriate for the voiced stop consonants /b/, /d/, and /g/ before different vowels (adapted from Delattre et al., 1955).

The acoustic consequences of coarticulation and other sources of contextually conditioned variability give rise to the problem of lack of invariance. The problem is that the
acoustic features of a given phoneme frequently vary as a function of the phonetic context in which the phoneme is produced. For example, the second formant transitions for syllable-initial stop consonants (e.g., in /ba/ or /da/), which provide cues for place of articulation (e.g., /b/ vs. /d/ vs. /g/), vary considerably depending on the following vowels (Delattre, Liberman, & Cooper, 1955). As sketched in Figure 1, the second formant transition for stop consonants in syllable-initial position do not uniquely specify the place of articulation across all vowels. In general, rising second formant (F2) indicates a labial stop, relatively flat F2 tends to be heard as alveolar, and a falling F2 produces velar perception. The Haskins researchers concluded that the acoustic cues for a given stop consonant were different depending on the vowel which followed the constant (Cooper et al., 1952).

It should also be noted that some researchers feel that the invariant acoustic features can be found in spectral transitions to and from stops. In addition, the stop burst spectrum can be a sufficient cue to the place of articulation. When the spectra of stop bursts are examined, labials tend to have diffusely falling or flat spectra, alveolars have diffusely rising patterns, and velars exhibit compact spectra (Blumstein and Stevens, 1979). Although the primary cues to place are spectral, VOT (voice onset time) and amplitude also play a role. For example, the unvoiced stops usually have
longer VOT, especially the velars. The burst intensity increases as the place of articulation becomes further posterior in the vocal tract. Finally, the motor theory suggested that the perception of a formant transition is different depending on whether it is perceived in the auditory model, where it sounds like a chirp, or in the phonetic model, where it prompts a "nonchirpy" consonant (Liberman and Mattingly, 1985).

The formant transition rate is critical for the perception of consonants by human listeners. For example, as shown in Figure 2, Liberman et al. (1956) reported that when steady formants were preceded by linearly rising formants, listener heard /be/ if the transition is fast and /we/ if it is slower (duration longer than 40 ms). With very slow transitions (duration > 100 ms), /ue/ was heard. When falling transitions were used, /ge/, /je/, and /ie/ were consequently heard as the transition slope decreased.

The slopes and directions of the second-formant transition are different for CV pairs with various consonants and vowels. The human auditory system can detect the fast frequency transition in the syllable-initial stop consonants, and thus can discriminate different stop consonants followed by different vowels. It is important to know how normal hearing listeners can use slope of frequency transition as a cue in the differentiation of CV or VC pairs. The way the auditory system process these dynamic sounds is of interest.
Figure 2: Spectrographic patterns with different transition directions and slopes. (After Liberman et al., 1956)

for psychoacoustical research.

1.1 Experiments of Discrimination in Frequency Transition

Elliott (Elliott et al., 1989) and Porter (Porter et al., 1991) examined the discrimination of speech-formant-like transitions in the presence of vowel-like steady states. Porter's results suggested that the discrimination differed as a function of transition direction - rising vs. falling. Both of them found that the signal duration had a primary effect on the discrimination. Discrimination was better for longer transitions. Porter (Porter et al., 1991) also reported that performance asymmetry existed in different adaptive directions. In Porter's second experiment, the shortest (30 ms) transitions displayed additional rate-related effects. The
JNDs for 30 ms transitions were smaller than those for the same-rate, 45 and 60 ms transitions. Discrimination was better for sharper transitions (comparing to the reference signal).

Elliott (Elliott et al., 1991) investigated the discrimination between transitions that were all rising or were all falling. They found that: (1) better discrimination occurred for transitions that diverge from a common frequency than for transitions that converge; (2) discrimination was better for longer transitions, when the rate of transition was fixed; (3) there was no discrimination difference between rising and falling signals.

Elliott and Porter used synthesized speech-formant-like stimuli (Klatt, 1980 and 1987) in their studies. However, frequency-modulated (FM) sinusoids have been used in many studies to approximate the range of the frequency excursion in the second formant transitions. In the present study, we assumed that the basic mechanism underlying discrimination of speech-like transitions can be illuminated by the study of simpler tonal glissandi. The signals were generated by moving the frequency over the same trajectories covered by formant transitions. We examined the listeners' abilities to extract and process information from dynamic sounds.

By using simple tone-glides, Cullen (Cullen et al., 1992) found that brief duration tone-glide discrimination was affected by the rate of transition and frequency-difference locus (initial vs. final). In addition, they suggested that
there were memory processes and/or backward masking according to the results of the differences between JNDs measured for initial and terminal frequency endpoints. The JNDs obtained for initial glides that fell in frequency were relatively "large". They explained this result by stating that the judgement was made on the basis of high frequency differences which were followed by low frequencies that swept across more basal regions. This led to an argument that a backward masking introduced by low frequency components affected the higher frequency, endpoint JND.

Dooley (Dooley and Moore, 1988) used FM sinusoids to examine the threshold changes in detecting frequency glides for different center frequencies and durations. In their study, subjects were asked to discriminate pairs of signals consisting of (1) two steady tones, (2) one steady tone and one up-glide, (3) one steady tone and one down-glide, and (4) one up-glide and one down-glide. They also reported that better discrimination occurred for longer durations and there is no significant difference in JNDs between up-glide and down-glide. Other studies using FM sinusoids include: (1) Collins and Cullen (1978) and Nabelek (1978), in which the detection thresholds of up-glides, down-glides and steady tones were measured; (2) Carlyon and Stubbs (1989) determining frequency modulation thresholds; and (3) Pollack (1968) who examined the thresholds for the detection of the direction and the rate of frequency change for pure tone-glides.
Some experimental results for the discrimination of the direction of frequency transition have been obtained in this laboratory (Neill, 1990; Feth, Neill and Hsu, 1991). A procedure was devised to determine the ability of listeners to discriminate between linear rising and falling frequency glides while varying the starting frequency of the glide signals. The roving frequency paradigm makes the discrimination task more like that involved in speech perception. Frequency Difference Limen (DLF) for all subjects were similar for fixed condition (no roving starting frequency) direction discrimination. There was a difference between thresholds for the fixed condition and those of the randomized conditions. The discrimination performance became poorer when randomized starting frequency procedure was introduced.

To create signals which approximate real formant transitions, a more complicated type of signal, the "moving filter", was introduced in the present study. This type of signal was generated by filtering a white Gaussian noise through a bandpass filter with linear moving center frequency. The filter center frequency moved linearly from a "starting" frequency to an "ending" frequency. The output of this operation was a "quasi-formant transition" and was named a "moving filter signal". In the present study, we examined the discrimination in slope of frequency transitions for FM sinusoids and moving filter signals.
1.2 Roving Starting Frequency Procedure

The initial and final pitch of signals causes a problem in the discrimination of frequency transitions. Nabelek, Nabelek and Hirsh (1970) illustrated the pitch cues associated with the initial and final frequencies of glide signals. Discrimination was influenced by pitch differences in the signals. In order to decrease the effect of the endpoint pitch cues, Pollack (1968) randomized the starting frequency of stimulus to check the end pitch effect. He concluded that if the randomized starting frequency procedure was employed, subjects respond to frequency changes over the duration of the signal rather than the initial or final pitch differences.

For the purpose of decreasing the effect of endpoint pitch cues, we designed an experimental procedure involving roving starting frequency. Here we refer to a procedure used in the study of intensity perception, in which a "roving" level of intensity is employed. In the intensity perception studies with the roving level paradigm, the listeners is required to detect a difference in intensity between two signals when the overall level of the signals varied between trials (Berliner and Durlach, 1973). Their procedure is different from the traditional fixed-overall-level procedure.

Running stop consonants do not have fixed starting frequencies in the formant transitions. Transition frequencies are different depending on the speaker and the sounds surrounding the transition. Thus, the roving frequency
paradigm should make the discrimination task more close to the processing of stop consonant transitions. The procedure used in this study is a frequency domain analog of the roving level paradigm used in intensity perception (Berliner & Durlach, 1973) and in profile analysis study (Green, 1988). One purpose of the present study is to determine the effect of random starting-frequency on the listener's ability to discriminate steep from flat frequency transitions.

In the present study, the starting frequencies of the glide tones are randomized in a two-cue, two alternative forced choice (2Q, 2AFC) adaptive procedure. The range corresponding to this starting frequency randomization is termed the "rove range". In order to examine the influence of randomization, the subjects were asked to distinguish between two linear glides differing only in the slope of their transitions.

1.3 Intensity Weighted Average of Instantaneous Frequency

By studying the discriminability of complementary pairs of two-tone complexes, Feth and coworkers (Feth, 1974; Feth and O'Malley, 1977; Feth et al., 1982) showed that pitch differences are proportional to the EWAIF (envelope weighted average of instantaneous frequency) differences between complex signals. There are some difficulties caused by the complexity of the EWAIF computation, especially for dynamic
(wideband) signals. The IWAIF (intensity weighted average of instantaneous frequency) model has been suggested as an alternative to the EWAIF model (Anantharaman, Krishnamurthy and Feth, 1992). The IWAIF model has comparable predictive performance to the EWAIF model. However, the computation involved in the IWAIF model is easier and much faster than that of the EWAIF model. Most important, the result of the IWAIF calculation is the "center of gravity" of the positive frequency portion of the signal spectrum. Therefore, in this study IWAIF was used to represent the "frequency" of signals.

Since the computation of IWAIF will involve using a finite, non-zero duration of the signal, a temporal integration mechanism is introduced for computing IWAIF. The temporal integration interval is the signal duration used to compute the IWAIF value. In this study, we use the short-term running IWAIF calculation. The idea is that given a signal \( s(t) \), \( 0 \leq t \leq T \), we compute a running IWAIF of the signal. The IWAIF value at time \( t_0 \) is computed based on the signal segment of \( s(t) \) between \( t_0 - T_w \leq t \leq t_0 \). Thus, \( T_w \) is the duration of the temporal integration (or short-term) window. Given a temporal window with fixed length, the short-term IWAIF calculation will generate a sequence of IWAIF values (IWAIF vector) for signal \( s(t) \). The IWAIF vector was used as the frequency track of any signal tested in this study.
1.4 The ASP's Excitation Pattern

Patterson and Holdsworth have developed the Auditory Sensation Processing (ASP) model (Patterson and Holdsworth, 1990). The ASP model is a functional model of the transformations which the auditory system employs to convert acoustic signals to the sensations that we hear in response to sounds. The ASP model has three main components which simulate various stages of processing of the auditory periphery. The first stage perform the spectral analysis by passing the incoming signal through a filter bank. The second stage contains temporal and cross-channel interaction to simulate the non-linear operations (compression, rectification, adaption and suppression) of the cochlea. Thus, the second stage serves as a "feature enhancement" mechanism. The third stage incorporates a triggered, quantized temporal integration to construct a "stabilized auditory image" (SAI). The SAI module can generate a sequence of images which is a representation of auditory/neural patterns of the dynamic spectral changes of complex signals. Located at the end of the second stage of ASP model, the software module, "genepn", can generate neural excitation patterns for any sound. The output of this module is named "EPN".

The excitation pattern is a representation used to describe the pattern of neural activity evoked by a given sound. It displays "neural activity" as a function of the characteristic frequency (CF) of the neurones being excited.
Psychoacoustically, the excitation pattern of a sound can be defined as the output of the auditory filters (Moore and Glasberg, 1983).

In this study, the ASP model served as a preprocessor to the IWAIF calculation. Both the short-term fast Fourier transform (FFT) and the excitation pattern for a given signal were used as the input to the IWAIF calculation. Since both the short-term FFT and the excitation pattern generate appropriate representations in frequency domain for linear glide signals, the results of using FFT and EPN were expected to be similar.

1.5 Computational Prediction Model for Discrimination of Frequency Transitions

In this study, we used a computational model to predict the listener's performance in discriminating frequency transitions. The computational model was based on the short-term IWAIF calculation and the independent channels (independent-noise) model proposed by Durlach et al. (1986). The assumptions of the independent-noise model can be summarized as follows. There is a linear filter bank which resolves the input signal into N frequency bands (or channels). The internal noise is Gaussian and independent across channels. The variance of the noise in each channel is independent of the signal presented. The central processing is ideal, so that the decision variable can be represented by the
logarithm of the likelihood ratio of the target probability to the reference probability. They pointed out that the channels did not have to be frequency bands. The consideration still hold if the channels represent a sequence of independent time intervals.

In the present study, the channels of the independent-noise model were regarded as independent time intervals. In each time interval, a short-term IWAIF calculation was performed and thus a sequence of IWAIF values (IWAIF vector) was generated for a frequency transition stimulus. Basically, the computational model derived a "sensitivity index" $d'$ for two different signals using their IWAIF vectors. The $d'$ was used as a measure of the discriminability between two signals. Given two signals, $s_1(t)$ and $s_2(t)$, the sensitivity index $d'$ is defined as (Durlach et al., 1986)

$$d' = \frac{1}{\sigma} \left[ \sum_{j=1}^{N} \Delta_j^2 - \frac{\sigma_R^2}{\sigma^2 + \sigma_R^2} \left( \sum_{j=1}^{N} \Delta_j \right)^2 \right]^{1/2}$$

(1)

where

$$\Delta_j = \log_2 (IWAIF_{1,j} - IWAIF_{2,j} + 1)$$

$$\sigma = \frac{\Delta \sqrt{N}}{d'}, \quad \Delta = \log_2 (\text{frequency DL} + 1)$$

$$\sigma_R = \frac{\log_2 (R_f + 1)}{\sqrt{12}}, \quad R_f = \text{roving frequency range.}$$

(2)

In the above expression, $IWAIF_1$ and $IWAIF_2$ are the IWAIF vector for $s_1(t)$ and $s_2(t)$ respectively, and frequency DL is the
"difference limen for frequency" which is obtained from the listening task of 79.4% correct performance (d'=1.16). The IWAIF vectors are the input of the computational model and the "frequency DL" and the "rove range" served as parameters in the model. Note that the definition in Equation (2) is not a normal definition of d'. The normal definition of d' is the ratio of the difference of means to their common standard deviation. The definition we used is the ratio of the log of mean difference to the log of their common standard deviation. A block diagram showing the model of computing d' for two signals is given in Figure 3. The results obtained in the listening experiment were then compared with the performance predicted by the model. The detailed discussion about the comparison between the model and listeners are described in Chapter III and IV.

1.6 Statement of Problem

Most of the studies in frequency transition discrimination have been motivated by the fact that the identification of many consonants in the speech stream depends upon frequency transitions. The tasks include detecting frequency transitions and discriminating among glides containing different transition rates, durations and directions. However, very few studies have been conducted to examine the transition slope discrimination with a roving starting-frequency paradigm. And, no study has provided a
computational model predicting listener performance in transition slope discrimination. The experiments conducted in this study were focused on examining the discriminability of slope in frequency transitions. The signals used in this study were FM sinusoids and moving filter signals.

The main purpose of this research was:

1. to investigate listener performance in the discrimination of frequency transition slope using FM
sinusoids and signals generated by a digital bandpass filter with moving center frequency.

(2) to examine the discriminability of frequency transition under different rove ranges of starting frequency.

(3) to develop a statistical decision theory model which can predict human listener performance in the discrimination of frequency transitions and.

(4) to evaluate the computational model by comparing model predictions with listener performance in transition slope discrimination.

The model will provide a new computational technique for estimating the discrimination of frequency transitions. The techniques include (1) calculating the IWAIF vectors for the Fourier transform and neural excitation pattern, and (2) applying statistic detection theory on the IWAIF vectors. The research will yield valuable and important results for psychoacoustics and consonant perceptions.
CHAPTER II
LITERATURE REVIEW

This chapter contains a review of literature relevant to the topics studied in this research. It includes the discrimination of frequency modulated tones, detection and discrimination of slope of frequency transitions, and the randomization of stimulus parameters to counteract an artifact. The Patterson’s Auditory Sensation Processing (ASP) model (Patterson and Holdsworth, 1990) will be reviewed. The development of the Intensity Weighted Average of Instantaneous Frequency (IWAIF) model from the Envelope Weighted Average of Instantaneous Frequency (EWAIF) model is provided. Finally, the application of the independent channels model is discussed.

2.1 Discrimination of Frequency Transitions

Elliott (Elliott et al., 1989) evaluated the discrimination of frequency transitions using rising and falling single-formant simulations. Subjects were asked to discriminate two transition stimuli containing different onset frequencies but same offset frequencies. All stimuli were generated with the Klatt (1980, 1987) synthesizer. Onset
frequencies of the rising transitions varied linearly from 942 to 1146 Hz in 17-Hz steps. Onset frequencies of the falling transitions varied linearly from 1772 to 1340 Hz in 36-Hz steps. All transitions ended at 1240 Hz. Signals contained a transition alone and signals contained a transition followed by a steady-state tone were examined. Stimulus durations were 30, 60, and 120 ms. Their data showed that: (1) some improvement in discrimination occurred after practice but primarily only for the stimuli with the shortest duration; (2) there was no discrimination difference between transition-only stimuli and transitions followed by steady-state sound; (3) there was no difference in discrimination between rising and falling transitions; and (4) discrimination was significantly better for the longest transitions. They also found that the just noticeable differences (JNDs) for the longest transitions, measured in Hz at transition onsets, were of approximately the same magnitude as JNDs for steady-state sounds having frequencies equal to midpoints of the transitions. They also indicated that their data supported the sampling theory of glissando perception (e.g., Pollack, 1968; Dooley and Moore, 1988).

Elliott et al. (1991) examined the discrimination of second-formant-like frequency transitions. They used four types of stimuli, two were rising transitions and the other two were falling transitions. For rising or falling transition, stimuli were classified as “SO” (signals had same
frequency onset but different frequency offsets) and "SF" (signals had same frequency offset but different frequency onsets). Signals were generated using Klatt (1980, 1987) synthesizer. Signal durations were 60 and 120 ms, and the rate of transition was the same for long and short signals. They reported that better discrimination, measured in Hz/ms, occurred for longer stimuli and for transitions that diverged from a common frequency. There were no difference between discrimination of rising and falling transitions. They indicated that their results suggested that, under that same conditions, discrimination of initial stop consonants that precede the same vowel should be poorer than discrimination of final consonants that follow the same vowel. However, some other researcher have reported the opposite results (e.g., House et al., 1965).

Porter et al. (1991) investigated the discrimination of formant transition onset frequency for different signal durations. The signals they used were analogs of second-formant-transition resonances together with steady-state, vowel-like resonances (modified Klatt routine). A 2Q, 2AFC adaptive procedure was employed and the subjects were asked to identify the interval (second or third) in which the different stimulus (with different transition slope) occurred. Their data showed that JNDS decreased with increase in transition duration. They suggested that longer transitions were associated with better frequency resolution and better
use of pitch and timbre cues. They also found that falling transitions were better discriminated than rising, and JNDS for discrimination of transitions with rates greater than the reference transition are smaller than those with rates less than the reference transitions (an asymmetric discrimination performance for different adaptive directions, from-above or from-below). In addition, the shortest duration (30 ms) signals displayed a rate-related effects. Under the fixed rate-of-transition (5 Hz/ms) condition, the JNDS for the shortest transition (30 ms) were smaller than those for the 45 nad 60 ms transitions. Porter et al. suggested that the small JNDS for shortest transitions may be due to the differences in the degree of dispersion of activity in the cochlea that high-rate transitions are effectively treated as non-time-varying wideband signals.

Cullen (Cullen et al., 1992) used isolated, linear, rising and falling 200 Hz, 40 ms reference tone-glides centered at 1950 Hz, and examined discrimination of target tone-glides with initial or final endpoint frequency differences. They also compared the discrimination of tone-glides using the references with different transition rates (40 ms glides with 2.5, 5.0, and 10 Hz/ms transition rates). Their data also showed that discrimination was better for the tone-glides with frequency differences at final endpoints; there was no significant difference for JNDS as a function of whether the glides were rising or falling; there
was no discrimination difference for the target signals with slower or faster transition rates than the reference signal. Note that it is not a contradiction with the asymmetric discrimination reported in Porter's experiment (1991); Cullen et al. didn't use different adaptive direction in their experiment, they compared the performance between rising and falling transitions. Cullen et al. also found that the JNDs increased with the increase in the transition rate of the reference glide. According to their results on the difference between the JNDs measured for initial and final frequency endpoints, they suggested that there were memory processes and/or backward masking in the discrimination of short tone-glides with high rate of frequency change. The JNDs obtained for initial glides that fell in frequency were relatively "large". They explained this result by stating that the judgement was made on the basis of high frequency differences which were followed by low frequencies that swept across more basal regions. This led to an argument that a backward masking introduced by low frequency components affected the higher frequency, endpoint JND.

Dooley and Moore (1988) used FM sinusoids to examine the threshold changes in detecting frequency glides for different center frequencies and durations. In their study, subjects were asked to distinguish two signals in a pair. Four types of signal-pairs are used: (1) two steady tones, (2) one steady tone and one up-glide, (3) one steady tone and one down-glide,
and (4) one up-glide and one down-glide. The thresholds for these four conditions were referred to as DLFs, UP-DLs, DOWN-DLs, and UPVSDOWN-DLs. They reported (1) better discrimination occurred for longer durations; (2) no significant difference between UP-DLs and DOWN-DLs; (3) no significant different between DLFs and UPVSDOWN-DLs; and (4) considerably higher UP-DLs and DOWN-DLs than either DLFs or the UPVSDOWN-DLs.

Collins and Cullen (1978) examined the detection of frequency transitions using tone glides. They claimed that the signals they used were analogous to the formant transitions of isolated stop consonants. The glides were in the frequency ranges of 200-700 Hz for F1 and 1200-1700 Hz for F2. Signal durations were 5 - 120 ms. Detection thresholds were measured for rising glides, falling glides, and steady-state tones. They reported that (1) the threshold for falling glide was higher than that of rising glide when duration was short ( < 35 ms); (2) the thresholds for both glides were higher than that of the fixed tone by an average of 3.3 dB for the F2-like transitions and 4.4 dB for the F1-like transitions.

Pollack (1968) investigated the discrimination of the direction of frequency change for pure tones. Initial frequencies were from 125 to 1000 Hz. Signal durations were 0.5, 1, 2 and 4 seconds. Subjects were asked to indicate the direction of the transition in a one interval task. The starting frequency of the glides was randomized so that the
responses were based on the direction of the frequency transition rather than on the frequency differences at the end of the signals. His data indicated that the subjects were able to discriminate the transition direction, and they did not rely on the comparison of the ending frequency to a mean starting frequency. Pollack stated that the listeners must have responded to the frequency transition along the signal duration when the starting frequency was randomized. He also found that there is a 1.3-fold and 1.8-fold increase in the threshold of frequency change for two different subjects when a random starting frequency was used.

Nabelek (1978) measured thresholds for rising glides, falling glides, and steady-state tones. The data showed that, in general, the thresholds for falling glides were higher than both those of rising glides and steady-state tones. Nabelek suggested that the differences could result from different time courses of neural decay and inhibition for constant tones, rising tones, and falling tones. The differences between rising and falling tones indicated that the phase spectra could influence signal detectability. Tsumura et al. (1973) examined the thresholds for the detection of frequency transitions. The frequency of the stimuli was linearly changed from 1000 Hz either during the whole signal duration or during a portion of it. Subjects were required to determine the just-noticeable-frequency difference of a steady tone and a tone containing either rising or falling transitions. They
reported that the thresholds for rising transitions were higher than those of falling transitions.

Some experimental results for the discrimination of the direction of frequency transition have been obtained in this laboratory (Neill, 1990; Feth, Neill and Hsu, 1991). They found that as the width of the roving starting frequency range were increased, subjects' thresholds for direction discrimination (DDT) were increased. The curves of psychometric functions remained parallel even when the starting frequency range was changed. This indicated that the subjects did not change their criterion of discrimination as the size of roving frequency changed. A roving frequency paradigm was designed to determine the ability of listeners to discriminate between linear rising and falling frequency glides while varying the starting frequency of each up glide signal and the ending frequency of each down glide signal. The roving frequency paradigm makes the discrimination task more like the processing of the auditory system that is involved in speech perception. Frequency DLs (Difference Limen) for all subjects were similar under fixed (no roving starting frequency) direction discrimination task. There were differences in DDTs between the fixed condition and the roving-frequency conditions. The direction discrimination performance became poorer when randomized starting frequency procedure was used.
2.2 Randomization of Stimulus Parameters

Some studies using frequency modulated (FM) sinusoidal tones indicated that the listener may use pitch cues derived from the frequency difference between endpoints to help in the determination of the direction of linear frequency transitions (Carlyon and Stubbs, 1989; Nabelek et al., 1970). Nabelek et al. (1970) found that pitch was matched near both the initial and the final frequencies for short signals having small frequency changes. When the duration and frequency transition increased, the pitch match moved toward the final frequency of the signal.

Henning (1965) employed a randomization technique in a study of frequency discrimination of high frequency (1000 - 15000 Hz) signals. He thought that in a frequency difference limen task, subjects reported loudness rather than frequency difference. In order to reduce the effect of the frequency dependent loudness cues, Henning (1965) employed a random-amplitude paradigm in his experiment. On each trial of a 2AFC procedure, one of the test tones, selected randomly, was attenuated by a random amount ranging from 8 to 20 dB in 2 dB steps. A poorer performance in frequency discrimination, as compared with the data obtained by Shower and Biddulph (1931), is observed above 3000 Hz when a random amplitude paradigm was used. This result suggested that the listeners use loudness differences at very high frequency (> 3000 Hz) to make decision for the frequency discrimination task.
Berliner and Durlach (1973) used a roving level procedure in the study of intensity perception. In their listening task, the subject was asked to detect an intensity difference between two sinusoidal signals. The overall level of the signals was randomly varied between trials. Green (1988) used a roving level paradigm in the study of profile analysis. Green examined the listener ability to detect changes in spectral shape. The profile analysis task required the listener to detect an increment in intensity of one single tone in a multi-tone complex array. Adding the increment to one of the tone in the array changes the shape of the spectrum. To prevent subjects from performing the task by monitoring the magnitude of the output of the single auditory filter centered at the frequency of the incremented component, the overall level of the whole stimulus was varied randomly from one stimulus to the next, over a 40 dB range. This roving level paradigm makes the magnitude of the output of any single filter an unreliable cue to the signal. The subjects were forced to make their decision based on the difference in overall spectral shape of the arrays. It is referred to as a "simultaneous" comparison rather than a "successive" comparison.

2.3 The Patterson's ASP Model

The human auditory system is a very sophisticated signal processor. A general model of peripheral auditory system can
be divided into three stages. The first stage is the simulation of the movement of the basilar membrane. The second stage would be the characteristics of the hair cells. The third stage is the response of the auditory nerves and the central auditory system. Many computational model has been created to simulate the function of the auditory periphery. One type of the auditory model is based on the simulation of the displacement of the basilar membrane (e.g., Payton, 1988; Beet, 1990). This type of auditory periphery model implements equations of motion to simulate the displacement of the basilar membrane. Since this type of model involves solving sets of differential equations, the computational load is usually heavy. The one-dimensional transition-line model was also introduced to simulate the motion of the basilar membrane (e.g., Deng and Geisler, 1987).

A more efficient auditory model for the first stage of the cochlear function is based on both psychophysical and physiological knowledge about the human auditory perception (e.g., Lyon, 1982, 1984; Seneff, 1988; Martens and Immerseel, 1990; Patterson and Holdsworth, 1990, 1991). This type of model simulates the first stage of cochlea action based on the auditory filter characteristics derived behaviorally from normal-hearing humans. In general, this type of auditory model uses a filter-bank to approximate the motion of the basilar membrane. The shape of the human auditory peripheral filter was well determined by Patterson (1974, 1976). The equivalent
rectangular bandwidth (ERB) of the auditory filter was measured by Patterson and Moore (1986; Moore, 1986). Patterson and Holdsworth have developed the Auditory Sensation Processing (ASP) model (Patterson and Holdsworth, 1990). The ASP model is a functional model of the transformations which the auditory system employs to convert acoustic signals to the sensations that we hear in response to sounds.

The ASP model contains three main components. The first stage is a spectral analysis, which is performed by passing the incoming signal to a "bank of gammatone auditory filters" (Patterson et al., 1988) and the output is a simulation of basilar membrane motion.

However, the hair cells are not just passive transducers. In particular, it has been suggested that the outer hair cells have an active function which enhances features that arise in the basilar membrane motion. The second stage contains temporal and cross-channel interaction to simulate the non-linear operations (compression, rectification, adaptation and suppression) of the cochlea. Patterson called the processing of the second stage as "feature enhancement". The transduction and feature enhancement are performed by a "bank of adaptive threshold generators". It includes a bank of logarithmic compressors and a bank of adaptation units that apply adaptation to the motion of membrane simultaneously in time and frequency domain. The operation of this structure is named as "two-dimensional adaptation". The output of the
second stage of the ASP model is a simulation of the neural activity pattern that flows from the cochlea up the auditory nerve to the cochlear nucleus.

The third stage incorporates a triggered, quantized temporal integration to convert the neural activity pattern into a "stabilized auditory image" (SAI). The SAI module can generate a sequence of images which is a representation of auditory/neural patterns of the dynamic spectral changes of complex signals. If the sound is perceived as steady, the SAI is stationary. If the sound has a dynamic spectrum, for example, formant transitions, the SAI changes dynamically to match the normal listener's "perceptions" of the spectral dynamics. Since the auditory image contains two-dimensional data, the rate of the data output of SAI is considerable high.

2.4 The IWAIF and EWAIF Model

The envelope weighted average of instantaneous frequency (EWAIF) model was originally developed to predict the discriminability of two-tone complexes (Feth, 1974). Feth and coworkers (Feth, 1974; Feth and O'Malley, 1977; Feth el al., 1982) had studied the discriminability of complementary pairs of two-tone complexes, which represented the simplest modulated tones in Voelcker's (1966a,b) unified modulation theory. They reported that the pitch differences in the complementary complex-tone pairs are proportional to the EWAIF differences between complex signals. The EWAIF for a signal
$s(t)$ $0 \leq t \leq T$ is defined, in the time domain, as (Anantharaman, Krishnamurthy and Feth, 1992)

$$EWAIF[s(t)] = \frac{\int_{0}^{T} e(t) f(t) \, dt}{\int_{0}^{T} e(t) \, dt}$$  \hspace{1cm} (3)

where $e(t)$ is the instantaneous envelope and $f(t)$ is the instantaneous frequency (IF).

The EWAIF model has been used to explain a variety of discrimination tasks in which the dominant cue is the spectral pitch of the signals. For example, Feth and Stover (1987) used the EWAIF model to explain an anomaly in the data of "profile analysis", in which greater sensitivity to an increment of one component amplitude was observed when additional components were added (Green, 1988). Feth and Stover tried to use EWAIF model to quantify the pitch change observable in the discrimination of complex stimuli. They indicated that the increment in amplitude of one component in the profile signal produced a difference in the EWAIF which listener could perceive as a noticeable pitch change. One drawback of the EWAIF model is that it is difficult to compute, especially for wideband signals. In the calculation of EWAIF, the envelope and the instantaneous frequency of the signal have to be determined separately. This involves taking the derivative of instantaneous phase and some times requires division of two near-zero numbers.
The intensity-weighted average of instantaneous frequency (IWAIF) model has been suggested as an alternative to the EWAIF model (Anantharaman, Krishnamurthy and Feth, 1992). The difference between the IWAIF and EWAIF is that the signal envelope is used to weight the instantaneous frequency (IF) in the EWAIF calculation, while the signal intensity (proportional to envelope-squared) is used to weight the IF in the IWAIF calculation. The IWAIF model has comparable predictive performance to the EWAIF model. However, the computation involved in the IWAIF model is much easier and faster than that of the EWAIF model.

The IWAIF of signal $s(t)$ is defined as

$$IWAIF[s(t)] = \frac{\int_0^T e^2(t) f(t) \, dt}{\int_0^T e^2(t) \, dt} \quad 0 \leq t \leq T$$

(4)

where $e(t)$ is the instantaneous envelope and $f(t)$ is the instantaneous frequency of the signal $s(t)$. As shown by Anantharaman et al. (1992), a much more convenient representation of the IWAIF of $s(t)$ was obtained for the frequency domain

$$IWAIF[s(t)] = \frac{\int_0^\infty f |S(f)|^2 \, df}{\int_0^\infty |S(f)|^2 \, df}$$

(5)

where $S(f)$ is the Fourier transform of $s(t)$.

One of the advantages of the IWAIF model is that the IWAIF can be calculated in the frequency domain using fast
Fourier transform (FFT). Suppose $s(t)$ is sampled at a rate $F_s$ and the $N$-point FFT of the signal $s(t)$ is $S[k]$, $k=0,1,...,N-1$. Then, the IWAIF of the $s(t)$ can be computed as:

$$IWAIF[s(t)] = \frac{\sum_{k=0}^{(N/2)-1} k\Delta f |S[k]|^2 \Delta f}{\sum_{k=0}^{(N/2)-1} |S[k]|^2 \Delta f}$$

$$= \Delta f \frac{\sum_{k=0}^{(N/2)-1} k|S[k]|^2}{\sum_{k=0}^{(N/2)-1} |S[k]|^2}$$

where $\Delta f = F_s/N$ is the frequency spacing between samples of the FFT.

The IWAIF provides both computational efficiency and an intuitive interpretation in the frequency domain. The IWAIF of a complex signal is the "center of gravity" of the positive portion of the signal spectrum (Anantharaman, Krishnamurthy and Feth, 1992).

2.5 The Independent Channels Model used for Discrimination predictions

Plomp (1970) examined the characteristics of timbre influenced by amplitude pattern and phase pattern for complex tones. In his study of the relation between timbre and amplitude, Plomp proposed a multi-channel model to calculate the distance in frequency spectrum between tones. The timbre difference between two tones with same loudness and pitch was expressed as
\[ D = \sqrt[p]{ \sum_{n=1}^{m} |L_{1,n} - L_{2,n}|^p } \]  

where \( L_{i,n} \) is the SPL of tone \( i \) in frequency band \( n \), and \( m \) is the total number of frequency bands. He also investigated the correlation between dissimilarity indices (measured from listening task) and the distance calculated by the above equation for different \( p \) values. He reported that the difference in frequency spectrum expressed in distance in a Euclidean space \((p=2)\) or in area between the spectrum curves \((p=1)\) correlated well with the differences in timbre observed by the listeners.

Florentine and Buus (1981) designed a multi-band (multi-channel) version of Zwicker's (1956, 1970) excitation-pattern model for intensity discrimination for pulsed sounds. They also evaluated the single-band model proposed by Zwicker. Florentine and Buus assumed, for the multiband model, that the performance of the intensity discrimination was determined by an optimum decision based on the information in all critical bands. The Zwicker's single-channel model assumed that the performance was determined by the critical band containing the fastest change in excitation with changing in stimulus level.

Florentine and Buus's model for intensity discrimination has four basic assumptions: (1) the excitation level in each critical band is defined by the Zwicker's excitation pattern; (2) the excitation level discrimination in the critical band is
independent across all bands; (3) the sensitivity-per-Bel of excitation level difference within a critical band is independent of excitation level (Weber’s law); (4) the sensitivity-per-Bel is a constant for all bands. The multiband model proposed by them for predicting the performance of intensity discrimination was written as

\[ d' = \sqrt{\sum_{i=1}^{24} d'_i^2} \]  

where

\[ d'_i = k \Delta L_i \]  

is the sensitivity in channel \( i \), \( \Delta L_i \) is the excitation level difference in channel \( i \), and \( k \) is a free parameter used to adjust the prediction to fit the listener’s data. Florentine and Buus used the model to predict the intensity discrimination for various stimuli: pure tones, partially masked tones, and wideband noise, and compared the predictions with some experiment data obtained from other researchers (e.g., Rabinowitz et al., 1976). Their data showed that the predictions made by the single-band model of Zwicker (1956, 1970) were usually qualitatively correct, but it can not account for some data in intensity discrimination of pulsed sounds. The multi-channel model, except at high frequencies, made predictions in good qualitative and quantitative agreement with the experiment data.
Durlach, Braida and Ito (1986) proposed an application of the independent channels (or independent-noise) model for the processing of broadband signals. Their basic independent-noise model was very similar to the model proposed by Plomp (1970) and Florentine and Buus (1981). The model contains a linear filter bank which resolved the input into \( N \) independent frequency bands (or channels). The internal noise added to each channel is Gaussian and independent across channels. The variance of the noise in each channel is independent of the stimulus presented. If an ideal central processing is assumed, the sensitivity index \( d' \) as a measure of the discriminability between two stimuli is derived as

\[
d' = \left[ \sum_{j=1}^{N} \left( \frac{\Delta_j}{\sigma_j} \right)^2 \right]^{1/2}, \quad \Delta_j = M_{ij} - M_{2j},
\]

where \( M_{ij} \) is the mean value of stimulus \( i \) in channel \( j \). They generalized the simple model by adding a common noise \( R \) representing the interchannel correlation to incorporate the roving-level condition employed in some discrimination experiments. Usually, the randomized roving-level is introduced by experimenter to force the subject to make decisions based on interchannel comparison rather than on the single-channel information. If the variance in each channel is assumed to be equal, the sensitivity index \( d' \) with the consideration of interchannel correlation is given by
where $\sigma_n^2$ is the variance of interchannel common noise $R$. Furthermore, they also included "central noise" into the model. The central noise was divided into two types: noise added after the decision stage, and noise added before the decision stage. They indicated that the model could be applied to any type of channels (e.g., frequency channel or time interval). The model is independent of whether monaural or binaural signal is presented and whether amplitude or phase information is measured.

Gagne and Zurek (1988) studied the ability of normal-hearing listeners to discriminate changes in the frequency of a single resonance filter. Beside investigating the dependence of the discrimination ability on the frequency, filter bandwidth and source signal, they also devised an auditory filter-bank model to account for the psychoacoustic data. In their experiment, the just-noticeable differences in resonance frequency were measured for stimuli generated with different frequencies and bandwidth using two types of source signals: sawtooth signal and flat-spectrum noise. Gagne and Zurek analyzed their results by using a general filter-bank model which was similar to the model proposed by Florentine and Buus (1981) and Durlach et al. (1986). Their model contained 19 peripheral bandpass filters. The spacing between adjacent center frequency was 1/3 oct. The shape of those bandpass filter was derived by Patterson (1974). Each filter

$$d' = \frac{1}{\sigma} \left[ \sum_{j=1}^{n} \Delta_j^2 - \frac{\sigma_n^2}{\sigma^2 + \sigma_n^2} \left( \sum_{j=1}^{n} \Delta_j \right)^2 \right]^{1/2}$$

(11)
was followed by square-law rectification, integration over the signal duration, and logarithmic conversion. A zero-mean Gaussian random variable representing internal noise was added to each band after the log conversion. The noise was assumed to independent with equal variance across bands. They also indicated that the variance of the internal noise provided a fitting parameter to the experiment data. Since there was a fitting parameter in the model, they could fit the model's prediction to the listener's performance very satisfactorily. They used two different schemes in their model: multi-channel (or spectral integration) and single-channel (or best-band) methods. The model derived a quantity $D$ which is a measure of the difference between two single-resonance stimuli. For the multi-channel model, $D$ was defined as

$$D(F, \Delta F) = \sqrt{\sum_{i=1}^{19} [L_i(F) - L_i(F+\Delta F)]^2},$$

(12)

where $L_i(F)$ is the level in channel $i$ resulting from a signal with resonance frequency $F$. For the best-band model, $D$ was defined as

$$D(F, \Delta F) = \max_i |L_i(F) - L_i(F+\Delta F)|.$$

(13)

Note that these two types of model are very similar to the multi-channel and single-band models described in Florentine and Buus's paper (1981). Gagne and Zurek reported that the best-band model accounted for the discrimination results better than the multi-channel model. Their conclusion did not
agree with the results of some other studies (e.g., Green, 1983; Florentine and Buus, 1981), in which spectral-integration (multi-channel) model had better performance.
CHAPTER III
METHODS

In this chapter, the numerical procedure for IWAIF calculations is described. Then, signals used in the study are specified. A description of the instrumentation is provided. Details of the design of the psychoacoustic experiment are given. Finally, the computational model used to predict listener performance is described in detail.

3.1 IWAIF Calculation

In this study, the IWAIF (intensity weighted average of instantaneous frequency) calculation was applied to both the Fourier transform and the EPN (Neural Excitation Pattern) of the test signals. In this section the calculation of IWAIF values will be discussed in detail.

Given a finite energy real signal $s(t)$, the IWAIF of $s(t)$ is defined as

$$\text{IWAIF}[s(t)] = \frac{\int_0^T e^2(t) f(t) \, dt}{\int_0^T e^2(t) \, dt} \quad 0 \leq t \leq T$$  \hspace{1cm} (14)

where $e(t)$ is the instantaneous envelope and $f(t)$ is the instantaneous frequency of the signal (Anantharaman,
It can be shown that the IWAIF can equivalently be computed as:

$$IWAIF[s(t)] = \frac{\int_{0}^{\infty} |S(f)|^2 df}{\int_{0}^{\infty} |S(f)|^2 df}$$  \hspace{1cm} (15)$$

where \( S(f) \) is the Fourier transform of \( s(t) \).

3.1.1 Computation of IWAIF using the Fourier transform

The above frequency domain representation (15) provides a simple and efficient procedure for computing the IWAIF of a signal. To compute the IWAIF value, all we need to obtain is the Fourier transform of \( s(t) \). This can be done by using the FFT algorithm. Suppose \( s(t) \) is sampled at a rate \( F_s \) and the \( N \)-point FFT of the \( j \)th frame (computing the short-term FFT) of signal \( s(t) \) is \( S[j,k], k=0,1,\ldots,N-1 \). Then, the IWAIF of the \( j \)th frame of \( s(t) \) can be computed as:

$$IWAIF_j[s(t)] = \frac{\sum_{k=0}^{(N/2)-1} k\Delta f |S[j,k]|^2 \Delta f}{\sum_{k=0}^{(N/2)-1} |S[j,k]|^2 \Delta f}$$  \hspace{1cm} (16)$$

where \( \Delta f = F_s / N \) is the frequency spacing between samples of the FFT. Furthermore, the IWAIF calculation based on the
Trapezoidal Rule, \( \int_a^b f(x) \, dx = \frac{(b-a)}{2} [f(a) + f(b)] \), can be written as:

\[
I_{WAIF_j}[s(t)] = \frac{\sum_{k=0}^{N/2-2} (k\Delta f|S[j,k]|^2 + (k+1)\Delta f|S[j,k+1]|^2) \Delta f}{\sum_{k=0}^{N/2-2} (|S[j,k]|^2 + |S[j,k+1]|^2) \Delta f}
\]

where \( \Delta f = \frac{F_s}{N} \) is the frequency spacing between samples of the FFT.

3.1.2 Computation of IWAIF using excitation pattern

In the ASP model (Patterson and Holdsworth, 1990), the filterbank is linear. The output of the filterbank is rectified and compressed logarithmically to represent the cochlear nonlinearity. Since the EPN output of the ASP model is represented on a quasi-log scale, the EPN data has to be exponentiated before we calculate the IWAIF value for EPN. This gives large values their proper weight. A detailed description of extracting EPN data from the ASP model is given in appendix A. Let \( EPN[k] \) be the EPN output obtained directly
from the ASP model, and $epn[k]$ be the exponentiated EPN data, the relation is:

$$epn[k] = 10^{EPN[k]/20}.$$  \hspace{1cm} (19)

Now, the computation of the IWAIF of excitation for a signal $s(t)$ can be described as follows.

Since all of the signals used in this study contain frequency transitions, sampling the EPN leads to a sequence of frames of excitation. Let $epn[j,k]$ be the jth frame of the excitation pattern of $s(t)$. Each frame of excitation contains $N$ points ($k = 0, 1, \ldots, N-1$), which are the excitation values of the $N$ channels. The IWAIF of the jth frame of the excitation of $s(t)$ can be represented as:

$$IWAIF_j[s(t)] = \frac{\sum_{k=0}^{N-1} f_k |epn[j,k]|^2 \Delta f_k}{\sum_{k=0}^{N-1} |epn[j,k]|^2 \Delta f_k}$$ \hspace{1cm} (20)

where $f_k$ is the center frequency of the $k^{th}$ channel and $\Delta f_k$ is the frequency spacing between channels. A more accurate calculation based on the Trapezoidal Rule for the numerical IWAIF integration is given below:

$$IWAIF_j[s(t)] = \frac{\sum_{k=0}^{N-2} (f_k |epn[j,k]|^2 + f_{k+1} |epn[j,k+1]|^2) (f_{k+1} - f_k)}{\sum_{k=0}^{N-2} (|epn[j,k]|^2 + |epn[j,k+1]|^2) (f_{k+1} - f_k)}.$$ \hspace{1cm} (21)
The above equation (21) was used for computing IWAIF in this study.

According to the above discussion (section 3.1.1 and 3.1.2), one IWAIF number will be obtained for each frame of signal. Therefore, if the signal \( s(t) \) has \( j \) frames, \( j \) IWAIF values will be obtained. The \( j \) IWAIF values can be written as a vector

\[
IWAIF[s(t)] = \begin{bmatrix}
IWAIF_1[s(t)] \\
IWAIF_2[s(t)] \\
\vdots \\
IWAIF_j[s(t)]
\end{bmatrix}
\]

where \( IWAIF_j[s(t)] \) is the IWAIF value for the \( j^{th} \) frame of \( s(t) \). This vector, \( IWAIF[s(t)] \), is called the **IWAIF vector** of signal \( s(t) \). For example, given a 100 ms tone glide with frequency transition from 1000 to 1600 Hz, the **IWAIF vector** calculated based on the FFT using a 10 ms hamming window becomes \( [ 1035.13 \ 1083.56 \ 1138.64 \ 1194.36 \ 1256.07 \ 1318.80 \ 1381.31 \ 1443.57 \ 1506.07 \ 1561.42 ]^T \). Generally speaking, the **IWAIF vector** will approximate the contour of frequency transition of the signal. The **IWAIF vector** is then used in developing the computational model that predicts the human listener performance in discriminating different frequency transitions.
The computer programs used for the IWAIF computation were created using MATLAB on the SUN workstation and TURBO Pascal on the IBM 386 PC.

3.2 Signals and Parameters

In this section the signals used in this study are specified and their parameters are defined. Also, the equipment used to generate signals (on-line and off-line) is described.

3.2.1 FM Sinusoids with Linear Frequency Modulation

Sinusoids with linear frequency modulation (FM) were used to test listeners' ability to discriminate the slope of frequency transitions. The signals were rising-frequency linear glides. The frequency excursion is defined as the difference between the starting frequency and the ending frequency of the signal. The slope of the signal is defined as the frequency excursion divided by the duration of the signal and is expressed in Hz/ms. Listeners were required to distinguish between two glides of equal duration that differed only in frequency excursion.

Because multi-listener adaptive procedures and roving frequency paradigms were included in our experiment, it was necessary to generate "real time" stimuli rather than stored signals. Using stored signals would require long disk I/O time.
and lots of disk storage. Therefore, in the listening experiments, signals were generated in real-time ("on-line") using an Ariel DSP-16 signal acquisition board (Ariel corporation, 1987) mounted in a Zenith PC/XT computer. All signals were generated at a 100 KHz sampling rate, and low-pass filtered at 8 kHz. To smooth the onset and offset of the signals, all signals were gated on and off with a Wilsonics (model BSIT) programmable cosine gate. The programs used to generate signals were implemented in TMS320C25 assembly language (Texas Instruments Incorporated, 1987) and could run independently on the Ariel DSP-16 board. The detailed information about generating signals in real-time using DSP-16 are given in the report written by Hsu and Feth (1992).

The signals used to test the computational decision model were generated "off-line" on an IBM compatible 386 PC or on a SUN workstation. Since we needed to compare the performance of the computational model with the performance of the human listeners, the signals used to test the model and human subjects were the same. The difference was that one was generated "on-line" and the other one was generated "off-line".

In the transition slope discrimination test, the starting frequency for each tone glide was selected at random from a uniform distribution centered at 1000 Hz. The width of this uniform frequency distribution is defined as the "rove range".
The rove ranges examined in this study include 0, 100, 200, 400, and 800 Hz. Three signal durations: 100 ms, 50 ms, and 25 ms, were tested. The frequency excursion of the reference tone glide was fixed at 400 Hz. The initial frequency excursion of the target tone glide was 400 Hz above or below the reference excursion. Here, "above" means the frequency excursion of the target is 800 Hz, initially. "Below" means the frequency excursion of the target is 0 Hz, initially. In the adaptive procedure, the frequency excursion of the target glide was increased (adapted from below) or decreased (adapted from above) toward the reference glide with three consecutive corrective correct responses, and increased or decreased away from the reference glide with each incorrect response. The adaptive procedure will be discussed latter in this chapter.

3.2.2 Frequency Transition Signals Generated by Linear Moving Filters

A more complicated type of signal, the "moving filter signals", was used to run the human subjects and test the computational decision model. A white Gaussian noise with zero-mean was filtered by a bandpass filter with a moving center frequency. The filter center frequency moved linearly from a "starting" frequency to an "ending" frequency. The output of this operation was a "quasi-formant transition". This process is shown in Figure 4.
White Noise Generator

A/D

Real-time Digital Filter with linear Moving Center Frequency bandwidth=70Hz gain=40dB

D/A

Attenuator and BSIT gate

Figure 4: Block diagram of generating real-time moving filter signals.
Since the filter used to generate these signals required the center frequency to be changed with time, a type of time-varying digital audio filter (Mourjopoulos, Kyriakis-Bitaros and Goutis, 1990) was used. The control parameters (center frequency, bandwidth, and gain) of the filter could be varied with time. However, in this study, only the center frequency was changed with time. For bandpass variable parameter filters, second-order transfer functions are required. Thus, the transfer function with time-varying coefficients $A_i(n)$ and $B_i(n)$ at the sampling instant $n$ has the form

$$H(z, n) = \frac{A_0(n) + A_1(n)z^{-1} + A_2(n)z^{-2}}{1 + B_1(n)z^{-1} + B_2(n)z^{-2}} \quad \text{(23)}$$

where each variable coefficient is a function of $\{g(n)\}$: filter gain, $\omega_0(n)$: center frequency, $W(n)$: bandwidth). Thus, the filter output $y(n)$ for any input $x(n)$ can be realized by the difference equation

$$y(n) = \sum_{i=0}^{2} A_i(n)x(n-i) - \sum_{i=1}^{2} B_i(n)y(n-i). \quad \text{(24)}$$

In this study, the filter gain and bandwidth were selected to be fixed at 40 dB and 70 Hz, respectively. The 70 Hz bandwidth was chosen because it is typical of second format bandwidths for speech (Klatt, 1977; Klatt, 1980). The coefficients $A_i(n)$ and $B_i(n)$ are defined by
\[ A_0(n) = \frac{1}{2} \{ 1 + a(n) + k(n) - [k(n) a(n)] \} \]
\[ A_1(n) = B_1(n) = b(n) [1 + a(n)] \]
\[ A_2(n) = \frac{1}{2} \{ 1 - k(n) + a(n) + [k(n) a(n)] \} \]
\[ B_2(n) = a(n). \]

The variables \( a(n), b(n), \) and \( k(n) \) are given by

\[
\begin{align*}
  a(n) &= \frac{1 - \tan \{ \pi W(n) / f_s \}}{1 + \tan \{ \pi W(n) / f_s \}} \\
  b(n) &= -\cos \{ 2\pi f_0(n) T_s \} \\
  k(n) &= 10^{g(n)/20}
\end{align*}
\]

where \( f_0(n) \) is filter center frequency, \( W(n) \) is filter bandwidth fixed at 70 Hz, and \( g(n) \) is filter gain fixed at 40 dB.

For the "off-line" signal generation, the moving filters described above were implemented on the SUN workstation as well as on the IBM 386 PC. The signals used for the listening experiments were generated on the Ariel DSP-16 board in real-time. The TMS320C25 assembly program which implements moving filters is listed in appendix B.

### 3.2.3 Summary of signals

The signals used in this study are summarized in Table I. The frequency excursions of the reference signals for both the tone glides and moving filter signals were fixed at 400 Hz. A 2Q, 2AFC (two-alternative force-choice) adaptive procedure with roving frequency paradigm was conducted for both adaptive directions (from above and below) in this study. The frequency
excursions for the target signals were higher (when adapted from above) or lower (when adapted from below) than 400 Hz. The starting frequencies of all signals were selected at random from within the rove range. A detailed description of the adaptive procedure employed for transition slope discrimination is given in section 3.4.5.

Table 1: Summary of signals.

<table>
<thead>
<tr>
<th>Duration</th>
<th>Direction</th>
<th>Tone glide</th>
<th>Moving filter signal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rove range of</td>
<td></td>
<td>Rove range of starting</td>
</tr>
<tr>
<td></td>
<td>starting frequency</td>
<td></td>
<td>frequency</td>
</tr>
<tr>
<td>100ms</td>
<td>adapted from above</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
<tr>
<td></td>
<td>adapted from below</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
<tr>
<td>50ms</td>
<td>adapted from above</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
<tr>
<td></td>
<td>adapted from below</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
<tr>
<td>25ms</td>
<td>adapted from above</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
<tr>
<td></td>
<td>adapted from below</td>
<td>0, 100, 200, 400, 800 Hz</td>
<td>0, 100, 200, 400, 800 Hz</td>
</tr>
</tbody>
</table>

3.3 A Computational Prediction Model

The IWAIF vector described in section 3.1 is used to determine the difference in slope of frequency transitions between pairs of signals. For a given signal, an IWAIF vector is generated, and the IWAIF vector is regarded as a "representation" of the signal. The IWAIF vectors for two
different signals are then used to compute a statistic for the discrimination of frequency transition.

**Figure 5:** Independent channels model using IWAIF.

The detection model is based on the scheme proposed by Durlach, Braida and Ito, (1986). Basically, the model derives the sensitivity index $d'$ which is a measure of the
discriminability between two stimuli. The independent channels model adopted in this study is illustrated in Figure 5. This model is essentially a typical independent-noise model which has been proposed many times (Plomp, 1970; Florentine and Buus, 1981 and Durlach, Braid and Ito, 1986). The present model assumes that the IWAIF calculation is performed in the auditory peripheral mechanism. The IWAIF value is obtained directly from the Fourier transform of each stimulus. Since the computational model is based on the IWAIF and psychoacoustic data (e.g. frequency DL) rather than the physiological data (e.g. displacement of the basilar membrane), the proposed model is essentially a functional model of cochlear processing.

In this study, the IWAIF calculation provides a kind of "feature enhancement" processing which was addressed in the ASP model (Patterson and Holdsworth, 1990). The short-term Fourier transform was performed by applying a finite temporal Hamming window on the signals. The Hamming window was used to reduce the sidelobes. The process of IWAIF calculation was assumed to be continuous along the signals. Given a signal with duration of T ms, the continuous IWAIF values were sampled every n millisecond, where n is the size of the time window. This would create N IWAIF values (N=T/n) for a signal, which was shown in Figure 5: the time channel 1 to N. Typically, the n was selected to be 10.
It is assumed that a zero mean Gaussian random variable representing internal noise is added to each time channel (see Figure 5). Further, we assume that the internal noise is statically independent across time channels. Also, we assume the variance of the noise in each channel is independent of which stimulus is presented. Finally, as usual, it is assumed that the central processing is ideal within the constraints imposed by the internal noise.

3.3.1 Slope Discrimination for Fixed Starting Frequency

Under the assumptions of the independent-noise model, the "IWAIF pattern" for a signal can be represented by an "IWAIF vector": \([IWAIF_1 \ldots IWAIF_n]\) where the \(IWAIF_j\) is the mean of the statistically independent Gaussian random variable corresponding to time channel \(j\). The variance of the random variable for time channel \(j\) is defined as \(\sigma^2_j\). Suppose we have two glide signals, \(s_1\) and \(s_2\), with the same starting frequency and IWAIF vector \([IWAIF_{11} \ldots IWAIF_{1n}]\) and \([IWAIF_{21} \ldots IWAIF_{2n}]\) respectively. The sensitivity index \(d'\) can be written as

\[
d' = \left[ \sum_{j=1}^{N} \left( \frac{\Delta_j}{\sigma_j} \right)^2 \right]^{1/2}, \quad \Delta_j = \log_2(|IWAIF_{1j} - IWAIF_{2j}| + 1). \tag{27}
\]

In equation (27), the definition of \(\Delta_j\) should be noted. The differences between IWAIF values are scaled logarithmically. Some simpler expression such as \(\Delta_j = IWAIF_{1j} - IWAIF_{2j}\) were tried, but the logarithmic scale provides more
reasonable values for $d'$. We put the term "$+1$" in the definition of $\Delta_j$, because we want to avoid computing $\log(0)$ when two IWAIF values are equal. In addition, the $\sigma_j$ in equation (27), should be discussed in more detailed. The $\sigma_j$ are the standard deviations of IWAIF values for the individual time channels. The IWAIF of a real signal is located at the "center of gravity" of its energy density spectrum. Thus, for a tone-glide, the $\sigma_j$ can be seen as the standard deviation of the instantaneous frequency $f_j$ at a specific instant $j$. In equation (27), if we consider the frequency discrimination of two steady (fixed frequency) tones, $\Delta_j$ is fixed and will be equal to the frequency difference limen (frequency DL) of the two tones. The frequency DLs vary with different center frequencies and different stimulus durations (Moore, 1972 and 1973). However, for simplicity of the model, we assume the frequency DLs are fixed over the frequency range (600 - 1400 Hz) concerned in this study. If we assume the $\sigma_j$ is also fixed along the signals, it can be obtained from equation (27) by making use of the value of frequency DL:

$$d' = \left[ \sum_{j=1}^{N} \left( \frac{\Delta_j}{\sigma_j} \right)^2 \right]^{1/2}, \quad \Delta_j = \log_2(|F_1 - F_2|+1) = \Delta, \quad \sigma_j = \sigma$$

(28)

where $F_1 - F_2$ is the frequency DL. Thus, we have

$$\sigma_j = \sigma = \frac{\Delta \sqrt{N}}{d'}, \quad \Delta = \log_2(\text{frequency DL} + 1) \text{ and } 1 \leq j \leq N.$$ 

(29)
where the $d'$ is the sensitivity index corresponding to a 2AFC (2-alternative forced choice) procedure for which the subject achieves 79.4% correct responses. The value 79.4% is obtained from the up-down adaptive procedures (Levitt, 1971). The up-down procedure will be described later. The value of $d'$ for 2AFC with 79.4% correct response rate is about 1.16 (Hacker and Ratcliff, 1979).

Note that equation (27) is useful only for the discrimination of fixed starting-frequency conditions. If a roving starting-frequency paradigm is considered, the computation is more complicated.

3.3.2 Slope Discrimination for Roved Starting Frequency

For the roving frequency paradigm, the IWAIF discrimination model described above can be modified in a way suggested by Durlach, Braida and Ito, (1986). A common noise variable $R$ representing the interchannel correlation is added to the variables $IWAIF_j$. It is assumed that $R$ is a statically independent, Gaussian random variable with mean zero and variance $\sigma_R^2$. The random variable $R$ can be interpreted as the result of a random starting frequency introduced in our experiment to force the listener to make interchannel comparisons (i.e., to attend to frequency transition slope) rather than respond on the basis of single time channel
information (i.e., respond to the frequency cue at the onset or offset of the stimulus).

After we introduce the random variable $R$, the IWAIF vector of a signal can be written as: $[IWAIF_1 \ldots IWAIF_N]$ where the $IWAIF_j$ are the mean of the Gaussian random variable corresponding to time channel $j$. The Gaussian random variables corresponding to the time channels have variances $\sigma^2_j + \sigma^2_R$, and covariances $\sigma^2_R$. Again, we assume for simplicity that

$$\sigma_j = \sigma, \quad 1 \leq j \leq N. \quad (30)$$

Under this assumption, the sensitivity index $d'$ for the roving frequency case is given by

$$d' = \frac{1}{\sigma} \left[ \sum_{j=1}^{N} \Delta_j^2 - \frac{\sigma^2_R}{\sigma^2 + \text{No}^2} \left( \sum_{j=1}^{N} \Delta_j \right)^2 \right]^{1/2} \quad (31)$$

where

$$\Delta_j = \log_2 (|IWAIF_{1j} - IWAIF_{2j}| + 1),$$

$$\sigma = \frac{\Delta \sqrt{N}}{d'}, \quad \Delta = \log_2 (\text{frequency DL} + 1),$$

$$\sigma_R = \frac{\log_2 (R_f + 1)}{\sqrt{12}}, \quad R_f = \text{roving frequency range}. \quad (32)$$

It should be noted that the variance of the common noise variable $R$ is taken as the variance of a random variable uniformly distributed over the interval of the range of roved starting frequency. That is, the probability density of $R$ is assumed to be Gaussian in the model but is approximated by a rectangular distribution.
3.3.3 Model Evaluation and Computer Realization

The sensitivity indices $d'$ for signals with different frequency transitions were obtained for various roving frequency ranges. From these data, we could find the required frequency transition differences (in Hz) for which the model predicted 79.4% correct response ($d' = 1.16$ in our study) for different roving frequency ranges. Then, the frequency transition differences predicted by the model was compared with the results obtained from the listeners.

The second method we used to evaluate the model was to calculate the efficiency of the listener (EL). The efficiency of the listener is defined as:

$$EL = \frac{d'}{d'_{m}} \times 100\%$$

where $d'_{m}$ is obtained by converting the listener's frequency transition differences to the corresponding sensitivity indices predicted by the model, and $d'$ is 1.16 for the 3up-1down adaptive procedure used in our study. Then, the efficiency of listener (EL) was used as an index which represent the percentage of discrimination efficiency of the subject relative to the model.

The computational prediction model was simulated by running computer programs. The computer programs were written based on MATLAB on the SUN workstation as well as by TURBO PASCAL on the IBM 386 PC. The flowchart showing the procedures
involved in using the computational prediction model as a predictor and the comparison between the model's performance and listener's result is shown in Figure 6. The listening task employing human subjects will be discussed in the next section.

3.4 Experiment using human listener

For the purpose of comparing our model's prediction to the human's listening ability, we conducted a frequency slope discrimination experiment using stimuli containing frequency transitions. In this section details of the experiment, including subjects, procedures, and equipment will be described.

3.4.1 Subjects

Three listeners with normal hearing (pure tone threshold better than 15 dB HL re ANSI 1969) participated in the study. The age of the subjects ranged from 20 to 25 years. All three listeners were female. Prior to data collection, the listeners were practiced at the task for two weeks (20 hours). The purpose of training the listeners was to let them become familiar with the characteristics of the stimuli, the environment and equipment, and the adaptive procedure employed in the study. Two of the listeners were undergraduate students in the Division of Speech and Hearing Science of The Ohio State University, and the other one was a student with
Figure 6: Diagram of the Procedures of the prediction model and how the model’s prediction is compared to the listener’s performance.

engineering background. All the subjects were paid $5 per hour plus a bonus of $1 per hour for the completion of the experiment.
3.4.2 Signals and Equipment

The stimuli used in the study have been described in Section 3.2. In summary, two types of signals were examined in the study: pure tone-glides (sinusoids with linear frequency modulation) and moving filter signals (white noise filtered by a linear moving filter). All the signals were "rising" (starting at low frequency and ending at high frequency) linear glides.

The frequency excursion of the reference signal (in a 2Q, 2AFC adaptive procedure) was fixed at 400 Hz. The starting frequencies of all signals were selected at random from a uniform distribution centered on 1000 Hz. The frequency rove ranges (the width of the uniform distribution) examined in the study were 0, 100, 200, 400, and 800 Hz. Signal durations were 100, 50, and 25 ms. The signals used in the study are summarized in Table 1.

Signals were generated "on-line" using an Ariel DSP-16 signal processing board mounted in a Zenith 159 microcomputer (IBM PC compatible). All signals were generated at 100 KHz sampling rate, and low-pass filtered at 8 KHz. Signals were gated on and off with a Wilsonics (model BSIT) cosine gate to make the signals be smooth at onset and offset. A block diagram of the equipment is given in Figure 7.

For all task, listeners were seated in separate sound isolated rooms facing a monitor and a small computer (Radio Shack Color II). The small computer was used to accept each
**Figure 7:** Block diagram of equipment for frequency slope discrimination task and frequency DL task. Note that the white noise generator is used only for generating moving filter signals.
subject's response and display the correct response feedback. All signals were presented at 50 dB SL through one side of a Sennheiser HD414SL headset. The subjects made their responses by pressing buttons on a response box. Correct response feedback was displayed immediately after each trial.

The experiment was controlled by the laboratory computer. Computer programs were implemented so that the 3 subjects were able to listen to different signals at the same time, and the computer could record responses separately. To perform this task, all the equipment controlled by computer program has to communicate with each other. The detailed descriptions of those programs are given in (Hsu and Feth, 1992). Figure 7 shows the block diagram for the instrumentation set up.

3.4.3 Detection Threshold for Signals Presented in Quiet

Detection thresholds for 50 ms signals were measured. A 2AFC adaptive procedure was employed. There were two intervals for each trial. Only one interval contained the signal. For each trial, the listener was asked to judge whether the first or the second interval had the signal. The order of the target interval was varied randomly from trial to trial. The magnitude (in dB) of the signal was reduced by 5 dB (or 3 dB when the up-down data became stable) at three consecutive correct responses and increased by 5 dB (or 3 dB) at any wrong response. With this paradigm, the quiet threshold was defined
as 79.4% correct performance (Levitt, 1971). Detection thresholds were measured for both the pure tone-glides and the moving filter signals. For the pure tone-glides, a 50 ms glide with frequency transition from 900 to 1100 Hz was used as the test signal. For the moving filter signals, a 50 ms sweep with transition from 1 kHz to 1.4 kHz was used.

Data were collected for 6 blocks of 50 trials each. For each block, a single estimate of threshold was obtained by averaging the level (dB) at the reversal points. A reversal point is defined as the point where the adaptive direction changes ("up to down" or "down to up"). In order to remove the bias of the starting point and keep even numbers of the reversal points, the first 2 or 3 reversal points are discarded. The signals used for frequency slope discrimination task were presented at 50 dB SL through the earphones.

3.4.4 Frequency Difference Limen

The difference limen for frequency (frequency DL) was necessary for the computation of our prediction model (see Section 3.4). Therefore, a measurement of the frequency DL was obtained for each subject. A 2Q, 2AFC (two-cue, two-alternative forced-choice) adaptive procedure (Levitt, 1971) was employed. Subjects were asked to identify the signal that was higher in pitch. There are 4 intervals for each trial. Only one of the middle two intervals contained the signal with
higher frequency (the target), the other three intervals presented the same lower frequency. The frequency of the target signal was decreased by a factor of two with three consecutive correct responses, and increased by a factor of two with any incorrect response. With this up-down procedure, the frequency DL is obtained for which the subject achieves 79.4% correct performance. Signals were presented at 50 dB SL.

The frequency DL for each subject was examined at frequencies of 600, 1000, and 1400 Hz for signal durations at 25, 50, and 100 ms. For each combination (frequency x duration), data were obtained from at least 6 blocks, or 300 trials (could be more than 6 blocks if subject's performance had large variance) of listening task. For each block, a single estimate of the frequency DL was calculated by averaging the values at the remaining reversal points by discarding the first 2 or 3 points. A reversal point is defined as the same as that described in the previous section.

3.4.5 Transition Slope Discrimination

A 2Q, 2AFC adaptive procedure, roving frequency paradigm was conducted for this experiment. Each trial contained four intervals. One stimulus was presented in each interval. Only the second or third interval contained the target which contained a different slope in frequency transition. The other three reference stimuli contained the same transition slopes.
The subject was asked to determine which interval (2nd or 3rd) contained the target. The starting frequency of each glide in each interval was randomly selected from a roving frequency range.

The interstimulus interval was 400 ms. The response interval was 1000 ms. In addition, there was a 600 ms delay between two trials.

Data were collected in three blocks of 50 trials each. Three subjects were able to listen at the same time, and response were recorded separately. The difference in frequency excursion between the target and reference was decreased by a factor of 2 after three consecutive correct responses, and increased by a factor of 2 following any incorrect response. Feedback of correct response was given immediately after each trial. By using this 3up-1down adaptive procedure, the transition slope discrimination threshold (TSDT) is an estimate of the 79.4%-correct point on the psychometric function in a constant-stimulus procedure.

A reversal is defined as the point where the direction changes. The TSDT was calculated by averaging the reversals after the first 2 or 3 points were deleted. Depending on the subject's performance, data were collected from six or more blocks for each signal condition. Subject performance was checked daily for improvement between session. If the improvement or large variance was noted, more blocks of listening task were provided.
Figure 8: Two adaptive directions in the up-down adaptive procedure.

Two adaptive directions (from-above and from-below) were examined in this study. In the “adapted-from-above” procedure, the ending frequency of the target was typically 400 Hz higher than that of the reference when the procedure started. That is the target frequency excursion was 800 Hz and the reference frequency excursion was 400 Hz, initially. The target stimulus was then adapted to approach the reference stimulus “from above” with more and more correct responses from subjects. In the “adapted-from-below” procedure, the ending frequency of the target was typically 200 Hz (or 400 Hz depending on the difficulty of the task) lower than that of the reference when the procedure started. That is the target frequency excursion was 200 Hz and the reference frequency excursion was 400 Hz, initially. The target stimulus was then adapted to approach the reference stimulus “from below” with more and more correct
responses from subjects. These two adaptive directions are illustrated in Figure 8.

Two types of signals: pure tone-glides and moving filter signals were examined in the frequency slope discrimination test. The data from human listeners were compared with the model's predictions.
CHAPTER IV
RESULTS

This chapter consists of two parts. The first part includes a presentation of the results of the listener discrimination of frequency transition slope for the pure tone-glide and moving filter stimuli. Also contained in the first part are the results of the listener frequency DL measures and detection thresholds. The second part of this chapter contains the performance of the IWAIF prediction model and a comparison between the model predictions and experimental results. Data are presented graphically. Some data are also tabulated in order to examine the precise values.

4.1 Psychoacoustic Data for Listening Experiment

4.1.1 Detection Threshold

The listener detection thresholds for pure tone-glides and moving filter signals were measured. The detection threshold is defined as the signal level resulting in 79.4% correct performance in a 2AFC 3up-1down adaptive procedure (Levitt, 1971). The detection thresholds (measured in dB SPL) for the three subjects are shown in Table 2. For subject #1
and #2, the detection thresholds for tone glide are smaller than those for the moving filter signal. But for Subject #3, the detection threshold for tone glide is larger than that for moving filter signal. Subject #3 may need more practice in the task for pure tone glides. The signals used for the other frequency discrimination tasks were presented 50 dB above the detection threshold of a given signal for each subject.

Table 2: Detection Threshold (in dB SPL) for Signals Presented in Quiet.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Pure tone glide</th>
<th>Moving filter signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>10.6</td>
<td>15.7</td>
</tr>
<tr>
<td>#2</td>
<td>13.9</td>
<td>17.7</td>
</tr>
<tr>
<td>#3</td>
<td>12.1</td>
<td>8.5</td>
</tr>
</tbody>
</table>

4.1.2 Frequency Difference Limen

Frequency difference limen (frequency DLs) for pure tones of different durations and center frequencies were determined to serve as a parameter in the computational model. They also provide a check for fixed-condition frequency slope discrimination. DLs were measured for each subject at 3 different center frequencies (600, 1000, 1400 Hz) and at 3 different durations (25, 50, 100 Hz). The frequency DLs represent the frequency difference for which the subject
achieves 79.4% correct response in a 2Q, 2AFC adaptive procedure.

The frequency DLs for all subjects across different signal conditions are shown in Figure 9 and Table 3. Since we are interested in the value of DLs for each subject, the data are plotted separately for individual subject.

Table 3: Listener Frequency DLs measured (and standard deviations) in Hz.

<table>
<thead>
<tr>
<th>subject</th>
<th>duration (ms)</th>
<th>center frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>600</td>
</tr>
<tr>
<td>#1</td>
<td>100</td>
<td>5.88(8.05)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5.94(2.70)</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>10.73(1.82)</td>
</tr>
<tr>
<td>#2</td>
<td>100</td>
<td>5.85(1.38)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>16.83(13.39)</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>22.44(6.45)</td>
</tr>
<tr>
<td>#3</td>
<td>100</td>
<td>4.63(2.35)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5.60(6.92)</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>13.23(3.83)</td>
</tr>
</tbody>
</table>

Frequency DLs obtained in this study are slightly larger than some previous measurements (e.g., Dooley & Moore, 1988). Their data were taken using 500 ms signals. Since DLs increase with the decrease in duration, larger values of frequency DLs are expected. In addition, the DLs in this study were obtained at the 79.4% correct point on the psychometric function, while Dooley et al. used the 70.7% correct point on the psychometric function.
Figure 9: Results of frequency DLs measured at 79.4% correct performance ($d' = 1.16$).
A two-factor analysis of variance, with center frequency (CF) and duration (D) as factors was performed for each subject. For subject #1, the analysis revealed a significant effect of CF \[ F(2, 50) = 13.43, \ p < 0.001 \]. The duration effect and interaction between CF and D were not significant. For subject #2, CF \[ F(2, 47) = 25.77, \ p < 0.001 \], D \[ F(2, 47) = 40.28, \ p < 0.001 \] and the interaction \[ F(4, 47) = 8.73, \ p < 0.001 \] were all significant effects. For subject #3, only duration was a significant effect \[ F(2, 54) = 11.40, \ p < 0.001 \], CF and the interaction were not significant. Multiple comparisons were also carried out for different levels of CF and duration for each subject. In general, for all three subjects, the values of frequency DL for 600 and 1000 Hz were very close. Frequency DLs were significantly higher at 1400 Hz for two subjects (#1 & #2). When the comparison was made across subjects, general speaking, subject #2 has the largest frequency DLs, and the DLs for subject #1 and #3 are very close.

Although the statistic analysis showed that the frequency DLs have higher values at 1400 Hz than that measured at lower frequencies, the difference can probably be reduced by giving more practice to the subjects. (By comparing some preliminary data obtained in this study to these final data shown in Table 3, we found that the differences of frequency DLs between these three center frequencies were decreased when more practices were give to subjects.) Therefore, for the simplicity of our computational model, the frequency DLs
measured at 1000 Hz were chosen to be the parameter used in the model. Note that the duration effect is still considered in the model, so different frequency DLs were used for different signal durations.

4.1.3 Frequency transition slope discrimination

The transition slope discrimination threshold (TSDT) was obtained for 79.4% correct performance by conducting a 2Q, 2AFC adaptive procedure with roving starting frequency. Each individual subject’s performance will be plotted separately in order to identify similarities across subjects. The functions for TSDT versus rove range are plotted for each subject.

4.1.3.1 Linear tone glides. Each subject’s performance (TSDT vs. rove range) is plotted in Figure 10 - Figure 12 for signal durations of 100, 50, and 25 ms. The results for the two adaptive directions (adapted from above and from below) are presented separately in order to indicate the asymmetries of different adaptive direction on slope discrimination. The fixed (no roving frequency) condition was plotted individually (at rove range = 0.1 Hz) in each figure. Different symbols represent different subjects: • - subject #1, ▼ - subject #2, and ■ - subject #3. Rove range is plotted on a logarithmic scale. Each datum point is derived from the average TSDT score for at least six blocks of 50 trials each.
From the data shown in Figure 10 – Figure 12, we can find that: (1) for all the conditions, TSDTs increase with the increase in rove range, (2) poorer discrimination occurs for the adaptive direction from below than for adaptation from above, (3) all three subjects have similar TSDT curve patterns, (4) the TSDT curves have similar shapes for different signal durations, but longer signals show slightly better performance, and (5) in general, the TSDT scores are consistent with the performance in frequency DLs; subjects having smaller frequency DLs perform better in the slope discrimination.

Table 4: TSDT (in Hz) for the tone glide slope discrimination. The center frequency of rove range is 1000 Hz.

<table>
<thead>
<tr>
<th>subject</th>
<th>duration (ms)</th>
<th>rove range (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>#1</td>
<td>100</td>
<td>39.4</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>60.6</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>59.9</td>
</tr>
<tr>
<td>#2</td>
<td>100</td>
<td>57.3</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>85.8</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>122.3</td>
</tr>
<tr>
<td>#3</td>
<td>100</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>47.8</td>
</tr>
</tbody>
</table>

In order to obtain a single TSDT value for each signal condition, the data from the two adaptive directions were averaged geometrically, although performance asymmetries
Figure 10: Transition slope discrimination threshold (TSDT) for 100 ms tone glides.
Figure 11: Transition slope discrimination threshold (TSDT) for 50 ms tone glides.
Figure 12: Transition slope discrimination threshold (TSDT) for 25 ms tone glides.
exist. The averaged TSDT for all signal conditions are shown in Table 4 for each subject.

The data shown in Table 4 are plotted in Figure 13. The data are plotted separately for the three subjects in order to see the differences among the subjects and the duration effect for each subject. A two-factor, repeated measures analysis of variance (ANOVA), with rove-range and duration as fixed factors and the data blocked across subjects (a random independent variable), revealed a significant effect of rove-range \( F(4, 8) = 83.59, p < 0.0001 \). The TSDT values increase with the increase in rove-range. The duration effect was not significant \( F(2, 4) = 1.42, p = 0.3422 \). Discrimination performance is not affected by the signal durations. There was a weak interaction between rove-range and duration \( F(8, 16) = 3.06, p = 0.0272 \). This was caused by the observation that the averaged TSDT for 25 ms was larger than that for 100 ms at 0 rove range but the averaged TSDT for 25 ms was smaller than that for 100 ms at 800 Hz rove range. In general, the TSDT values for subject #1 and #3 were close; subject #2 had higher TSDT values. Since rove-range was significant factor in the ANOVA, paired comparisons were carried out between specific means to determine which of the levels were significantly different from each other. This was done by performing a post hoc t-test, equivalent to Fisher's least-significant-difference test, on the means for each independent variable. A summary of the results are presented
Figure 13: Transition slope discrimination threshold for tone glides. Data are plotted separately for three subjects.
in Table 5.

Table 5: Pairwise comparison matrix for different levels of the two factors: rove-range and duration, and the random variable: subject. The asterisks indicate the significant level of the difference between a given means: ***=\( p<0.0001 \); **=\( p<0.001 \); *=\( p<0.01 \); NS = not significant. Signal used: tone glides. The numbers in parentheses are the TSDT difference between two different rove ranges.

<table>
<thead>
<tr>
<th>rove-range</th>
<th>0</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>rove-range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>.</td>
<td>***(42)</td>
<td>***(71)</td>
<td>***(120)</td>
<td>***(172)</td>
</tr>
<tr>
<td>100</td>
<td>.</td>
<td>**(29)</td>
<td>**(78)</td>
<td>**(130)</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>.</td>
<td></td>
<td>***(49)</td>
<td>***(101)</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>.</td>
<td></td>
<td></td>
<td>***(52)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The slope discrimination performances for subject #1 and #3 are very close. The performance for subject #2 is poorer than that for the others. This is consistent with their frequency DL measures (A larger frequency DL leads to poorer performance in transition slope discrimination). Rove-range is a significant factor on the discrimination performance. The multiple comparison for rove-range showed that the TSDT for different rove-ranges are significantly different from each other. The effect of signal duration is not significant. The data are plotted in a different way (in Figure 14) to show the difference between subjects for three durations. It can be seen that subject #2 has somewhat poorer performance and the results for the other two subjects are very close.
Figure 14: Transition slope discrimination threshold for tone glides. Data are plotted separately for three signal durations.
Elliott et al. (1991) examined the discrimination of second-formant-like frequency transitions. Our TSDT data for the fixed condition are generally consistent with this earlier study. Discrimination is better for the signals with longer durations. The TSDTs obtained in the present study for 100 ms signals are slightly smaller than the JNDs measured at 70% correct performance for 120 ms in Elliott’s report. The reason of this difference can be considered that the reference signal used in Elliott’s experiment having very low transition rate (0.2 Hz/ms) but the reference signal we used having higher transition rate (4 Hz/ms). The methods used to generate the stimuli and the experiment procedures also differed.

4.1.3.2 Moving filter signals. The transition slope discrimination threshold (TSDT) was measured for moving filter signals. The results are shown in Figure 15 to Figure 17 for signal durations of 100, 50, and 25 ms respectively. Again, the data for fixed condition are plotted separately in the figures. The same symbols are used for each listener as before. Rove range is shown on a logarithmic scale. Each datum point represents an average TSDT score obtained from at least six blocks of 50 trials each.

Data from the two adaptive directions are plotted separately. Again performance asymmetries occurred for all signal durations and for all subjects. Discrimination performance was better for the adapted-from-above condition.
From these figures, we can see that: (1) the shape of TSDT curve for the moving filter signal is similar to that for the pure tone glide, (2) discrimination becomes poorer when rove-range increases, (3) discrimination is better for longer signals; duration has a stronger effect in moving filter signals than that in tone glides, (4) all three subjects have similar TSDT patterns, and (5) in general, the TSDT scores are consistent with the results of frequency DLs; listeners having smaller frequency DLs perform better in the slope discrimination.

Again the TSDTs from both adaptive directions were averaged to obtain a single TSDT measure for each signal condition. The averaged TSDTs for all the signal conditions are given in Table 6.

Table 6: TSDT (in Hz) for the moving filter slope discrimination. The center frequency of rove range is 1000 Hz.

<table>
<thead>
<tr>
<th>subject</th>
<th>duration (ms)</th>
<th>rove range (Hz)</th>
<th>0</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#1</td>
<td>100</td>
<td>108.7</td>
<td>126.9</td>
<td>166.7</td>
<td>224.1</td>
<td>268.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>124.0</td>
<td>184.3</td>
<td>222.2</td>
<td>236.3</td>
<td>275.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>191.8</td>
<td>200.5</td>
<td>211.8</td>
<td>260.2</td>
<td>334.5</td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>100</td>
<td>144.2</td>
<td>176.7</td>
<td>209.2</td>
<td>238.7</td>
<td>283.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>192.5</td>
<td>198.1</td>
<td>239.8</td>
<td>289.2</td>
<td>295.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>238.7</td>
<td>242.0</td>
<td>248.6</td>
<td>290.0</td>
<td>336.6</td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td>100</td>
<td>103.9</td>
<td>133.9</td>
<td>178.0</td>
<td>203.1</td>
<td>290.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>111.1</td>
<td>139.7</td>
<td>187.7</td>
<td>220.1</td>
<td>290.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>168.9</td>
<td>184.1</td>
<td>195.6</td>
<td>276.6</td>
<td>326.5</td>
<td></td>
</tr>
</tbody>
</table>
The data shown in Table 6 are plotted in Figure 18. The data are plotted separately for individual subject in order to see the duration effect for each subject. A two-factor, repeated measures analysis of variance (ANOVA) was conducted. Rove-range and duration were fixed factors and the data were blocked across subjects (a random independent variable). The analysis revealed a significant effect of rove-range \([F(4, 8)=73.14, \ p<0.0001]\), and duration \([F(2,4)=54.31, \ p=0.0013]\). The TSDT scores increase with the increase in rove-range, and the discrimination performance is better for longer signals. Statistical analysis showed that there was a weak interaction between rove-range and duration \([F(8,16)=2.92, \ p=0.0323]\). It appears from Figure 18 that the difference between durations may be large at narrow rove-ranges but slightly smaller at wide rove-ranges. This could account for the interaction between rove-range and duration. The analysis also showed that the performances of the three subjects are significantly different.

Since rove-range and duration were significant factors in the ANOVA, paired comparisons (post hoc t-test) were performed between specific means to determine which of the levels were significantly different from each other. The multiple comparison for each independent variable was done in the same way described in the previous section. A summary of the results is presented in Table 7.
Frequency slope discrimination
100ms 400Hz sweep, slope=4Hz/ms
moving filter signals

**Figure 15:** Transition slope discrimination threshold (TSDT) for 100 ms moving filter signals.
Frequency slope discrimination
50ms 400Hz sweep, slope=8Hz/ms
moving filter signals

Figure 16: Transition slope discrimination threshold (TSDT)
for 50 ms moving filter signals.
Frequency slope discrimination
25ms 400Hz sweep, slope=16Hz/ms
moving filter signals

Figure 17: Transition slope discrimination threshold (TSDT)
for 25 ms moving filter signals.
Figure 18: Transition slope discrimination threshold for moving filter signals. Data are plotted separately for three subjects.
Table 7: Pairwise comparison matrix for different levels of the two factors: rove-range and duration, and the random variable: subject. The asterisks indicate the significant level of the difference between a given means: ***=p<0.0001; **=p<0.001; *=p<0.01; NS = not significant. Signal used: moving filter signals. The numbers in parentheses are the TSDT difference between two different levels.

<table>
<thead>
<tr>
<th>rove-range</th>
<th>0</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>rove-range 0</td>
<td>**(22)</td>
<td>***(53)</td>
<td>***(95)</td>
<td>***(146)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.</td>
<td>***(31)</td>
<td>***(73)</td>
<td>***(124)</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>.</td>
<td>.</td>
<td>***(42)</td>
<td>***(93)</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>***(51)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>duration</th>
<th>25</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>duration 25</td>
<td>.</td>
<td>***(33)</td>
<td>***(57)</td>
</tr>
<tr>
<td>50</td>
<td>.</td>
<td>.</td>
<td>***(24)</td>
</tr>
<tr>
<td>100</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

In general, the discrimination performance for subjects #1 and #3 are not different from each other. The performance for subject #2 is worse than that for the other two. This is consistent with their frequency DL measures (larger frequency DLs lead to poorer performance in transition slope discrimination). The rove-range is a very strong factor in the moving filter experiment. The multiple comparisons showed that the TSDT for all rove-ranges are significantly different from each other. The effect of signal duration is also significant. The TSDTs for all durations differ from each other; short signals lead to poorer discrimination. The data are plotted in a different way in Figure 19 to show the difference between
subjects for three durations. It can be seen that subject #2 has somewhat poorer performance and the results for the other two subjects are very close.

4.1.3.3 Comparison between the discrimination of pure tone glide and moving filter signal. In order to examine differences in discrimination between the tone-glide and moving-filter signals, an averaged TSDT across signal duration was obtained and shown in Figure 20 for three listeners individually. A three-factor ANOVA revealed that signal type is a significant effect \( F(1, 2) = 694.38, p = 0.0014 \). It is very clear in Figure 20, that the TSDTs for moving filter signals are higher than the TSDTs for pure tone glides. All three subjects show the same pattern of results. It seems that the difference between the two types of signal is large for narrow rove range and a little bit smaller at wide rove range, but that difference is not significant. On average, moving filter signal discrimination had a TSDT score about 76 Hz higher than the TSDT measured for tone glides.

TSDT values were averaged across subject to examine the difference between the two types of signals for different signal durations. The results are shown in Figure 21. The discrimination for moving filter signal was poorer for all signal durations. The difference for longer duration may be larger than the difference for shorter duration. But, again, it is not statistically significant.
Figure 19: Transition slope discrimination threshold for moving filter signals. Data are plotted separately for three signal durations.
Frequency slope discrimination for tone glides and moving filter signals

Figure 20: TSDT score averaged across signal durations for tone glides and moving filter signals. Data are plotted separately for three subjects.
Figure 21: TSDT score averaged across subjects for tone glides and moving filter signals. Data are plotted separately for three durations.
4.2 Performance of the Computational Model

The computational model implemented in this study was based on an independent-channels model in which the frequency DL and rove range were employed as parameters, and the IWAIF vectors served as model input. (See chapter 3) The model calculated a sensitivity index $d'$ to predict the discriminability between two stimuli. The predicted $d'$ is (from chapter 3):

$$d' = \frac{1}{\sigma} \left[ \sum_{j=1}^{N} \frac{\Delta_j}{\sigma^2} - \frac{\sigma_R^2}{\sigma^2 + N\sigma_R^2} \left( \sum_{j=1}^{N} \Delta_j \right)^2 \right]^{1/2} \quad (34)$$

where $\sigma_R$ is equal to zero if the case of fixed starting frequency is considered.

Predictions of the model are presented graphically so that we can compare the model's results with listener performance. The model was evaluated in three aspects. First, the model was examined mathematically. This means no real signal was used in calculating the model output. In this case, we assumed that the IWAIF vector contained frequency values which were located on the linear transition. Second, digital samples of the tone glides were used to compute the model output. Finally, the moving filter signals were used to calculate the model output. In addition, we demonstrate the use of Patterson and Holdsworth's (1990) ASP (auditory sensation processing) model as a preprocessor to the IWAIF calculation.
4.2.1 Mathematical Model

Performance

In this case, we want to examine the model's performance without IWAIF calculations. Thus, an "ideal" IWAIF vector was assigned to a specific signal. The "ideal" IWAIF vector was created to fit the linear transition exactly. For example, the IWAIF vector for a transition from 1000 to 1600 Hz would be designated as \[ [1030 \ 1090 \ 1150 \ 1210 \ 1270 \ 1330 \ 1390 \ 1450 \ 1510 \ 1570]^T \] rather than a computed IWAIF vector e.g., \[ [1035.13 \ 1083.56 \ 1131.64 \ 1194.36 \ 1256.07 \ 1318.80 \ 1381.31 \ 1443.57 \ 1506.07 \ 1561.42]^T \].

The reference signal was a rising tone glide from 1000 Hz to 1400 Hz. The \( d' \) was the sensitivity index between the reference signal and a target signal having a larger frequency excursion. Note that, in the prediction model, there is no performance difference between adapted-from-above and adapted-from-below, since the model only considered the distance between two signals. The \( d' \) calculated by using equation (34) with the parameter, \( \sigma \), evaluated from the frequency DL for listener #1 is shown in Figure 22. The difference in frequency excursion between target and reference is presented as "DF" in Figure 22. Thus, "DF=0 Hz" means that both the target and reference excursion are 400 Hz, while "DF=200 Hz" means the target excursion is 600 Hz and the reference excursion is 400 Hz. Figure 22 shows the \( d' \) for target excursion from 400 Hz up to 600 Hz. Rove range is
Figure 22: The $d'$ prediction for subject #1 with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

plotted on a logarithmic scale. Each curve represents the $d'$ for a single excursion difference under various rove-ranges.

The predicted $d'$ increases logarithmically with the increase in excursion difference. The $d'$ becomes smaller when the rove-range increases. This means that model performance becomes poorer as rove-range increases. Note that the results
obtained from the listening experiment also exhibited the same tendency. When examining the plot at the point where $d' = 1.16$ for zero rove-range (the fixed condition), we can see that the excursion difference is about 10 Hz. The mean of the excursion difference, which is located at the middle of the frequency transition, is about 5 Hz. If we think frequency DL is an important key to the frequency slope discrimination, a reasonable assumption is that the frequency DL should be equal to the mean of the just detectable excursion difference. From Figure 22, we find that for $d' = 1.16$, the frequency DL (4.71 Hz) is located at about the mean of the corresponding excursion difference (about 10 Hz). Therefore, from the above inspections, the model works appropriately. However, the data obtained from the listening task showed that the TSDTs for fixed conditions were usually more than double the values of frequency DL (see Table 3 and 4). Thus, while playing an important role, the frequency DL is not the only determinant in the frequency transition slope discrimination.

To show the model's prediction for different frequency DLs used as parameters, the prediction of $d'$ for subject #2 (frequency DL=8.09 Hz) and subject #3 (frequency DL=7.03 Hz) are presented in Figure 23 and Figure 24 respectively. From these figures, we can find that the predicted $d'$ for different frequency DLs have the same tendency (decreasing discriminability with an increase in rove-range). Further, the predicted performance for larger frequency DLs is poorer than
that for smaller frequency DLs. So the model prediction is consistent with the listener performance.

Figure 23: The $d'$ prediction for subject #2 with a frequency DL equals 8.09 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

The model's predictions were compared with the discrimination performance of the listeners. To do this, we first drew a horizontal line at $d'=1.16$ in the $d'$ versus
Figure 24: The $d'$ prediction for subject #3 with a frequency DL equals 7.03 Hz. Signals are 100 ms tone glides. Ideal IWAIF.

rove-range plot. We then find the "required excursion difference" (that is the TSDT) for which the model predicts 79.4% correct performance for different rove-ranges. Figure 25 shows the comparison between the performances (TSDT vs. rove-range) of the prediction model and listener #1. In Figure 25, the curves with open squares represent the model's
prediction calculated for different d’s at different rove-ranges. The filled circles represent the difference in frequency excursion at d’=1.16 for different rove-ranges exhibited by subject #1.

We can see that both the model and listener show poorer discrimination performance as rove-range is increased, but the model predicts much better performance than the listener obtains. The model predicts smaller TSDTs than the listener produces. The TSDT values obtained by the subject at 97.9% correct (d’=1.16) corresponds with the model prediction of about 91% correct (d’=1.9). The mathematical model uses the information contained in the signals optimally to calculate the predictions. If we assume that the human listeners could only use part of the signal information, the performance discrepancy between the model and listeners is understandable.

In order to make the comparison more quantitative, the "efficiency of the listener" (EL) designated by $d’/d’_m$ was obtained:

$$EL = \frac{d’}{d’_m} \times 100\%$$  \hspace{1cm} (35)

where $d’_m$ is obtained by converting the listener’s TSDT to the corresponding sensitivity index predicted by the model and $d’$ is equal to 1.16 for a 2AFC 3up-1down adaptive procedure employed in the listening task. IF EL is equal to 100%, it means the model and listener obtain the same $d’$. An EL smaller than 100% indicates that listener’s performance is poorer than
The model's prediction and the result of subject #1
Frequency DL=4.71 Hz at d'=1.16
Signal : 100 ms tone glides

Figure 25: The comparison of the discrimination (TSDT) between the prediction model and subject #1. Open square: model; Filled circle: subject. Ideal IWAIF.

the model's prediction. An EL larger than 100% means the model predicts poorer performance than the subject obtains.

Therefore, The EL served as an index which presented the relative discrimination efficiency of the subject when the experimental results were compared to the model prediction.
Since the computational model predicted a much better discrimination performance than the listener measured, a EL value smaller than 100% was expected.

The EL values calculated from the TSDTs for subject #1, for 100 ms tone glides, is shown in Table 8. The $d'_m$ for different rove ranges were obtained (from Figure 25). Then, the EL values are calculated by dividing 1.16 by those $d'_m$. The ELs for different rove ranges did not have the same values. There is a weak trend that, except the fixed frequency condition, ELs decrease with the increase in rove-range. This can be observed, from Figure 25, that the listener’s TSDT curve over rove-range is steeper than those TSDT curves with constant $d'$ predicted by the model. This indicates that the performance difference between the model and subject is smaller for narrow rove ranges and is larger for wide rove ranges.

Since under the fixed condition (without roving frequency), the listeners can use the offset pitch cue to do the discrimination task, the fixed condition is treated as a special case in our experiment. Particularly, the suprising high efficience (98.64%, see Table 8) at fixed condition for subject #3 shows the pitch cue effect clearly. Thus, the averaged EL values shown in Table 8 do not include the data from the fixed condition. The averaged EL values were obtained by averaging the ELs of the roving frequency conditions. From Table 8, the averaged EL value for subject #1 is about
Thus, the model's prediction is better than the performance of subject #1.

Comparisons between the model and listeners #2 and #3 are plotted in Figure 26 and Figure 27 respectively. From these figures, we can find that for all the subjects, the discrimination predicted by the model is much better than the listener's performance. For subject #2, the obtained excursion differences would have resulted in a $d' = 1.75$ as predicted by the model. For subject #3, the experimental TSDT curve occupied a wider range of model predictions. At the fixed condition, this listener's performance (located at $d' = 1.2$) is close to the model's prediction. When the rove-range equals 800 Hz the listener's TSDT is close to a predicted $d' = 1.8$. Furthermore, it seems that the difference between the model prediction and subject performance is smaller for narrow rove range. The difference becomes larger when rove range increases. This can also be seen in Table 8 where, in general, the values of $d'_m$ increase slightly with the increase in rove range.

The EL values calculated for subjects #2 and #3 are also listed in Table 8. From the table, The model has the closest prediction for subject #3, because subject #3 has the highest efficiency (69.88%). Due to the differences in listening ability among subjects, the EL values for the three subjects differ from each other. However, all the EL values are less than 100%, and thus we can say that, in general, the model's
prediction is better than the subject's performance. Since the model is designed to predict human performance, a perfect model would predict human performance as close as possible. When we say "the model's prediction is better than subject's performance", we do not mean the model is a better model; we just mean that the efficiency of the listener is less than 100 percent. A model is perfect if EL value is 100 percent. The average efficiency of discrimination in frequency transition slope across subjects is 66.04% for the roving frequency condition and is 75.98% for the fixed condition.

In order to examine the model prediction for different signal durations, the predicted d's (d' vs. rove-range) for 50 and 25 ms signals are shown in Figure 28 and Figure 29 respectively. The model's prediction of performance (TSDT vs. rove-range) generated from the d' plot for each duration is presented in Figure 30 and Figure 31 respectively. Note that only the data from subject #1 are shown here. The other two subjects produced similar results.

Short signal durations, of course, produced poorer discrimination predictions (see Figure 28 and Figure 29). However, since the value of the frequency DL also changed with the signal duration, it was not easy to evaluate the effect of the duration on the model's prediction. The signal duration did show a weak effect on the model prediction, when we tried to keep the frequency DLs fixed for different duration...
Table 8: Listener’s EL values for the three subjects. (100 ms tone glide, ideal IWAIF)

<table>
<thead>
<tr>
<th>subject #1</th>
<th>EL value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rove range (Hz)</td>
<td>d'_m</td>
</tr>
<tr>
<td>0</td>
<td>1.916</td>
</tr>
<tr>
<td>100</td>
<td>1.776</td>
</tr>
<tr>
<td>200</td>
<td>1.855</td>
</tr>
<tr>
<td>400</td>
<td>1.939</td>
</tr>
<tr>
<td>800</td>
<td>1.957</td>
</tr>
<tr>
<td><strong>Average</strong> (without 0 rove range)</td>
<td>1.882</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #2</th>
<th>EL value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rove range (Hz)</td>
<td>d'_m</td>
</tr>
<tr>
<td>0</td>
<td>1.687</td>
</tr>
<tr>
<td>100</td>
<td>1.693</td>
</tr>
<tr>
<td>200</td>
<td>1.698</td>
</tr>
<tr>
<td>400</td>
<td>1.784</td>
</tr>
<tr>
<td>800</td>
<td>1.808</td>
</tr>
<tr>
<td><strong>Average</strong> (without 0 rove range)</td>
<td>1.746</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #3</th>
<th>EL value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rove range (Hz)</td>
<td>d'_m</td>
</tr>
<tr>
<td>0</td>
<td>1.176</td>
</tr>
<tr>
<td>100</td>
<td>1.512</td>
</tr>
<tr>
<td>200</td>
<td>1.629</td>
</tr>
<tr>
<td>400</td>
<td>1.742</td>
</tr>
<tr>
<td>800</td>
<td>1.785</td>
</tr>
<tr>
<td><strong>Average</strong> (without 0 rove range)</td>
<td>1.667</td>
</tr>
</tbody>
</table>
Model's prediction and the result of subject #2

Frequency DL=8.09 Hz at d'=1.16
Signal : 100 ms tone glides

![Graph showing comparison between model prediction and subject #2's result.]

**Figure 26:** The comparison of the discrimination (TSDT) between the prediction model and subject #2. Open square: model; Filled circle: subject. Ideal IWAIF.

...conditions. But, since the subjects exhibited different frequency DLs for different signal durations, the model were computed using the frequency DLs measured from the listening task. The point we address here is that the model's predictions for short signal durations did not show any...
Figure 27: The comparison of the discrimination (TSDT) between the prediction model and subject #3. Open square: model; Filled circle: subject. Ideal IWAIF.

difference in tendency from the results for 100 ms described earlier. The model's prediction for short signal duration was highly depended on frequency DL (small frequency DL led to better performance) and rove range (performance was poorer for wide rove range).
Model's prediction of d' frequency DL = 5.81 Hz at d' = 1.16 subject #1
Signal: 50 ms tone glide

Figure 28: The d' prediction for subject #1 with a frequency DL equals 5.81 Hz. Signals are 50 ms tone glides. Ideal IWAIF.
Model's prediction of d'
frequency DL = 7.64 Hz at d'=1.16 subject #1
Signal: 25 ms tone glide

Figure 29: The d' prediction for subject #1 with a frequency DL equals 7.64 Hz. Signals are 25 ms tone glides. Ideal IWAIF.
Model's prediction and the result of subject #1
Frequency DL=5.81 Hz at d' = 1.16
Signal: 50 ms tone glide

Figure 30: The comparison of the discrimination (TSDT) between the prediction model and subject #1 for 50 ms signals. Ideal IWAIF.
Model's prediction and the result of subject #1
Frequency DL=7.64 Hz at d' = 1.16
Signal: 25 ms tone glide

Figure 31: The comparison of the discrimination (TSDT) between the prediction model and subject #1 for 25 ms signals. Ideal IWAIF.
4.2.2 Model Performance for Tone Glides

In this section, digital samples of the tone glides were sent to the IWAIF calculation. Then, the IWAIF vectors for the signals were used as the input of the prediction model. Since, in general, the IWAIF vector for a given glide signal will approximate the ideal linear frequency trajectory of the signal, the model's prediction should be similar to that described in the previous section in which the model was evaluated mathematically. Because the variance was increased by using "real" signals, the performance predicted by the model was expected to be poorer than that obtained for the mathematical evaluation in the previous section.

The prediction of $d'$ ($d'$ versus rove-range) for subject #1, #2 and #3 are plotted in Figure 32, Figure 33, and Figure 34 respectively. The predicted $d'$ increases logarithmically with the increase in frequency excursion difference. Since the use of real IWAIF calculation introduced some deviation in transition of frequency, the $d'$ does not increase very smoothly. When comparing model prediction between this section and the previous section, we found that the $d'$ obtained from using real IWAIF calculation was slightly smaller than the $d'$ calculated ideally (under the assumption that IWAIF vector fitted the transition exactly).

The predicted performance of the model was constructed in the same way described in section 4.2.1. The comparison of
Model's prediction of $d'$ using IW AIF
Frequency DL=4.71 Hz at $d'=1.16$ (79.4% correct)
Signal: 100 ms tone glide subject: #1

![Graph showing the $d'$ prediction for subject #1 with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides.]

**Figure 32:** The $d'$ prediction for subject #1 with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides.
Figure 33: The d' prediction for subject #2 with a frequency DL equals 8.09 Hz. Signals are 100 ms tone glides.
Figure 34: The $d'$ prediction for subject #3 with a frequency DL equals 7.03 Hz. Signals are 100 ms tone glides.
discrimination performance (TSDT vs. rove-range) between the model and listeners are shown in Figure 35, Figure 36, and Figure 37 for subjects #1, #2 and #3 respectively. Due to the variance introduced by using real signals, the curves of model prediction do not change smoothly with rove-range and show some extent of deviations. In general, the performance predicted by the model in this case is slightly poorer than the ideal model performance (in section 4.2.1). The TSDT predicted by the model increased with the increase in rove range. This is consistent with the data from the listening experiment. However, the results still showed that the model predicted much better performance than the subjects produced.

Again, the relative efficiency of discrimination for each subject is listed in Table 9. The efficiency of listener (EL) is calculated in the same way defined in section 4.2.1. The trend observed in section 4.2.1 that the EL value decreased slightly with the increase in rove-range is not clear for subject #1 and #2. While comparing the EL values obtained here to the EL values obtained by using ideal IWAIF (comparing Table 9 and Table 8), we can find that for subject #1 and #2, the efficiency of listener increases about 2 percent; for subject #3, the EL values are very close to each other. On average, the efficiency of the listener is 67.24% for roving frequency conditions and is 75.35% for fixed condition. These numbers indicate that the model predicts a much better performance than the subject obtains.
Model's prediction and the result of subject #1
frequency DL=4.71 Hz at d'=1.16
Signal : 100 ms tone glide

Figure 35: The comparison of the discrimination (TSDT) between the prediction model and subject #1. Open square: model; Filled circle: subject.
Model's prediction and the result of subject #2
frequency DL=8.09 Hz at $d'=1.16$
Signal: 100 ms tone glide

Figure 36: The comparison of the discrimination (TSDT) between the prediction model and subject #2. Open square: model; Filled circle: subject.
Model's prediction and the result of subject #3
frequency DL=7.03 Hz at d'=1.16
Signal : 100 ms tone glide

Figure 37: The comparison of the discrimination (TSDT) between the prediction model and subject #3. Open square: model; Filled circle: subject.
Table 9: Listener's EL values for the three subjects. (100 ms tone glide)

<table>
<thead>
<tr>
<th>subject #1</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1.911</td>
<td>60.70</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.758</td>
<td>65.98</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.825</td>
<td>63.56</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>1.885</td>
<td>61.54</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>1.883</td>
<td>61.60</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>(without 0 rove range)</td>
<td>1.838</td>
<td>63.17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #2</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1.683</td>
<td>68.92</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.644</td>
<td>70.56</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.606</td>
<td>72.23</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>1.764</td>
<td>65.76</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>1.747</td>
<td>66.40</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>(without 0 rove range)</td>
<td>1.690</td>
<td>68.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #3</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1.203</td>
<td>96.43</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.532</td>
<td>75.72</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.625</td>
<td>71.38</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>1.751</td>
<td>66.25</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>1.761</td>
<td>65.87</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>(without 0 rove range)</td>
<td>1.667</td>
<td>69.81</td>
</tr>
</tbody>
</table>
4.2.3 Model Performance for Moving Filters

In this section, the IWAIF vector of the moving filter signals were calculated and sent to the prediction model. Because the use of the moving filter signals introduced a larger variance in the results of IWAIF calculation, the performance predicted by the model was expected to be much poorer than that obtained for the linear tone glides in the previous section. For moving filter signals, it is useless to generate a $d'$ versus rove-range plot for various excursion differences unless the average of a single signal is computed. Due to the difficulty of generating an averaged $d'$ plot for different frequency excursions, we only calculated the model's predictions at the subjects' TSDTs obtained from the listening task. Then the efficiencies of the listeners were computed.

In order to calculate the model prediction, the frequency difference limen (frequency DL) for 100 ms moving filter signals centered at 1000 Hz were measured. The frequency DLs obtained in this case were larger than that measured for tone glides. The frequency DL for moving filter signal was 9.23, 15.03, and 13.8 Hz for subject #1, #2, and #3 respectively.

The relative efficiency of discrimination for each subject is listed in Table 10. The efficiency of listener (EL) is calculated in the same way defined in section 4.2.1. In general, the model still predicted much better performance than the listeners produced. There is not a clear tendency
that Table 10: Listener's EL values for the three subjects. (100 ms moving filter signal)

<table>
<thead>
<tr>
<th>subject #1</th>
<th>EL value</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>2.153</td>
<td>53.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>1.766</td>
<td>65.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>1.796</td>
<td>64.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>1.758</td>
<td>65.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>800</td>
<td>1.719</td>
<td>67.48</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>without 0</strong></td>
<td></td>
<td>1.760</td>
<td>65.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>rove range</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #2</th>
<th>EL value</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1.841</td>
<td>63.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>1.719</td>
<td>67.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>1.670</td>
<td>69.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>1.570</td>
<td>73.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>800</td>
<td>1.553</td>
<td>74.69</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>without 0</strong></td>
<td></td>
<td>1.628</td>
<td>71.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>rove range</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subject #3</th>
<th>EL value</th>
<th>rove range (Hz)</th>
<th>$d'_m$</th>
<th>$1.16/d'_m \times 100$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1.844</td>
<td>62.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>1.652</td>
<td>70.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>1.668</td>
<td>69.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>1.596</td>
<td>72.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>800</td>
<td>1.605</td>
<td>72.27</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>without 0</strong></td>
<td></td>
<td>1.630</td>
<td>71.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>rove range</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

the listener's efficiency increases or decreases with the change in rove-range. The results of the listening test showed
that the discrimination performance for moving filter signal was much poorer than that for pure tone glide (see section 4.1.3.3). While comparing the EL values for moving filter signal to the EL values for tone glides (comparing Table 10 and Table 9), we found that all subjects exhibited slightly higher efficiencies for moving filter signals under roving frequency conditions, but lower efficiencies under fixed condition. The averaged (across three subjects) EL value for roving frequency condition is 69.5% and is 59.93% for fixed condition. Under the roving frequency condition, the difference of EL between these two types of signals was 2.26%. Since the difference was so small, it indicated that the model also predicted poorer discrimination performance for moving filter signals. That is, the model also predicted a positive TSDT shift from tone glides (low TSDT) to moving filter signals (high TSDT). The amount of the TSDT shift due to the different stimuli was not investigated for the model. However, since the EL difference between the tone glides and the moving filter signals was small, the TSDT shift for the model should be comparable to the TSDT difference obtained from the listeners.

4.2.4 Using the ASP Model as a Preprocessor

In this section, we will demonstrate the use of the ASP model as a preprocessor to the IWAIF calculation. The stimuli
examined in this study were signals with linear frequency transitions. Thus, both short-term FFT analysis and the ASP's excitation pattern generated appropriate representations in frequency domain for the glide signals. Given a tone glide, the results of IWAIF calculations for both the neural excitation pattern and the short term FFT were expected to be very similar.

An example of d' predicted by the model using IWAIF vectors calculated from the excitation pattern is shown in Figure 38. The signals were 100 ms tone glides. Frequency DL was selected to be 4.71 Hz, which is obtained from subject #1. Comparing this figure to Figure 32, we can see that the model's prediction became slightly poorer if the excitation pattern was used. In other word, the model's prediction became closer to the listener's performance. One should note that when we say "the model's prediction became poorer if the excitation pattern was used", we did not mean this is a poorer model. Actually, a model using excitation pattern is a better model because its prediction becomes closer to human data. It seems that the use of excitation pattern is more approximate to the cochlear processing. The improvement is small but it is positive.

The performance (TSDT vs. rove-range) predicted by the model using excitation pattern is shown in Figure 39. This figure was generated from Figure 38. The comparison between Figure 39 and Figure 35 shows that the use of excitation
pattern leads to a prediction which is slightly closer to subject's performance. The EL values were calculated and listed in Table 11. The data shows the trend that the EL values decrease with the increase in rove-range. The average of the efficiency of listener is 66.52% for roving frequency conditions and the efficiency for fixed condition is 83.63%. Comparing these numbers respectively to 63.17% and 60.70%, which were shown in Table 9 for the model using FFT analysis, we found that higher efficiency occurred for the use of excitation pattern. This indicates that the use of the ASP model generates a prediction closer to the listener performance. Note that the efficiency discussed here is the efficiency for listener not for the computational model.

Table 11: Listener's EL values for the model using EPN data. (100 ms tone glide)

<table>
<thead>
<tr>
<th>subject #1</th>
<th>100 ms tone glide</th>
<th>EL value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d'_m$</td>
<td>$1.16/d'_m \times 100 %$</td>
</tr>
<tr>
<td>0</td>
<td>1.387</td>
<td>83.63</td>
</tr>
<tr>
<td>100</td>
<td>1.651</td>
<td>70.26</td>
</tr>
<tr>
<td>200</td>
<td>1.719</td>
<td>67.48</td>
</tr>
<tr>
<td>400</td>
<td>1.803</td>
<td>64.34</td>
</tr>
<tr>
<td>800</td>
<td>1.813</td>
<td>63.98</td>
</tr>
<tr>
<td><strong>Average</strong> (without 0 rove range)</td>
<td>1.747</td>
<td>66.52</td>
</tr>
</tbody>
</table>
Figure 38: The $d'$ prediction for subject #1 with a frequency DL equals 4.71 Hz. Signals are 100 ms tone glides. EPN data are used in the model.
Excursion Difference (Hz)

Model's prediction and the result of subject #1
frequency DL=4.71 Hz at d'=1.16
Signal : 100 ms tone glide (using EPN)

Figure 39: The comparison of the discrimination (TSDT) between the prediction model and subject #1. EPN data is used in the model.
The main purpose of this dissertation was to investigate the discrimination of frequency transition slopes and to devise a computational model to predict the performance of listeners. A roving frequency paradigm was employed to reduce the pitch artifacts and make the task more like that required of the auditory system in the perception of speech. In this chapter, we give a general discussion of the results we obtained and relate them to similar work in the literature. Some possible considerations and modifications to the model will be suggested so that the model can better predict human transition slope discrimination.

5.1 Frequency Discrimination for Steady-state Tones and Transition Stimuli

Elliott et al. (1989) reported that the just noticeable difference (JND) for the longest (120 ms) transitions, measured in Hz at transition onsets (without roving starting frequency), were of approximately the same magnitude as JNDs for steady-state sounds having frequencies equal to midpoints of the transitions. Some other studies also indicated that at
long durations (approximately 300 ms or longer), discrimination among transitions is similar to that for comparable steady-state signals (Shower and Biddolph, 1931; Tsumura et al., 1973; Hartmann and Klein, 1980). However, Tyler et al. (1983), using 50 ms stimuli, reported a smaller frequency threshold (14.8 Hz) for discriminating steady-state tones and larger threshold for discriminating transitions (59.6 Hz). Our data, for the fixed starting frequency condition, is consistent with Tyler's results. Dooley and Moore (1988) used 500 ms stimuli to measure the frequency threshold for discriminating between steady-state and transition tones. They reported that the UP-DLs (frequency threshold for steady-state tone vs. rising tones) and DOWN-DLs (frequency threshold for steady-state tone vs. falling tone) are higher than the DLFs (frequency threshold for discriminating steady-tone), on average, by a factor of 2.3. Our experimental data show that, for the fixed starting frequency condition, the TSDTs for transition slope discrimination are much higher than the simple frequency DLs. The point we are going to address here is that there is a disagreement in the relationship between frequency DLs for steady-tone and transition discrimination threshold measures. Some researcher indicated that the magnitudes of these two measures are similar (Elliott et al., 1989; Shower and Biddolph, 1931; Tsumura et al., 1973; Hartmann and Klein, 1980), while some others reported the thresholds for
transition discrimination are much higher than the frequency DLs (Tyler et al., 1983; Dooley and Moore, 1988; present study). The duration of signals and their center frequency also have to be considered as factors; long and short signals will not provide the same results, and threshold measured at different center frequencies will be different.

In our computational model we consider the frequency DL as an important key to frequency transition slope discrimination. Further, we assume that the frequency DL should be equal to the mean of the just detectable transition slope difference. However, the data obtained from the listening task showed that the TSDTs for fixed starting frequency conditions were usually more than double that of the frequency DL (see Table 3 and 4). Thus, although the frequency DL plays an important role in the discrimination of transition slopes, it is not the only determinant for the discrimination. We will discuss the use of an alternative measure rather than the frequency DL as an parameter of the computational model in section 5.6.

Elliott et al. (1991) examined the discrimination of second-formant-like frequency transitions. Our TSDT data at fixed starting frequency conditions are generally consistent with this earlier study. Discrimination is better for the signals with longer durations. The TSDTs obtained in the present study for 100 ms signals are slightly smaller than the JNDs measured at 71% correct performance for 120 ms in
Elliott's report. Some possible reasons for this difference include (1) the reference signal used in Elliott's experiment had very low transition rates (0.2 Hz/ms) compared with the higher transition rate (4 Hz/ms) used in the present study; (2) The stimuli used in the two experiments differed; Elliott used signals synthesized by Klatt procedure (Klatt, 1987) and we used FM sinusoids; (3) In Elliott's experiment 2, a three-interval, forced-choice (3IFC) non-adaptive procedure was used, while in our experiment, a 2Q, 2AFC 3up-1down adaptive procedure was employed; (4) Elliott's data were extracted from the psychometric functions, our data were obtained from the adaptive procedure.

5.2 Effect of Roving Starting Frequency
The results of this investigation show that the introduction of a roving starting frequency significantly influences subject performance on the transition slope discrimination task. In general, using a random starting frequency leads to a decrease in discriminability for all subjects. Transition slope discrimination performance is further degraded as the size of roving frequency range is increased. Thus TSDTs increase with an increase in the rove-range for all subjects. The same tendency is also exhibited when the moving filter signals is used. Neill (1990) found that the curves of psychometric functions remained parallel even the starting frequency range was changed. She
indicated that the subjects did not change their criterion for discrimination of transition direction with either the introduction of, or an increase in size of, the random starting frequency range. The 2Q, 2AFC procedure used in our experiment eliminate the effect of changes in criterion or motivation which can make performance changes. Thus, we can assume that the subjects used the same criterion for both the fixed and roving frequency transition slope discrimination task. When the roving frequency procedure is used, the decrease in listener sensitivity to slope change indicates that listeners are unable to do the discrimination as well as the fixed condition. The deterioration in performance, when random starting frequencies are used, provides support for the assumption that listeners use endpoint pitch cues in "fixed" frequency conditions to help them in the discrimination of slope. When the starting frequencies are randomized, the pitch cue is unusable.

Our finding of better performance for fixed versus roving starting frequency is generally in agreement with the results of some earlier studies (Pollack, 1968; Henning, 1965). Pollack showed that an increase in glide direction discrimination threshold was observed with randomization of starting frequency. Pollack reported 1.3-fold and 1.8-fold increases in the threshold extent-of-frequency change for two different subjects when a random starting frequency was used. The results obtained in this study can not be directly
compared with Pollack's data due to differences in experimental designs.

Nabelek et al. (1970) studied the pitch of tone bursts containing linear frequency changes. They stated that (1) if frequency change and duration were not too large, the bursts were matched in pitch to frequencies near the middle of the burst; (2) for large frequency changes, bursts were matched by either initial or final frequency. This result supports the idea that subjects can use pitch cues at either the initial or the final point of the transitions.

Henning (1965) also reported that if the random amplitude paradigm was employed, the frequency dependent loudness cues could not be used by the subjects to make listening task decisions. Henning's results suggested that the JNDs obtained by using roving amplitude paradigm more accurately reflect the frequency resolution of the auditory system. The roving starting frequency procedure employed in the present study appears to control the pitch artifact interference successfully for the task of transition slope discrimination.

The performance change with the introduction of roving starting frequency is easy to explain using the computational model. The d' predicted by the model is represented by
where $\sigma_r^2$ is the variance of randomized starting frequency. It is always positive and proportional to the size of the roving frequency range. It is clear that if a roving frequency range is introduced ($\sigma_r^2 \neq 0$), the second term of the above equation will not be zero, and the resulting $d'$ will be smaller than the $d'$ predicted for the fixed starting frequency condition, (in which $\sigma_r^2 = 0$). This means introducing a roving frequency range will decrease the $d'$ and thus degrades discrimination performance. Also $d'$ will decrease with an increase in the size of roving frequency range because increasing rove-range will increase $\sigma_r^2$. Given other conditions unchanged, the equation shows that increasing the rove-range decreases $d'$. The value of $d'$ will asymptotically approach a minimum boundary for an infinitely large rove-range size. This suggests that there is a "saturation" in the performance of transition slope discrimination. That is, the performance can not continue to fall when some large roving frequency range is reached. In the present experiment, some data for the condition of adapted-from-below showed saturations in discrimination performance, but the averaged data did not display the "saturation" phenomena very clearly. The size of the roving frequency range used in our experiment may not be large enough to exhibit this aspect.
5.3 Duration Effect for Transition Discrimination

Our finding that increased duration yields improved frequency resolution for moving filter signals (i.e., better discrimination performance) is consistent with some other reports (Nabelek and Hirsh, 1969; Nabelek et al., 1970; Elliott et al., 1989, 1991; Porter et al., 1991). This phenomena may be related to the "sampling theory" of transition discrimination defined by Dooley and Moore (1988). Dooley and Moore proposed that transition discrimination was based on the perception of differences in frequency between samples of the stimuli taken at different points in time. According to their "sampling theory", under the same transition rate, the TSDTs should become smaller with increased stimulus duration because the longer stimuli sweep through larger frequency range and would allow opportunities for taking more samples. However, in our experiment, the transition rate changed with the change in signal duration. The transition rate for 100 ms glide was 4 Hz/ms, while the transition rate for 25 ms glide was 16 Hz/ms. All the signals sweep through the same frequency range. Cullen (Cullen et al., 1992), using short stimuli (40 ms), showed that discrimination was poorer for faster transitions (using stimuli with longer duration may produce opposite results). Our data showed that the duration effect was weak for the discrimination of tone-glides, but for the moving filter signals, signal
duration was a significant factor (see Table 7). Thus, although the rate-of-transition may be a factor in slope discrimination, the "sampling theory" can still be used to explain the duration effects in our experiment results. As the signal duration increases, the listeners get more chances to sample along the signal and thus, they can make a more accurate estimation of frequencies for transitions.

5.4 The Asymmetric Discrimination for Different Adaptive Directions

The TSDTs obtained in the present experiment showed that the discrimination was performed better for the adapted-from-above condition. Adapted-from-below procedures always led to higher TSDT values for all three subjects and for all of the signal conditions (durations and rove-range) and types. The reason for this asymmetric performance is not obvious. In a study of discrimination in formant transition onset frequency, Porter et al. (1991) also report a kind of performance asymmetry for different adaptive directions. Porter et al. proposed that this asymmetry could be explained by a "rate cue" which is based on the excitation-envelope mechanism. They stated that "the phenomenon can be explained on the basis of the observation that the growth in perceived envelope magnitude with increases in the physical change in the signal (Terhardt, 1968; Vogel, 1975); that is, given a particular reference variation, a given physical increase in
magnitude of the variation will produce a larger change in the resulting excitation envelope than an equivalent physical decrease." Thus, increments will produce smaller JNDs than decrements. This explanation can be applied to our experiment in which the transition slope was increased or decreased.

In our experiment, the adapted-from-above procedure gradually decreases the transition slope and leads to a reversal point where the subject can not distinguish the difference between two signals. Then, the slope of the target transition was increased to make a large frequency excursion difference, and the increase in target slope produced a change in the excitation pattern (envelope). The adapted-from-below has a reverse situation, in which the target slope was decreased at the reversal point to make a larger difference between two signals. According to the excitation-envelope mechanism proposed by Porter et al., the change in the excitation envelope caused by the adapted-from-above procedure will be greater than the excitation envelope change caused by the adapted-from-below procedure. Thus, the adapted-from-above procedure resulted in smaller TSDT values.

It will be easier to understand the above argument if we explain by using a figure. Three excitation patterns corresponding to the offsets of three signals are shown in Figure 40. The curve on the left represents a signal offset for the adapted-from-below procedure. The middle curve is the offset excitation pattern for the reference signal. The curve
Offset Neural Excitation Patterns for different adaptive directions

Figure 40: Neural excitation patterns corresponding to the signal offsets for different adaptive directions and the reference.

on the right illustrates a signal offset for the adapted-from-above procedure. It can be seen that, in the adapted-from-above procedure, the signal sweeps through wider frequency range than the frequency range swept by the signal for adapted-from-below procedure. Therefore the changes in the excitation pattern introduced by the adapted-from-above procedure are greater than that introduced by the adapted-from-below procedure. Note that this explanation can only be used for the fixed frequency condition. The performance asymmetry also existed for roving frequency condition, and the reason is not clear.
5.5 Discrimination Difference between Tone-glides and Moving Filter Signals

The data of our experiment show that the TSDTs obtained for moving filter signals are significantly higher (by a factor of 1.55, or about 76 Hz = 215-139 Hz) than the TSDTs measured for pure tone-glides. This result is independent of signal duration, rove-range, and listener. Our computational model also predicts poorer discrimination performance for moving filter signals. Since the moving filter introduces a larger variance in the signals and the calculation of IWAIF, the poorer performance predicted for moving filter signals is not surprising. It can also be understood that since the moving filter signals have wider bandwidth and produce larger variance in the auditory system, the poorer discrimination is expected for the listener.

Although the model performs transition discrimination much better than the human listeners do, the model predicts similar results for the listeners on moving filter signals and tone-glides. For the moving filter signals, the averaged efficiency of listener (EL) relative to the model is 69.50% (not include 0 rove range). For the tone-glides, the EL value (not include 0 rove range) is 67.24%. The difference between the prediction of these two types of signal is small (only 2.26 percent). This result implies that the listeners may use the same psychoacoustical mechanism to discriminate frequency transitions for both tone-glides and moving filter signals.
The sampling theory (Dooley and Moore, 1988) can be applied on both moving filter signals and tone-glides for the duration effect on discrimination performance. That is, better discrimination performance occurred for longer signals.

Gagne and Zurek (1988) used glide sounds generated by a second-order filter to examine the just-noticeable change in resonance frequency. Their data showed that the just-noticeable change in resonance frequency increased with the increases in center frequency and filter bandwidth. They indicated that the results could be explained by Weber's law and the concept of critical band that the auditory analysis bandwidths increase with center frequency. Our experiment results at the fixed starting frequency condition are consistent with Gagne and Zurek's data. According to their explanation, since the moving filter introduces a wider bandwidth than tone-glide does, the poorer discrimination performance for moving filter signal is expected.

Another possible explanation for the poorer discrimination performance for moving filter signals is derived in the time domain. The zero crossing rate for the tone glide is about fixed, while the zero crossing rate for the moving filter signal is irregular. It can be explained that distinguishing between two signals with fixed zero crossing rate is easier for listener and becomes harder if the signals contain irregular zero crossing rates.
5.6 Possible Modifications for the Prediction Model

In general, the performance of the computational model is consistent with the subjects data. However, the model predicts much better performance than the subjects obtained. In order to make the present model more close to human listener performance, several suggestions are discussed in this section.

In the present study, we use the "short-term", running IWAIF calculation as the input of the decision model. It is based on the assumption that the listener samples the frequency along the duration of the stimulus, and use the sampled frequency vector to do the discrimination task. Long-term IWAIF is a consideration to be used as an alternative to short-term IWAIF. That is only one IWAIF value is obtained for a stimulus. Long-term IWAIF can be use to test if the listener make the decision based on integrated information of the whole transition signal rather than the temporal information along the transition duration. Note that the single-channel model proposed by Gagne and Zurek (1988) and Florentine and Buus is not applicable to our study because we employ a roving starting frequency procedure and the single-channel model is useless under this condition. The reason is that, in our case, the maximum difference between two signals usually occurs at the offset of the signals and roving frequency paradigm makes the offset pitch cue unusable.
In Florentine’s study (Florentine and Buus, 1981), they used a sensitivity fitting parameter \( k \) in the calculation of \( d' \). The model used by Gagne and Zurek (1988) did not contain the variance of the internal noise and they stated that the variance of the internal noise provided a fitting parameter to the experimental data. If a fitting parameter is added into the prediction model, the listeners’ performance can be fitted to the model’s prediction satisfactorily. We can assume the \( d' \) predicted by the model is proportional to the ideal sensitivity index \( d' \) obtained with the listener. Since we already have the averaged value of listener’s efficiency. We can predict the performance for a specified subject very accurately. For example, for subject #1, we know the \( d'_m \) is about 1.8 (see chapter 4, Table 9), then we can calculate the TSDT at \( d' = 1.8 \) using the model, and the TSDT calculated will be a good prediction for subject #1. The reason is that we already know the model’s prediction is, on average, comparable to the subject’s (#1) performance at \( d' = 1.8 \).

In the present study the IWAIF vectors for two different stimuli are sent to the computational model. We can also use different measures, which are still based on IWAIF, as the input of the model. One example is to use the difference of consecutive IWAIF values rather than the IWAIF values per se. Since our purpose is to discriminate two signals with different frequency transition slopes, the use of difference between two consecutive IWAIF for the discrimination task is
an appropriate assumption. Suppose the number of channels is $N$, the modified equation used to predict $d'$ is shown as follows:

$$d' = \frac{1}{\sigma} \left[ \sum_{j=1}^{N-1} \Delta_j^2 - \frac{\sigma_R^2}{\sigma^2 + (N-1) \sigma_R^2} \left( \sum_{j=1}^{N-1} \Delta_j \right)^2 \right]^{1/2}$$  \hspace{1cm} (37)$$

and

$$\Delta_j = \log_2(|IWAIF_{dj}-IWAIF_{dj}|+1)$$

$$\sigma = \frac{\Delta \sqrt{N-1}}{(N-1)d''}, \Delta = \log_2(TSDT_f+1)$$

$$\sigma_R = \frac{\log_2(R_f+1)}{\sqrt{12}}, \text{ } R_f = \text{roving frequency range}.$$  \hspace{1cm} (38)

where $IWAIF_{dj}$ is difference between the IWAIF in $j+1^{th}$ channel and the IWAIF in $j^{th}$ channel for signal $i$, and $TSDT_f$ is the threshold measured at fixed starting frequency condition. The idea is that the data obtained for the fixed starting frequency condition is used as a parameter of the model. Using this calculation, we may model some results of the 100 ms tone-glide discrimination for subject #1. The TSDT measured for the fixed starting frequency condition for subject #1 is 39.4 Hz. Figure 41 shows the predicted $d'$ from zero excursion difference to 800 Hz excursion difference using the ideal IWAIF values. The shape of the $d'$ curves in Figure 41 are similar to that shown in chapter 4 (see Figure 22). But, the $d'$ calculated by this new idea is smaller than that predicted before (see Figure 22). The model’s performance (TSDT vs.
rove-range) is plotted in Figure 42 with the data obtained by subject #1. We can see that the TSDT values obtained by the subject at 79.4% correct (d' = 1.16) corresponds with the model prediction for d' = 1.2. This results show that the prediction calculated by the model with the new calculation method can better fit listener performance. The predicted d’ for real 100 ms tone-glides are plotted in Figure 43 and the predicted TSDT values are shown in Figure 44. From Figure 44, we can see that the model’s prediction is very close to the performance of the subject. It seems that the use of difference of IWAIF as the input of the model is a good choice. However, before we can draw any conclusions, further studies have to be done.

In order to reduce the performance of the computational model, we can consider a more complicated concept for the interchannel correlation. One possibility is that interchannel correlation is not constant (the present model has a constant interchannel correlation named R, see section 3.3.2.). That is the perceived information at a specified time may affect the perception in previous time or the perception after that time. This consideration may reflect the concept of forward masking and backward masking, although we don’t know how much influence is produced by the masking in the discrimination of frequency transition slopes.
Model's prediction of $d'$ using slope difference

$DF=39.4$ Hz at fixed starting frequency

Signal: 100 ms, Subject: #1

Figure 41: The $d'$ prediction for subject #1 with TSDT equals 39.4 Hz at fixed condition. Signals are 100 ms tone glides. Ideal IWAIF.
Model's prediction and the result of subject #1
DF=39.4 Hz at fixed starting frequency
Signal: 100 ms tone glides (using slope difference)

Figure 42: The comparison of the discrimination (TSDT) between the prediction model and subject #1. Filled square: model; Filled circle: subject. Ideal IWAIF.
Model's prediction of $d'$ using IWAIF and slope
DF=39.4 Hz at fixed starting frequency for $d'=1.16$
Signal: 100 ms tone glides subject: #1

Figure 43: The $d'$ prediction for subject #1 with TSDT equals 39.4 Hz at fixed condition. Signals are 100 ms tone glides.
Figure 44: The comparison of the discrimination (TSDT) between the prediction model and subject #1. Filled square: model; Filled circle: subject.
5.7 Future Studies

5.7.1 Using Different Stimulus

In our experiment, pure tone-glides and moving filter signals generated by filtering white noise with a moving bandpass filter were used to test the discrimination of frequency transition slopes. It is interesting to examine the discriminability for some different stimuli. For example, the stimuli used in Gagne and Zurek’s experiment (1988) were generated by exciting a second-order filter with a sawtooth signal. The line spectrum of this sawtooth signal fell at a rate of about 6 dB/oct, similar to the spectrum of the glottal source wave. This type of source signal can be used on our moving filter structure to generate stimulus more like real speech. Some other periodic signals which are more close to glottal pulse may also be considered as the source waves of the moving filters.

The signals used in our experiment contained linear frequency changes. It is also interesting to use signals contain non-linear frequency changes. For example, to determine a measure of temporal acuity for signals with frequency transitions, we can devise a discrimination task in which listeners are asked to distinguish between a linear glide tone and a signal covering the same frequency change via a multiple-step trajectory (step frequency transitions). To generate the “step frequency transitions”, all we need to do...
is to redefine the frequency trajectory for the moving filter. In fact, the moving filter designed in our experiment can be adjusted to generate stimuli with any predefined frequency trajectory.

5.7.2 The Size of Roving Frequency Range

In the study of profile analysis using roving level (Green, 1988), subjects' performance exhibited a flat U-shaped pattern for rove ranges up to 40 dB. Performance does not show much change with the change in roving level range if the range is less than 40 dB. If a rove range larger than 40 dB is used, threshold change is observed. The 40 dB roving level range was therefore chosen for profile tasks. Our data do not show such a boundary. Performance decreases continuously with the increase in the size of rove-range up to 800 Hz. Our data do not show a point where the performance reach a plateau (or saturation). In section 5.2, we also discussed that the prediction model showed a performance plateau; performance cannot be more poorer when rove-range reached a large size.

Neither very small nor very large size of roving starting frequency ranges were examined in the present study. Future study can include these ranges. Larger range can be used to test if there is a performance decreasing boundary outside the largest range (800 Hz) tested in this study. We can use small
rove range in the task to see how much a range is need to cause the initial decrease in discrimination performance.
APPENDIX A

EXTRACTION OF EPN DATA FROM THE ASP MODEL

The Neural Excitation Pattern (EPN) is generated by \texttt{genepn} which is an individual module in the Auditory Sensation Processing (ASP) model (Patterson and Holdsworth, 1990). \texttt{genepn} generates a sequence of frames of excitation pattern. Each frame contains an array of numbers which represents the output across the channels at an instant of time. An example of neural excitation pattern for a 100 ms tone glide, starting at 1000 Hz and ending at 1600 Hz, is shown in Figure 45. Ten frames of excitation patterns are plotted in the graph. Each curve represents a frame of excitation which is sampled every 10 ms from the continuous neural excitation pattern of the ASP model. The first curve on the left corresponds to the onset of the signal and the last curve on the right corresponds to the offset of the signal. A frequency transition can be seen by visual inspection of the figure. The peaks of the excitation patterns move from the lower frequency (about 1000 Hz) to the higher frequency (about 1600 Hz).

To produce an output file from \texttt{genepn}, set the option \texttt{output = on} in the ASP model. \texttt{Genepn} will generate a binary file which contains the values of excitation patterns. In the
Figure 45: The Neural Excitation Pattern for a 100 ms tone glide starting at 1000 Hz and ending at 1600 Hz. Each curve represents a frame of data which is sampled every 10 ms from the continuous excitation pattern of ASP model.

ASP model the frequency is scaled in terms of the ERB (Equivalent Rectangular Bandwidth). The ERB can be described by the following equation (Glasberg and Moore, 1990):

\[
ERB = 24.7 \left(4.37f/1000 + 1\right),
\]

where \( f \) is frequency in Hz.

In order to calculate IWAIF values for excitation patterns conveniently, we need to transform the frequency scale from ERB to Hertz. The relation between number of ERBs (or ERBR, the effective rectangular bandwidth rate) and
frequency (in Hz) can be derived by integrating the reciprocal of ERB with respect to frequency (Glasberg and Moore, 1990).

\[
\text{Number of ERBs (ERBR)} = 21.3323 \times \log_{10}(4.37f/1000 + 1),
\]

(40)

and

\[
f = (\exp(\frac{\text{ERBR}}{9.2645}) - 1) \times 1000/4.37,
\]

(41)

where \( f \) is frequency in Hz.

By using the above equations, the frequencies corresponding to the ERB scale excitation pattern can be evaluated as follows:

1. Transform minimum and maximum center frequencies (Hz) to ERBRs. The minimum and maximum center frequencies can be obtained from the parameters \text{mincf}_afb and \text{maxcf}_afb in \text{genepn} module.

2. Calculate the ERBRs for all the channels. The number of channels in filter is obtained from the parameter \text{channels}_afb in \text{genepn} module.

3. Transform the ERBRs obtained in step 2 to frequencies in Hz.

The frequency scale in Figure 44 has been transformed to Hz from ERBs.

After we transformed the frequency scale from ERB to Hz, the next step was to extract the magnitude of excitation. A binary file was generated by \text{genepn} with the \text{output} option on. A computer program was implemented to translate the binary file containing the EPN values to an ASCII text file. The
ASCII file then was used for the input to IWAIF calculations. The purpose of using ASCII text file was that the program used for IWAIF calculation was written in MATLAB (The MathWorks Inc., 1990) and MATLAB could read ASCII code file.\footnote{The newest version of MATLAB (v 4.0) is able to read external binary files.}
APPENDIX B

TMS320C25 FUNCTIONS FOR THE IMPLEMENTATION OF MOVING FILTER

The program listed here is the TMS320C25 (Texas Instruments Incorporated, 1987) assembly program used to generate moving filter signals on DSP-16 board. The RESMON monitor (Ariel Corporation, 1987) which occupies the first 1 kbyte of the program memory is not included in this appendix. This appendix contains two user functions starting at the address 1024 and the filter parameter table.

; USER FUNCTIONS FOR filter.ASM
; sampling frequency is set to 25000 Hz dual DAC and ADC channels
; parameter table starts from 57344(E000)
; EQUATES FOR DATA PAGE ZERO VARIABLES
; On-board RAM and dedicated registers
; Start of on-board data RAM
;
ONE: EQU 96 /Constant = 1
ZERO: EQU 97 /Constant = 0
COMMAND: EQU 98 ;Host command word
HOSTBUF: EQU 99 ;Buffer host status
DUMMY: EQU 100 ;Dummy register
TBLADDR: EQU 101 ;point to the address of parameter table
RAMTYPE: EQU 102 ;6 for 1M buffer RAM
ALL1: EQU 103 ;ALL1 = FFFFH
TEMP: EQU 104
OTEMP: EQU 123
DISCARD: EQU 124 ;discard the beginning
TEST1: EQU 125
Y: EQU 105 ;DAC output buffer
Y1: EQU 106
Y2: EQU 107
X: EQU 108 ;ADC input buffer
X1: EQU 109
X2: EQU 110
STARTF: EQU RM_USER1 ;Start frequency for the moving filter
STOPF: EQU RM_USER2 ;Stop frequency for the moving filter
FSTEP: EQU RM_USER3 ;Frequency step, usually = 1
STEPSEG: EQU RM_USER4 ;sample points for each frequency step
UPDOWN: EQU RM_USER5 ;0=up 1=down
;constant filter coefficients x 2^15
CAO: EQU RM_USER6 ;A0[n]
CA2: EQU RM_USER7 ;A2[n]
CB2: EQU RM_USER8 ;B2[n]
;
DRR1: EQU 0 ;A/D Input data register
IMR1: EQU 4 ;Interrupt mask register
GREG1: EQU 5 ;Global memory allocation register
DRAM: EQU 512 ;Address of internal RAM B0 after CNFD
ORG 1024 ;user programs must begin at 1024 or greater!

;************************ Initialize 320 *************************

;user function #0 initialization

INIT:

DINT ;Disable interrupt
ROVM ;reset overflow mode
RSXM ;reset sign extension mode
SPM 0 ;set product mode for no-shift
FORT 0 ;Configure for 16 bit word serial data
LDPK 0 ;Address base RAM with data memory pointer
LACK 0E0H ;Data memory 57344-65535, program 0-8191
SACL GREG1 ;Write to global memory allocation register

;data ram 8k page 448-511, program ram 8k

;Load constants

ZAC
SACL ZERO ;ZERO is zero
OUT ZERO,REFRESH ;TURN REFRESH ON
SACL Y ;Out DAC buffer init 0
LACK 1 ;Load 1
SACL ONE ;One is one
SUB ONE,1
SACL ALL1 ;ALL1 = FFFFH
LACK 6
SACL RAMTYPE ;RAMTYPE = 6 for 1M buffer
OUT ZERO,WRDAC ;set DAC output to zero

; LRLK AR0,3500
; set up loop counter for transfer of parameter tables
; from pma to global dma
LRLK AR1,57344
;AR1 points to start of wave table dest. in global
LARP AR1 ;point to AR1
LALK PARTBL ;acc points to start of parameter tables in pma

COPY:

TBLR TEMP ;get a value from pma table
ADDS ONE ;bump table ptr
SACL OTEMP ;save the ptr
ZALS TEMP ;get data in acc
SACL *+,0,AR0 ;save it in global RAM; NEWARP = AR0
ZALS OTEMP ;get wave table ptr again
BANZ COPY,*-, AR1 ;continue till done, NEWARP = AR1

;Transfer program to internal RAM block 0 for fast execution

XFER:

CNFD ;Configure as data memory for xfer
LARP 0 ;Set address register pointer
LRLK 0, DRAM ;Point to destination in B0 RAM as data
RPTK 255 ;Fill page of program RAM
BLKP 1250, *+ ;Block move from LOOP to AR0
CNFP ;Configure RAM block 0 as program memory

INITEND:

IN HOSTBUF,RDSTAT ;Read host status
BIT HOSTBUF,15-1 ;TDR status bit 1
BBNZ INITEND ;If host port ready for data
OUT ALL1,WRHOST ;Output FFFFH to I/O port 15
RET

;****** user function #1. moving filter processing

FILTER:

DINT ;Disable interrupt
SPM 0 ;set product mode for no-shift
FORT 0
LDPK 0
LACK 0E0H
SACL GREG1
LALK 57344
ADDS STARTF
SACL TEMP
LAR AR1,STEPSEG ;STEPSEG saved in AR1
LAR AR3,TEMP ;AR3 POINTS TO THE STARTING POINT IN THE TABLE
LALK 57344
ADDS STOPF
SACL TEMP
LAR AR2,TEMP ;AR2 saves the STOP freq
LARP AR3
SOVM ;set overflow mode
SSXM ;set sign extension mode
ZAC
SACL X1
SACL X2
SACL Y1
SACL Y2
CNFP ;make sure B0 is program RAM
;REENT:
LACK 22 ;load AC with 22, enable RINT, AINT, BINT
SACL IMR1 ;Set interrupt mask register
B CHECK + PRAM ;Branch to internal program

;***** Foreground loop *****
ORG 1250
LOOP:
 IN HOSTBUF, RDSTAT ;Read host status
 BIT HOSTBUF, 15-0 ;Test host status
 BBZ LOOP + PRAM ;Branch if host data received
HOSTDAT:IN COMMAND, RDHOST ;Read host data
 B ENDPF + PRAM ;Any input will stop the function
CHECK:
 BIOZ OKANDGO + PRAM ;ok to start if we are in channel B time
 slot
 B CHECK + PRAM ;else wait for one to begin
OKANDGO:
 LDPK 0 ;switch to page 0
 EINT
 B LOOP + PRAM ;Branch to foreground loop
;
ENDPF1:
 POPD TEMP ;pop top of stack out
ENDPF:
 DINT ;Disable interrupt
 ZALS ALL1
 SACL DUMMY ;Dummy = FFFFH
HOSTP2:
 IN HOSTBUF, RDSTAT ;Read host status
 BIT HOSTBUF, 15-1 ;TDR status bit 1
 BBNZ HOSTP2 + PRAM ;If host port ready for data
 OUT DUMMY, WRHOST ;Output FFFFH to I/O port 15
 LACK 0E0H ;Reset Data memory
 SACL GREG1 ;Write to global memory allocation register
 LACK 0
SACL IMR1 ;Set IMR zero
RET ;Return

;***** Channel A output interrupt service routine *****
AINT:
OUT Y, WRDAC ;Load DAC LATCH
NOP
SXF ;Set XF flag

ZALS STEPSEG
SUBS ONE
SACL STEPSEG
BGZ NOEND1+PRAM ;continue if greater than zero
SAR AR1,STEPSEG ;restore STEPSEG

;update AR3, the pointer of parameter table
SAR AR3,TBLADDR ;AR3 -> TBLADDR
ZALS UPDOWN ;TEST GOING UP OR DOWN
BNZ CNTDOWN ;NOT ZERO GOING DOWN
ZALS TBLADDR
ADDS FSTEP
SACL TBLADDR
LAR AR3,TBLADDR ;TBLADDR -> AR3
SAR AR2,TEMP ;AR2 -> TEMP
SUBS TEMP
BLEZ NOEND1+PRAM ;not end if start freq <= stop freq
B ENDPRL+PRAM ;if start > stop then end

CNTDOWN:
ZALS TBLADDR
SUBS FSTEP
SACL TBLADDR
LAR AR3,TBLADDR ;TBLADDR -> AR3
SAR AR2,TEMP ;AR2 -> TEMP
SUBS TEMP
BLZ NOEND1+PRAM ;If start < stop then end

NOEND1:
EINT
RET

;BINT:
OUT X, WRDAC ;Load DAC LATCH
NOP
RXF ;Reset XF flag

NOEND2:
EINT
RET

PRAM: EQU 0FF00H - 1250
;this special equate is needed for offsetting addresses into the 32020's
;internal program RAM. 0FF00H=65280. This equate must be used when
;we configure block B0 as the internal program RAM.
;1200 is the address where the original program starts.

INFINT:
BIOZ BINP+PRAM ;if channel B input data
LAC DRR1
SACL X
LAC X,15 ;ACC <- X*2^15
SUB X1,15
ADD Y1,15
LT X2 ;X2 -> T
MPY CA2 ;X2*CA2 -> P
LTD X1 ; X1 -> T, X1 -> X2, ACC+P -> ACC
MPY * ; X1*(dma, A1) -> P
LTS X ; X -> T, ACC-P -> ACC
DMOV X ; X -> X1
MPY CA0 ; X*CA0 -> P
LTA Y2 ; Y2 -> T, ACC+P -> ACC
MPY CB2 ; Y2*CB2 -> P
LTS Y1 ; Y1 -> T, ACC-P -> ACC
DMOV Y1 ; Y1 -> Y2
MPY * ; Y1*(dma, B1) -> P
APAC ; ACC+P -> ACC
SACH Y, 1 ; ACCH*2 -> Y
DMOV Y ; Y -> Y1

; ENDINP:
  EINT
  RET

; BINP:
  LAC DRRL ; do nothing if channel B input
  EINT
  RET

; filter parameters for A1=B1
; center freq from 0 to 3500 Hz
; band width 70 Hz
; gain 40 dB
; sample freq 25000 Hz

; PARTBL:
DW 7DC4H, 7DC4H, 7DC4H, 7DC4H, 7DC4H
DW 7DC4H, 7DC4H, 7DC4H, 7DC4H, 7DC4H
DW 7DC4H, 7DC4H, 7DC4H, 7DC4H, 7DC4H
DW 7DC3H, 7DC3H, 7DC3H, 7DC3H, 7DC3H
DW 7DC3H, 7DC3H, 7DC3H, 7DC3H, 7DC3H
DW 7DC2H, 7DC2H, 7DC2H, 7DC2H, 7DC2H
DW 7DC2H, 7DC1H, 7DC1H, 7DC1H, 7DC1H
DW 7DC1H, 7DC1H, 7DC0H, 7DC0H, 7DC0H
DW 7DC0H, 7DC0H, 7DBFH, 7DBFH, 7DBFH
DW 7DBFH, 7DBFH, 7DBEH, 7DBEH, 7DBEH
DW 7DBEH, 7DBEH, 7DBDH, 7DBDH, 7DBDH
DW 7DBDH, 7DBCH, 7DBCH, 7DBCH, 7DBCH
DW 7DBBH, 7DBBH, 7DBBH, 7DBBH, 7DBAH
DW 7DBAH, 7DBAH, 7DB9H, 7DB9H, 7DB9H
DW 7DB8H, 7DB8H, 7DB8H, 7DB8H, 7DB7H
DW 7DB7H, 7DB7H, 7DB6H, 7DB6H, 7DB6H
DW 7DB5H, 7DB5H, 7DB4H, 7DB4H, 7DB4H
DW 7DB3H, 7DB3H, 7DB3H, 7DB2H, 7DB2H
DW 7DB2H, 7DB1H, 7DB1H, 7DB0H, 7DB0H
DW 7DB0H, 7DAFH, 7DAFH, 7DAEH, 7DAEH
DW 7DADH, 7DAOH, 7DADH, 7DACH, 7DACH
DW 7DABH, 7DABH, 7DAAH, 7DAAH, 7DA9H
DW 7DA9H, 7DA8H, 7DA8H, 7DA7H, 7DA7H
DW 7DA7H, 7DA6H, 7DA6H, 7DA5H, 7DA4H
DW 7DA4H, 7DA3H, 7DA3H, 7DA2H, 7DA2H
DW 7DA2H, 7DA1H, 7DA0H, 7DA0H, 7D9FH
DW 7D9FH, 7D9EH, 7D9EH, 7D9EH, 7D9CH
DW 7D9CH, 7D9BH, 7D9BH, 7D9AH, 7D99H
DW 7D99H, 7D98H, 7D98H, 7D97H, 7D96H
DW 7D96H, 7D95H, 7D95H, 7D94H, 7D93H
DW 7D93H, 7D92H, 7D91H, 7D91H, 7D90H
| DW   | 7D90H, 7D8FH, 7D8EH, 7D8EH, 7D8DH |
| DW   | 7D8CH, 7D8BH, 7D8BH, 7D8AH, 7D89H |
| DW   | 7D89H, 7D88H, 7D87H, 7D87H, 7D86H |
| DW   | 7D85H, 7D84H, 7D84H, 7D83H, 7D82H |
| DW   | 7D82H, 7D81H, 7D80H, 7D7FH, 7D7FH |
| DW   | 7D7EH, 7D7DH, 7D7CH, 7D7CH, 7D7BH |
| DW   | 7D7AH, 7D79H, 7D78H, 7D78H, 7D77H |
| DW   | 7D76H, 7D75H, 7D74H, 7D74H, 7D73H |
| DW   | 7D72H, 7D71H, 7D70H, 7D6FH, 7D6FH |
| DW   | 7D6EH, 7D6DH, 7D6CH, 7D6BH, 7D6AH |
| DW   | 7D6AH, 7D69H, 7D68H, 7D67H, 7D66H |
| DW   | 7D65H, 7D64H, 7D63H, 7D63H, 7D62H |
| DW   | 7D61H, 7D60H, 7D5FH, 7D5EH, 7D5DH |
| DW   | 7D5CH, 7D5BH, 7D5AH, 7D59H, 7D58H |
| DW   | 7D57H, 7D56H, 7D55H, 7D54H, 7D54H |
| DW   | 7D53H, 7D52H, 7D51H, 7D50H, 7D4FH |
| DW   | 7D4EH, 7D4DH, 7D4CH, 7D4BH, 7D4AH |
| DW   | 7D49H, 7D48H, 7D47H, 7D46H, 7D45H |
| DW   | 7D44H, 7D43H, 7D42H, 7D41H, 7D40H |
| DW   | 7D3FH, 7D3EH, 7D3DH, 7D3CH, 7D3AH |
| DW   | 7D39H, 7D38H, 7D37H, 7D36H, 7D35H |
| DW   | 7D34H, 7D33H, 7D32H, 7D31H, 7D30H |
| DW   | 7D2FH, 7D2DH, 7D2CH, 7D2BH, 7D2AH |
| DW   | 7D29H, 7D28H, 7D27H, 7D26H, 7D24H |
| DW   | 7D23H, 7D22H, 7D21H, 7D20H, 7D1FH |
| DW   | 7D1DH, 7D1CH, 7D1BH, 7D1AH, 7D19H |
| DW   | 7D18H, 7D16H, 7D15H, 7D14H, 7D13H |
| DW   | 7D12H, 7D10H, 7D0FH, 7D0EH, 7D0DH |
| DW   | 7D0BH, 7D0AH, 7D09H, 7D08H, 7D07H |
| DW   | 7D05H, 7D04H, 7D03H, 7D01H, 7D00H |
| DW   | 7CF9H, 7CF7H, 7CF6H, 7CF5H, 7CF3H |
| DW   | 7CF2H, 7CF1H, 7CEFH, 7CEEH, 7CEDH |
| DW   | 7CEBH, 7CEAH, 7CE9H, 7CE7H, 7CB6H |
| DW   | 7CE5H, 7CE3H, 7CE2H, 7CE1H, 7CDFH |
| DW   | 7CDBH, 7CDDH, 7CDBH, 7CDAH, 7CD8H |
| DW   | 7CD7H, 7CD6H, 7CD4H, 7CD3H, 7CD1H |
| DW   | 7CD0H, 7CCFH, 7CCDH, 7CCCH, 7CAH |
| DW   | 7CC9H, 7CC7H, 7CC6H, 7CC5H, 7CC3H |
| DW   | 7CC2H, 7CC0H, 7CBFH, 7CBDH, 7CBCH |
| DW   | 7CB9H, 7CB7H, 7CB6H, 7CB4H, 7CB4H |
| DW   | 7CB3H, 7CB1H, 7CB0H, 7CAEH, 7CADH |
| DW   | 7CABH, 7CA9H, 7CA8H, 7CA7H, 7CA5H |
| DW   | 7CA4H, 7CA2H, 7CA1H, 7C9FH, 7C9EH |
| DW   | 7C9CH, 7C9AH, 7C97H, 7C96H, 7C95H |
| DW   | 7C94H, 7C93H, 7C91H, 7C8FH, 7C8EH |
| DW   | 7C8CH, 7C8BH, 7C89H, 7C87H, 7C86H |
| DW   | 7C84H, 7C83H, 7C81H, 7C7EH, 7C7EH |
| DW   | 7C7CH, 7C7AH, 7C79H, 7C77H, 7C75H |
| DW   | 7C74H, 7C72H, 7C70H, 7C6FH, 7C6DH |
| DW   | 7C6BH, 7C6AH, 7C68H, 7C66H, 7C65H |
| DW   | 7C63H, 7C61H, 7C60H, 7C5EH, 7C5CH |
| DW   | 7C5AH, 7C59H, 7C57H, 7C55H, 7C54H |
| DW   | 7C52H, 7C50H, 7C4EH, 7C4DH, 7C4BH |
| DW   | 7C49H, 7C47H, 7C45H, 7C44H, 7C42H |
| DW   | 7C40H, 7C3EH, 7C3DH, 7C3BH, 7C39H |
| DW   | 7C37H, 7C35H, 7C34H, 7C32H, 7C30H |
| DW   | 7C2EH, 7C2CH, 7C2AH, 7C29H, 7C27H |
| DW   | 7C25H, 7C23H, 7C21H, 7C1FH, 7C1EH |
| DW   | 7C1CH, 7C1AH, 7C18H, 7C16H, 7C14H |
| DW   | 7C12H, 7C10H, 7C0FH, 7C0DH, 7C0BH |
| DW   | 7C09H, 7C07H, 7C05H, 7C03H, 7C01H |
DW 7BFFH, 7BFDH, 7BFBH, 7BFAH, 7BF8H
DW 7BF6H, 7BF4H, 7BF2H, 7BF0H, 7BEEH
DW 7BEC8H, 7BEAH, 7BE8H, 7BE6H, 7BE4H
DW 7BE2H, 7BEOH, 7BDEH, 7BDCH, 7BDAH
DW 7B8DH, 7B6DH, 7B4DH, 7B2DH, 7B0DH
DW 7BCEH, 7BCCH, 7BCAH, 7BC8H, 7BC6H
DW 7BC4H, 7BC2H, 7BC0H, 7BEEH, 7BBCH
DW 7BB9H, 7BB7H, 7BB5H, 7BB3H, 7BB1H
DW 7BAFH, 7BADH, 7BA9H, 7BA7H, 7BA5H
DW 7B5AH, 7B58H, 7B56H, 7B54H, 7B52H, 7B4FH
DW 7B4DH, 7B4BH, 7B49H, 7B46H, 7B44H
DW 7B42H, 7B38H, 7B36H, 7B34H, 7B32H, 7B2FH, 7B2DH
DW 7B2BH, 7B28H, 7B26H, 7B24H, 7B21H
DW 7B1FH, 7B1CH, 7B1AH, 7B18H, 7B15H
DW 7B13H, 7B11H, 7B0EH, 7B0CH, 7B0AH
DW 7B07H, 7B05H, 7B02H, 7B00H, 7AFEH
DW 7AFBH, 7AF9H, 7AF6H, 7AF4H, 7AF1H
DW 7AEFH, 7AEDH, 7AEAH, 7AE8H, 7AE5H
DW 7AE3H, 7AEOH, 7ADEH, 7AD8H, 7AD9H
DW 7AD6H, 7AD4H, 7AD2H, 7ACFH, 7ACDH
DW 7ACAH, 7AC8H, 7AC5H, 7AC3H, 7AC0H
DW 7ABEH, 7ABBH, 7A99H, 7AB6H, 7AB3H
DW 7AB1H, 7AAEH, 7AA9H, 7AA6H, 7AA3H
DW 7AA4H, 7AA2H, 7A9FH, 7A9DH, 7A9AH
DW 7A97H, 7A95H, 7A92H, 7A90H, 7A8DH
DW 7A8AH, 7A88H, 7A85H, 7A83H, 7A80H
DW 7A7DH, 7A7BH, 7A78H, 7A76H, 7A73H
DW 7A70H, 7A6EH, 7A6BH, 7A68H, 7A66H
DW 7A63H, 7A6BH, 7A5EH, 7A5BH, 7A58H
DW 7A56H, 7A53H, 7A50H, 7A4EH, 7A4BH
DW 7A48H, 7A46H, 7A43H, 7A40H, 7A3EH
DW 7A3BH, 7A38H, 7A35H, 7A33H, 7A30H
DW 7A2DH, 7A2AH, 7A28H, 7A25H, 7A22H
DW 7A1FH, 7A1DH, 7A1AH, 7A17H, 7A14H
DW 7A12H, 7A0FH, 7A0CH, 7A09H, 7A06H
DW 7A04H, 7A01H, 79FEH, 79FBH, 79F8H
DW 79F6H, 79F3H, 79F0H, 79EDH, 79E8H
DW 79E7H, 79E5H, 79E2H, 79DFH, 79DCH
DW 79D9H, 79D6H, 79D4H, 79D1H, 79CEH
DW 79CBH, 79C8H, 79C5H, 79C2H, 79BFH
DW 79BCH, 79BAH, 79B7H, 79B4H, 79B1H
DW 79AEH, 79ABH, 79A8H, 79A5H, 79A2H
DW 799FH, 799CH, 7999H, 7996H, 7994H
DW 7991H, 798EH, 798BH, 7988H, 7985H
DW 7982H, 797FH, 797CH, 7979H, 7976H
DW 7973H, 7970H, 796DH, 796AH, 7967H
DW 7964H, 7961H, 795EH, 795BH, 7958H
DW 7955H, 7952H, 794EH, 794BH, 7948H
DW 7945H, 7942H, 793FH, 793CH, 7939H
DW 7936H, 7933H, 7930H, 792DH, 792AH
DW 7927H, 7923H, 7920H, 791DH, 791AH
DW 7917H, 7914H, 7911H, 790EH, 790AH
DW 7907H, 7904H, 7901H, 78FEH, 78FBH
DW 78F8H, 78F4H, 78F1H, 78EEH, 78EBH
DW 744EH, 7449H, 7445H, 7440H, 743CH
DW 7437H, 7433H, 742FH, 742AH, 7426H
DW 7421H, 741DH, 7418H, 7414H, 740FH
DW 740BH, 7406H, 7402H, 73FDH, 73F9H
DW 73F4H, 73F0H, 73EBH, 73E7H, 73E2H
DW 73DEH, 73D9H, 73D5H, 73D0H, 73CCCH
DW 73C7H, 73C3H, 73BEH, 73BAH, 73B5H
DW 73B1H, 73ACCH, 73A7H, 73A3H, 739EH
DW 739AH, 7395H, 7391H, 738CH, 7387H
DW 7383H, 737E7H, 737AH, 7375H, 7370H
DW 736CH, 7367H, 7363H, 735EH, 7359H
DW 7355H, 7350H, 734BH, 7347H, 7342H
DW 733DH, 7339H, 7334H, 732FH, 732BH
DW 7326H, 7321H, 731DH, 7318H, 7313H
DW 730FH, 730AH, 7305H, 7301H, 72FCH
DW 72F7H, 72F2H, 72EEH, 72E9H, 72E4H
DW 72E0H, 72DBH, 72D6H, 72D1H, 72CDH
DW 72C8H, 72C3H, 72BCH, 72BAH, 72B5H
DW 72B0H, 72ABH, 72A6H, 72A2H, 729DH
DW 7298H, 7293H, 728FH, 728AH, 7285H
DW 7280H, 727BH, 7276H, 7272H, 726DH
DW 7268H, 7263H, 725EH, 725AH, 7255H
DW 7250H, 724BH, 7246H, 7241H, 723CH
DW 7238H, 7233H, 722EH, 7229H, 7224H
DW 721FH, 721AH, 7215H, 7210H, 720CH
DW 7207H, 7202H, 71FDH, 71F8H, 71F3H
DW 71EEH, 71E9H, 71E4H, 71DFH, 71DAH
DW 71D5H, 71D0H, 71CBH, 71C6H, 71C2H
DW 71B8H, 71B3H, 71AEH, 71A9H
DW 71A4H, 719FH, 719AH, 7195H, 7190H
DW 718BH, 7186H, 7181H, 717CH, 7177H
DW 7172H, 716DH, 7168H, 7163H, 715DH
DW 7158H, 7153H, 714EH, 7149H, 7144H
DW 713FH, 713AH, 7135H, 7130H, 712BH
DW 7126H, 7121H, 711CH, 7117H, 7111H
DW 710CH, 7107H, 7102H, 70FDH, 70F8H
DW 70F3H, 70EEH, 70E8H, 70E3H, 70EBH
DW 70D9H, 70D4H, 70CFH, 70CAH, 70C4H
DW 70B8H, 70BAH, 70B5H, 70B0H, 70ABH
DW 70A5H, 70A0H, 709BH, 7096H, 7091H
DW 708CH, 7086H, 7081H, 707CH, 7077H
DW 7071H, 706CH, 7067H, 7062H, 705DH
DW 7057H, 7052H, 704DH, 7048H, 7042H
DW 703DH, 7038H, 7033H, 702DH, 7028H
DW 7023H, 701DH, 7018H, 7013H, 7008H
DW 7008H, 7003H, 6FFE8H, 6FF8H, 6FF3H
DW 6FFE8H, 6FE8H, 6FE3H, 6FDEH, 6FD8H
DW 6FD3H, 6FCEH, 6FC8H, 6FC3H, 6FBEH
DW 6FB8H, 6FB3H, 6FAEH, 6FA8H, 6FA3H
DW 6F9BH, 6F96H, 6F93H, 6F80H, 6F88H
DW 6F83H, 6F7DH, 6F78H, 6F72H, 6F6DH
DW 6F68H, 6F62H, 6F5DH, 6F57H, 6F52H
DW 6F4CH, 6F47H, 6F42H, 6F3CH, 6F37H
DW 6F31H, 6F2CH, 6F26H, 6F21H, 6F1BH
DW 6F16H, 6F10H, 6F0BH, 6F05H, 6F00H
DW 6EFAH, 6EF5H, 6EEFH, 6EEAH, 6EE4H
DW 6EDFH, 6ED9H, 6ED4H, 6ECEH, 6EC9H
DW 6EC3H, 6EBEH, 6EB8H, 6EBAH, 6EA3H
DW 6EA8H, 6EA2H, 6E9DH, 6E97H, 6E91H
DW 6E8CH, 6E86H, 6E81H, 6E7BH, 6E76H
DW 6E70H, 6E6AH, 6E65H, 6E5FH, 6E5AH
DW 6E54H, 6E4EH, 6E49H, 6E43H, 6E3EH
DW 5DC2H, 5DBAH, 5DB2H, 5DAAH, 5DA3H
DW 5D9BH, 5D93H, 5DBBH, 5D83H, 5D7BH
DW 5D73H, 5D6BH, 5D63H, 5D5BH, 5D53H
DW 5D4BH, 5D43H, 5D3BH, 5D33H, 5D2BH
DW 5D23H, 5D1BH, 5D13H, 5D0BH, 5D03H
DW 5CFBH, 5CF3H, 5CEBH, 5CE3H, 5CDBH
DW 5CD3H, 5CCBH, 5CC2H, 5CBAH, 5CB2H
DW 5CAAH, 5CA2H, 5C9AH, 5C92H, 5C8AH
DW 5C82H, 5C7AH, 5C72H, 5C6AH, 5C62H
DW 5C5AH, 5C51H, 5C49H, 5C41H, 5C39H
DW 5C31H, 5C29H, 5C21H, 5C19H, 5C10H
DW 5C08H, 5C00H, 5BF8H, 5BF0H, 5BECF
DW 5B0EH, 5B07H, 5B0FH, 5B9EH, 5B96H
DW 5B96H, 5B88H, 5B7EH, 5B75H, 5B6DH
DW 5B65H, 5B55H, 5B4CH, 5B44H
DW 5B3CH, 5B34H, 5B2BH, 5B23H, 5B1BH
DW 5B13H, 5B0AH, 5B02H, 5FAFH, 5FA2H
DW 5AE9H, 5AE1H, 5AD9H, 5AD1H, 5AC8H
DW 5AC0H, 5AB8H, 5AB0H, 5AA7H, 5A9FH
DW 5A97H, 5A8EH, 5A86H, 5A7EH, 5A76H
DW 5A6DH, 5A65H, 5A5DH, 5A54H, 5A4CH
DW 5A44H, 5A3BH, 5A33H, 5A2BH, 5A22H
DW 5A1AH, 5A12H, 5A09H, 5A01H, 59F8H
DW 59F0H, 59FEH, 59DFH, 59D7H, 59CFH
DW 59C6H, 59BEH, 59B5H, 59ADH, 59A5H
DW 599CH, 5994H, 598CH, 5983H, 597BH
DW 5972H, 596AH, 5961H, 5959H, 5951H
DW 5948H, 5940H, 5937H, 592FH, 5926H
DW 591EH, 5915H, 590DH, 5905H, 58FCH
DW 58F4H, 58EBH, 58E3H, 58DAH, 58D2H
DW 58C9H, 58C1H, 58B8H, 58B0H, 58A7H
DW 589FH, 5896H, 588EH, 5885H, 587DH
DW 5874H, 586CH, 5863H, 585BH, 5852H
DW 584AH, 5841H, 5838H, 5830H, 5827H
DW 581FH, 5816H, 580EH, 5805H, 57FDH
DW 57F4H, 57EBH, 57E3H, 57DAH, 57D2H
DW 57C9H, 57C0H, 57B8H, 57AFH, 57A7H
DW 579EH, 5795H, 578DH, 5784H, 577CH
DW 5773H, 576AH, 5762H, 5759H, 5750H
DW 5748H, 573FH, 5736H, 572EH, 5725H
DW 571CH, 5714H, 570BH, 5702H, 56FAH
DW 56F1H, 56E8H, 56E0H, 56D7H, 56CEH
DW 56C6H, 56BDH, 56B4H, 56ACH, 56A3H
DW 569AH, 5691H, 5689H, 5680H, 5677H
DW 566FH, 5666H, 565DH, 5654H, 564CH
DW 5643H, 563AH, 5631H, 5629H, 5620H
DW 5617H, 560EH, 5606H, 55FDH, 55F4H
DW 55EBH, 55E2H, 55DAH, 55D1H, 55C8H
DW 55B6H, 55A3H, 55A5H, 55A9H, 55C9H
DW 5593H, 558AH, 5502H, 5579H, 5570H
DW 5567H, 555EH, 5555H, 554DH, 5544H
DW 553BH, 5532H, 5529H, 5520H, 5517H
DW 550FH, 5506H, 54FDH, 54F4H, 54EBH
DW 54E2H, 54D9H, 54D0H, 54C7H, 54B6H
DW 54B6H, 54A4H, 549BH, 5492H
DW 5489H, 5480H, 5477H, 546EH, 5465H
DW 545CH, 5454H, 544BH, 5442H, 5439H
DW 5430H, 5427H, 541EH, 5415H, 540CH
DW 5403H, 53FAH, 53F1H, 53E8H, 53DFH
DW 53D6H, 53CDH, 53C4H, 53BBH, 53B2H
DW 53A9H, 53A0H, 5397H, 538EH, 5385H
<table>
<thead>
<tr>
<th>DW</th>
<th>537CH, 5373H, 536AH, 5361H, 5358H</th>
</tr>
</thead>
<tbody>
<tr>
<td>DW</td>
<td>534FH, 5346H, 533DH, 5334H, 532BH</td>
</tr>
<tr>
<td>DW</td>
<td>5322H, 5318H, 530FH, 5306H, 52FDH</td>
</tr>
<tr>
<td>DW</td>
<td>52F4H, 52EBH, 52E2H, 52D9H, 52D0H</td>
</tr>
<tr>
<td>DW</td>
<td>52C7H, 52B7H, 52B5H, 52ABH, 52A2H</td>
</tr>
<tr>
<td>DW</td>
<td>5299H, 5290H, 5287H, 527EH, 5275H</td>
</tr>
<tr>
<td>DW</td>
<td>526CH, 5263H, 5259H, 5250H, 5247H</td>
</tr>
<tr>
<td>DW</td>
<td>523EH, 5235H, 522CH, 5223H, 5219H</td>
</tr>
<tr>
<td>DW</td>
<td>5210H, 5207H, 51FEH, 51F5H, 51ECH</td>
</tr>
<tr>
<td>DW</td>
<td>51E2H, 51D9H, 51D0H, 51C7H, 51BEH</td>
</tr>
<tr>
<td>DW</td>
<td>51B4H, 51ABH, 51A2H, 5199H, 5190H</td>
</tr>
<tr>
<td>DW</td>
<td>5186H, 517DH, 5174H, 516BH, 5162H</td>
</tr>
<tr>
<td>DW</td>
<td>5158H, 514FH, 5146H, 513DH, 5133H</td>
</tr>
<tr>
<td>DW</td>
<td>512AH, 5121H, 5118H, 510EH, 5105H</td>
</tr>
<tr>
<td>DW</td>
<td>50FCH, 50F3H, 50E9H, 50EOH, 50D7H</td>
</tr>
<tr>
<td>DW</td>
<td>50CEH, 50C4H, 50BBH, 50B2H, 50A8H</td>
</tr>
<tr>
<td>DW</td>
<td>509FH, 5096H, 508DH, 5083H, 507AH</td>
</tr>
<tr>
<td>DW</td>
<td>5071H, 5067H, 505EH, 5055H, 504BH</td>
</tr>
<tr>
<td>DW</td>
<td>5042H, 5039H, 502FH, 5026H, 501DH</td>
</tr>
<tr>
<td>DW</td>
<td>5013H, 500AH, 5001H, 4FF7H, 4FEEH</td>
</tr>
<tr>
<td>DW</td>
<td>4FE5H, 4FDBH, 4FD2H, 4FC8H, 4FBFH</td>
</tr>
<tr>
<td>DW</td>
<td>4FB6H, 4FACH, 4FA3H, 4F9AH, 4F90H</td>
</tr>
<tr>
<td>DW</td>
<td>4F87H, 4F7DH, 4F74H, 4F6BH, 4F61H</td>
</tr>
<tr>
<td>DW</td>
<td>4F58H, 4F4EH, 4F45H, 4F3BH, 4F32H</td>
</tr>
<tr>
<td>DW</td>
<td>4F29H, 4F1FH, 4F16H, 4F0CH, 4F03H</td>
</tr>
<tr>
<td>DW</td>
<td>4EF9H, 4EF0H, 4EE6H, 4EDDH, 4ED4H</td>
</tr>
<tr>
<td>DW</td>
<td>4ECAH, 4EC1H, 4EB7H, 4EAEH, 4EA4H</td>
</tr>
<tr>
<td>DW</td>
<td>4E9BH, 4E91H, 4E88H, 4E7EH, 4E75H</td>
</tr>
<tr>
<td>DW</td>
<td>4E6BH, 4E62H, 4E58H, 4E4FH, 4E45H</td>
</tr>
<tr>
<td>DW</td>
<td>4E3CH, 4E32H, 4E29H, 4E1FH, 4E16H</td>
</tr>
<tr>
<td>DW</td>
<td>4E0CH, 4E03H, 4DF9H, 4DF0H, 4DE6H</td>
</tr>
<tr>
<td>DW</td>
<td>4DDCH, 4DD3H, 4DC9H, 4DC0H, 4DB6H</td>
</tr>
<tr>
<td>DW</td>
<td>4DADH, 4DA3H, 4D9AH, 4D90H, 4D86H</td>
</tr>
<tr>
<td>DW</td>
<td>4D7DH, 4D73H, 4D6AH, 4D60H, 4D57H</td>
</tr>
<tr>
<td>DW</td>
<td>4D4DH, 4D43H, 4D3AH, 4D30H, 4D26H</td>
</tr>
<tr>
<td>DW</td>
<td>4D1DH, 4D13H, 4D0AH, 4D00H, 4CF6H</td>
</tr>
<tr>
<td>DW</td>
<td>4CEDH, 4CE3H, 4CDAH, 4CD0H, 4CC6H</td>
</tr>
<tr>
<td>DW</td>
<td>4CBDH, 4CB3H, 4CA9H, 4CA0H, 4C96H</td>
</tr>
<tr>
<td>DW</td>
<td>4C8CH, 4C83H, 4C79H, 4C6FH, 4C66H</td>
</tr>
<tr>
<td>DW</td>
<td>4C5CH, 4C52H, 4C49H, 4C3FH, 4C35H</td>
</tr>
<tr>
<td>DW</td>
<td>4C2CH, 4C22H, 4C18H, 4C0EH, 4C05H</td>
</tr>
<tr>
<td>DW</td>
<td>4BFBH, 4BF1H, 4BE8H, 4BDEH, 4BD4H</td>
</tr>
<tr>
<td>DW</td>
<td>4BCAH, 4BC1H, 4BB7H, 4BADH, 4BA3H</td>
</tr>
<tr>
<td>DW</td>
<td>4B9AH, 4B90H, 4B86H, 4B7CH, 4B73H</td>
</tr>
<tr>
<td>DW</td>
<td>4B69H, 4B5FH, 4B55H, 4B4CH, 4B42H</td>
</tr>
<tr>
<td>DW</td>
<td>4B07H, 4AFDH, 4AF4H, 4AEAH, 4AE0H</td>
</tr>
<tr>
<td>DW</td>
<td>4AD6H, 4ACCH, 4AC2H, 4AB9H, 4AAFH</td>
</tr>
<tr>
<td>DW</td>
<td>4AA5H, 4A9BH, 4A91H, 4A88H, 4A7EH</td>
</tr>
<tr>
<td>DW</td>
<td>4A74H, 4A6AH, 4A60H, 4A56H, 4A4CH</td>
</tr>
<tr>
<td>DW</td>
<td>4A43H, 4A39H, 4A2FH, 4A25H, 4A1BH</td>
</tr>
<tr>
<td>DW</td>
<td>4A11H, 4A07H, 49FDH, 49F4H, 49EAH</td>
</tr>
<tr>
<td>DW</td>
<td>49E0H, 49D6H, 49CCH, 49C2H, 49B8H</td>
</tr>
<tr>
<td>DW</td>
<td>49AEH, 49A4H, 499AH, 4991H, 4987H</td>
</tr>
<tr>
<td>DW</td>
<td>497DH, 4973H, 4969H, 495FH, 4955H</td>
</tr>
<tr>
<td>DW</td>
<td>494BH, 4941H, 4937H, 492DH, 4923H</td>
</tr>
<tr>
<td>DW</td>
<td>4919H, 490FH, 4905H, 48F8H, 48F1H</td>
</tr>
<tr>
<td>DW</td>
<td>48E7H, 48DDH, 48D4H, 48CAH, 48C0H</td>
</tr>
<tr>
<td>DW</td>
<td>48B6H, 48ACH, 48A2H, 4898H, 488EH</td>
</tr>
<tr>
<td>DW</td>
<td>4884H, 487AH, 4870H, 4866H, 485BH</td>
</tr>
<tr>
<td>DW</td>
<td>4851H, 4847H, 483DH, 4833H, 4829H</td>
</tr>
<tr>
<td>DW</td>
<td>481FH, 4815H, 480BH, 4801H, 47F7H</td>
</tr>
</tbody>
</table>
DW 3B28H, 3B1DH, 3B12H, 3B07H, 3AFCH
DW 3AF1H, 3AE5H, 3ADAH, 3ACFH, 3AC4H
DW 3AB9H, 3AAEH, 3A33H, 3A98H, 3A8DH
DW 3A82H, 3A77H, 3A6CH, 3A61H, 3A56H
DW 3A4BH, 3A3FH, 3A34H, 3A29H, 3A1EH
DW 3A13H, 3A08H, 39FDH, 39F2H, 39E7H
DW 39DCH, 39D0H, 39C5H, 39BAM, 39AFH
DW 39A4H, 3999H, 398EH, 3983H, 3977H
DW 396CH, 3961H, 3956H, 394BH, 3940H
DW 3934H, 3929H, 391EH, 3913H, 3908H
DW 38FDH, 38F1H, 38E6H, 38DBH, 38D0H
DW 38C5H, 38B9H, 38AEH, 38A3H, 3888H
DW 388DH, 3881H, 3876H, 386BH, 3860H
DW 3855H, 3849H, 383EH, 3833H, 3828H
DW 381CH, 3811H, 3806H, 37FBH, 37EFH
DW 37E4H, 37D9H, 37CEH, 37C2H, 37B7H
DW 37ACH, 37A1H, 3795H, 378AH, 377FH
DW 3774H, 3768H, 375DH, 3752H, 3746H
DW 373BH, 3730H, 3725H, 3719H, 370EH
DW 3703H, 36F7H, 36ECH, 36E1H, 36D5H
DW 36CAH, 36BFH, 36B3H, 36A8H, 369DH
DW 3691H, 3686H, 367BH, 366FH, 3664H
DW 3659H, 364DH, 3642H, 3637H, 362BH
DW 3620H, 3614H, 3609H, 35FEH, 35F2H
DW 35B7H, 35DCH, 35D0H, 35C5H, 35B9H
DW 35AEH, 35A3H, 3597H, 358CH, 3580H
DW 3575H, 356AH, 355EH, 3553H, 3547H
DW 353CH, 3530H, 3525H, 351AH, 350EH
DW 3503H, 34F7H, 34E5H, 34E0H, 34D5H
DW 34C9H, 34BEH, 34B2H, 34A7H, 349CH
DW 3490H, 3485H, 3479H, 346EH, 3462H
DW 3457H, 344BH, 3440H, 3434H, 3429H
DW 341DH, 3412H, 3406H, 33FBH, 33EFH
DW 33E4H, 33D8H, 33CDH, 33C1H, 33B6H
DW 33AAH, 339FH, 3393H, 3387H, 337CH
DW 3370H, 3365H, 3359H, 334EH, 3342H
DW 3337H, 332BH, 331FH, 3314H, 3308H
DW 32FDH, 32F1H, 32E6H, 32DAH, 32CEH
DW 32C3H, 32B7H, 32A1H, 32A0H, 3294H
DW 3289H, 327DH, 3272H, 3266H, 325AH
DW 324FH, 3243H, 3238H, 322CH, 3220H
DW 3215H, 3209H, 31FDH, 31F2H, 31E6H
DW 31DBH, 31CFH, 31C3H, 31B8H, 31A0H
DW 31A0H, 3195H, 3189H, 317DH, 3172H
DW 3166H, 315AH, 314EH, 3143H, 3137H
DW 312BH, 3120H, 3114H, 3108H, 30FDH
DW 30F1H, 30E5H, 30DAH, 30CEH, 30C2H
DW 30B6H, 30A5H, 309FH, 3093H, 3088H
DW 307CH, 3070H, 3064H, 3059H, 304DH
DW 3041H, 3035H, 302AH, 301EH, 3012H
DW 3006H, 2FFBH, 2FEFH, 2FE3H, 2FD7H
DW 2FCCH, 2FC0H, 2FB4H, 2FA8H, 2F9CH
DW 2F91H, 2F85H, 2F79H, 2F6DH, 2F61H
DW 2F56H, 2F4AH, 2F3EH, 2F32H, 2F26H
DW 2F1BH, 2F0FH, 2F03H, 2FE7H, 2EBBH
DW 2EDFH, 2ED4H, 2EC8H, 2EBCH, 2EB0H
DW 2EA4H, 2E98H, 2E8DH, 2E81H, 2E75H
DW 2E69H, 2E5DH, 2E51H, 2E45H, 2E3AH
DW 2E2EH, 2E22H, 2E16H, 2E0AH, 2DFEH
DW 2DF2H, 2DE6H, 2DDAH, 2DCFH, 2DC3H
DW 2DB7H, 2DBBH, 2DF9H, 2D93H, 2D87H
DW 2D7BH, 2D6FH, 2D63H, 2D57H, 2D4CH
| DW   | 2D40H, 2D34H, 2D28H, 2D1CH, 2D10H |
| DW   | 2D04H, 2CF8H, 2CECH, 2CE0H, 2CD4H |
| DW   | 2CC8H, 2CBCH, 2CB0H, 2CA4H, 2C98H |
| DW   | 2C8CH, 2C80H, 2C74H, 2C68H, 2C5CH |
| DW   | 2C50H, 2C44H, 2C38H, 2C2CH, 2C20H |
| DW   | 2C14H, 2C08H, 2BFCH, 2BF0H, 2BE4H |
| DW   | 2BD8H, 2BCCH, 2B0CH, 2B04H, 2BA8H |
| DW   | 2B9CH, 2B90H, 2B84H, 2B78H, 2B6CH |
| DW   | 2B60H, 2B54H, 2B48H, 2B3CH, 2B30H |
| DW   | 2B24H, 2B18H, 2B0CH, 2B00H, 2AF4H |
| DW   | 2AE8H, 2ADCH, 2ACFH, 2AC3H, 2AB7H |
| DW   | 2AABH, 2A9FH, 2A93H, 2A87H, 2A7BH |
| DW   | 2A6FH, 2A63H, 2A57H, 2A4AH, 2A3EH |
| DW   | 2A32H, 2A26H, 2A1AH, 2A0EH, 2A02H |
| DW   | 29F6H, 29EAH, 29DDH, 29D1H, 29C5H |
| DW   | 29B9H, 29ADH, 29A1H, 2995H, 2988H |
| DW   | 297CH, 2970H, 2964H, 2958H, 294CH |
| DW   | 2940H, 2933H, 2927H, 291BH, 290FH |
| DW   | 2903H, 28F6H, 28EAH, 28DEH, 28D2H |
| DW   | 28C6H, 28BAH, 28ADH, 28A1H, 2895H |
| DW   | 2889H, 287DH, 2870H, 2864H, 2858H |
| DW   | 284CH, 283FH, 2833H, 2827H, 281BH |
| DW   | 280FH, 2802H, 27F6H, 27EAH, 27DCH |
| DW   | 27D1H, 27C5H, 27B9H, 27ADH, 27A0H |
| DW   | 2794H, 2788H, 277CH, 276FH, 2763H |
| DW   | 2757H, 274AH, 273EH, 2732H, 2726H |
| DW   | 2719H, 270DH, 2701H, 26F4H, 26E8H |
| DW   | 26DCH, 26D0H, 26C3H, 26B7H, 26ABH |
| DW   | 269EH, 2692H, 2686H, 2679H, 266DH |
| DW   | 2661H, 2654H, 2648H, 263CH, 262FH |
| DW   | 2623H, 2617H, 260AH, 25FEH, 25F2H |
| DW   | 25A7H, 259BH, 258FH, 2582H, 2576H |
| DW   | 256AH, 255DH, 2551H, 2544H, 2538H |
| DW   | 252CH, 251FH, 2513H, 2506H, 2FAAH |
| DW   | 24EEH, 24E1H, 24D5H, 24C8H, 24BCH |
| DW   | 24B0H, 24A3H, 2497H, 248AH, 247EH |
| DW   | 2471H, 2465H, 2459H, 244CH, 2440H |
| DW   | 2433H, 2427H, 241AH, 240EH, 2401H |
| DW   | 23F5H, 23E8H, 23DCH, 23D0H, 23C3H |
| DW   | 23B7H, 23AAH, 239EH, 2391H, 2385H |
| DW   | 2378H, 236CH, 235FH, 2353H, 2346H |
| DW   | 233AH, 232DH, 2321H, 2314H, 2308H |
| DW   | 22F8H, 22E2H, 22D6H, 22C9H |
| DW   | 22BDH, 22B0H, 22A3H, 2297H, 228AH |
| DW   | 227EH, 2271H, 2265H, 2258H, 224CH |
| DW   | 223FH, 2233H, 2226H, 2219H, 220DH |
| DW   | 2200H, 21F4H, 21E7H, 21DBH, 21CEH |
| DW   | 21C1H |
BIBLIOGRAPHY


Cullen, J. K., Houtsma, A. J. M., Collier, R. (1992, February) Discrimination of brief tone-glides with high rates of frequency-change. Poster contributed at the meeting of Association for Research in Otolaryngology, St. Petersburg, FL.


Zwicker, E. (1956). Die elementaren grundlagen zur 
bestimmung der informationskapazitat des gehors. 
Acustica, 6, 365-381.

consequences of the ear's frequency analysis, in 
Frequency Analysis and Periodicity Detection in 
Hearing, edited by R. Plomp and G. F. Smoorenberg 
(Sijthoff, Leiden), pp. 376-396.