Published in final edited form as:

J Acoust Soc Am. 2011 January ; 129(1): 1–4. doi:10.1121/1.3518745.

Frequency discrimination duration effects for Huggins pitch and narrowband noise (L)

Christopher J. Plack^{a)}

Human Communication and Deafness Division, University of Manchester, Manchester, M13 9PL, United Kingdom

Martine Turgeon and Stuart Lancaster

Department of Psychology, Lancaster University, Lancaster, LA1 4YF, United Kingdom

Robert P. Carlyon and Hedwig E. Gockel

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Abstract

Frequency difference limens (FDLs) were measured for Huggins pitch (HP) stimuli, consisting of a 30-Hz wide band of interaurally decorrelated noise in a diotic low-pass noise and for 30-Hz wide bands of diotic narrowband noise presented in a diotic low-pass noise background. FDLs at a 400-ms duration for the two stimulus types were equated by adjusting the level of the narrowband noise relative to the background. The effects of duration on the FDLs were then measured for center frequencies of 300, 600, and 900 Hz. Although the results were compromised by floor effects at 900 Hz, at 300 and 600 Hz, the duration effects were very similar for the HP and narrowband noise stimuli, with a large improvement in performance between 100 and 400 ms. In contrast to previous results for pure tones, the effect of duration was independent of frequency. The results suggest that: (1) Binaural and monaural pitches may be processed using a common mechanism; (2) discrimination performance for HP and low-sensation-level narrowband noise stimuli is not determined by the number of waveform periods.

I. INTRODUCTION

Huggins pitch (HP) is produced by presenting the same wideband noise to both ears, except for a narrow frequency band that is interaurally decorrelated (Cramer and Huggins, 1958). HP stimuli have a clearly musical pitch corresponding to the center frequency of the band (Akeroyd *et al.*, 2001). Frequency discrimination for HP stimuli is best between 200 and 800 Hz and declines beyond this range (Hartmann and Zhang, 2003; Santurette and Dau, 2007).

Information about the periodicity of sounds may be represented either by the place of maximum excitation in the cochlea or by a temporal code based on neural phase locking. Several results suggest that the effect of duration on the frequency difference limen (FDL) for pure tones increases greatly as frequency is decreased from 2000 to 250 Hz (Moore, 1973b; Micheyl *et al.*, 1998), although other results at 500 and 2000 Hz suggest a weaker interaction of frequency and duration (Hall and Wood, 1984; Freyman and Nelson, 1986). Moore argued that the duration effect is inconsistent with a place mechanism for frequencies

below 4 kHz. Instead, the duration effect may be a consequence of the frequency of these tones being encoded temporally and of the temporal analysis requiring a minimum number of periods to make an accurate frequency estimate (Plack and Carlyon, 1995; Micheyl *et al.*, 1998).

HP stimuli provide an interesting test case for this account of pitch processing. Although there is phase locking in the auditory nerve to the activity at each place in the cochlea excited by the noise, there is no monaural information (temporal or otherwise) concerning the frequency of the decorrelated band. However, the superior olivary complex (SOC) in the brainstem combines the temporal information from the two ears (Yin and Kuwada, 2010), and the center frequency of the HP band could be represented temporally (by phase locking) *after* extraction by the binaural mechanisms in the SOC (Akeroyd and Summerfield, 1999). It has been argued that FDLs for monaural narrowband noise (NB) stimuli (similar to the output of a binaural equalization–cancellation process applied to HP stimuli) show variations with bandwidth consistent with a temporal code (Moore, 1973a). If HP is coded temporally, and if the use of temporal information is limited by the number of periods of the waveform, then we might expect frequency-dependent duration effects that are similar to those for pure tones. Such a finding would be consistent with the notion that pitch stimuli are processed using a common temporal mechanism. Conversely, the phenomenon of “binaural sluggishness” (Grantham and Wightman, 1978) suggests that the processing of interaural decorrelation requires a long time window with an “equivalent rectangular duration” of about 100 ms (Culling and Summerfield, 1998). We might expect, therefore, that the pitch salience of HP stimuli would decrease dramatically at short durations compared to that for NB, since the binaural mechanism would not have enough information to process the decorrelated band optimally.

The present experiment examined these issues by studying the effects of duration on FDLs for HP and a salience-matched NB (the FDL was assumed to be a measure of pitch salience). Two (non-competing) hypotheses were tested:

- (1) Binaural and monaural pitches are processed using the same mechanism. In this case, we expected similar effects of duration for HP and NB stimuli.
- (2) Frequency discrimination for low-frequency stimuli is limited by the number of periods of the waveform. In this case, we expected a strong interaction of frequency and duration on the FDL.

II. METHOD

The HP stimuli consisted of a Gaussian noise with a spectrum level of 40 dB. The noise was low-pass filtered at 2 kHz and presented diotically, except for a 30-Hz wide frequency region in which a π phase shift was generated between the ears. The NB stimuli consisted of a diotic 40-dB spectrum level Gaussian noise low-pass filtered at 2 kHz, to which was added an independent 30-Hz wide band of diotic rectangular NB. The nominal center frequency of the NB and HP band was 300, 600, or 900 Hz. The results of a pilot experiment at 300 Hz using stimulus durations of 25, 50, 100, 200, and 400 ms (described below) demonstrated that both the HP and salience-matched NB stimuli produced very high FDLs when the duration was less than 100 ms. Hence, only durations of 100, 200, and 400 ms were tested in the main experiment. Durations included 10-ms raised-cosine onset and offset ramps. All stimuli were presented *via* Sennheiser HD580 circumaural headphones in a double-walled sound-attenuating booth. Lights on a computer monitor located outside the booth flashed on and off concurrently with each stimulus presentation and provided feedback at the end of each trial.

The main experiment was conducted in two phases. In phase 1, we determined, for each frequency, the level of the NB relative to the background that produced a similar FDL to that for the HP stimulus. A 400-ms duration was used. The spectrum level of the NB was either 3, 6, or 9 dB greater than the spectrum level of the background low-pass noise. For each listener and each frequency, the base-ten logarithm of the FDL was plotted against the relative level of the NB and fitted with a linear least-squares regression line. The intersection of this line with the FDL for the HP stimulus was used to read off the relative level of the NB that would produce an approximately equal FDL to that for the HP stimulus. In phase 2, the NB levels were fixed, individually for each listener, at the levels determined in phase 1. For both the HP stimuli and the NB stimuli, FDLs were measured at each of the three durations at each frequency.

FDLs were measured using a three-interval, three-alternative forced-choice task. The interstimulus interval was 500 ms. Two intervals contained the standard stimulus with a 300, 600, or 900-Hz center frequency. One interval (chosen at random) contained a comparison stimulus with a higher center frequency, which the listener was asked to select by pressing a key on a computer keyboard. The percentage difference between the frequencies of standard and comparison was varied adaptively using a geometric track with a two-down, one-up rule. The maximum allowed difference was 200%. If the track reached 200%, incorrect responses did not result in an increase in the percentage difference. A block of trials consisted of 16 reversals (changes in track direction). The step size was a factor of 2 for the first 4 reversals and a factor of 1.414 for the remaining 12 reversals. For each block, the FDL was taken as the geometric mean of the frequency difference at the last 12 reversals. Within a set of blocks, each condition was presented once in a random order. Four blocks were completed for each condition, and the geometric mean FDL was taken as the final estimate, with the proviso that any FDL more than two standard deviations from the mean was not included in the final mean. Only one block was discarded in phase 1 (for listener 4, 300-Hz HP). Ten blocks were discarded in phase 2, all for listener 3 (no more than one block per condition).

Prior to the main experiment, a pilot experiment was conducted using 300-Hz HP and NB stimuli to determine the range of durations to use in the main experiment. The pilot experiment also consisted of two phases and was conducted in the same way as the main experiment, except that only two blocks were completed per condition, and the relative level of the NB for use in phase 2 was calculated from the mean results from phase 1. The spectrum level for the NB stimuli was set at 7 dB above the level of the background for all listeners for phase 2 of the pilot experiment.

Four normal-hearing listeners (aged 20–24 yr) were tested in the main experiment, and five different normal-hearing listeners (aged 19–22 yr) were tested in the pilot experiment. All listeners were given at least 1 h of training on the conditions before data collection began.

III. RESULTS AND ANALYSIS

The geometric mean data from phase 2 of the pilot experiment are shown in Fig. 1(A). The FDL for both HP and NB stimuli decreased with increasing duration. Although there appears to be a difference between NB and HP stimuli, particularly at durations of 100 and 200 ms, the variability was high for the NB stimuli, and the effect was not consistent across listeners. A two-way analysis of variance (ANOVA) (stimulus type $\times$ duration) conducted on the logarithms of the individual results revealed a highly significant effect of duration [$F(4,16) = 40.7$, $p < 0.00005$] but no significant effect of stimulus type and no significant interaction. For all listeners at 25 and 50 ms for both the HP and NB stimuli and for three listeners at 100 ms for the NB stimuli, the track reached the maximum allowed difference of

200% in one or both blocks, and the performance was close to floor in these cases. For this reason, only durations of 100, 200, and 400 ms were tested in phase 2 of the main experiment.

The geometric mean data from phase 1 of the main experiment are shown in Fig. 1(B). Thresholds are higher than those from some other studies (Hartmann, 1993; Santurette and Dau, 2007). For example, Santurette and Dau (2007) reported a mean FDL of 2.3% for a 500-Hz, 500-ms HP stimulus. The discrepancy might be explained by the larger decorrelation bandwidth (16%) used by these authors. While the present HP results at 300 and 600 Hz are similar, the 900-Hz center frequency produced a FDL about five times greater. Santurette and Dau (2007) also reported a reduction in HP discrimination performance for frequencies above ~800 Hz. The results for the NB stimuli show less variation across frequency, although, as expected, FDLs increase for low levels of the NB relative to the background. For each listener, the relative levels of the NB that produced similar FDLs to the HP stimuli were estimated using the procedure described in Sec. II. Across listener, the relative spectrum levels used in phase 2 were 6.2–8.5 dB at 300 Hz, 6.2–7.4 dB at 600 Hz, and 3.4–4.4 dB at 900 Hz.

The geometric mean results of phase 2 of the main experiment are shown in Fig. 1(C). At 300 and 600 Hz and for both HP and NB stimuli, a large effect of duration was observed, with performance deteriorating at short durations. The pattern of results at 300 and 600 Hz is very similar for the HP and NB stimuli, and the FDLs are similar, showing that they were well matched in salience. In contrast, the matching procedure does not appear to have been successful at 900 Hz. The mean FDL for the HP stimuli is lower than that for the NB stimuli at 400 ms. Comparison with the results of phase 1 [Fig. 1(B)] reveals that this difference was largely due to an improvement in performance for the HP stimulus, presumably a practice effect. (It is unclear why practice effects were not observed at the other frequencies and for the NB stimuli at 900 Hz, but it could be related to the very high thresholds for the HP stimuli at 900 Hz in phase 1, suggesting that listeners found it difficult initially to identify the pitch cue.) In addition, when the FDL exceeded 122%, the NB center frequency was greater than the cutoff frequency of the background (2 kHz), presumably leading to an increase in salience. This was likely to have limited FDLs for the 900-Hz NB conditions, artifactually reducing the size of the duration effect. For this reason and on account of the failure to produce a salience match to the (low salience) HP stimulus, the 900-Hz results were not included in the main analysis.

All analyses were conducted on the logarithms of the individual results of phase 2. A three-way ANOVA (stimulus type $\times$ frequency $\times$ duration) was conducted using the results at 300 and 600 Hz. There was no significant main effect of stimulus type [$F(1,3) = 1.02$, $p = 0.39$] nor of frequency [$F(1,3) = 0.99$, $p = 0.39$], but the effect of duration was highly significant [$F(2,6) = 112$, $p < 0.00005$]. None of the interactions was significant. A separate two-way ANOVA was conducted on the HP results at all the three frequencies. There was a highly significant effect of frequency [$F(2,6) = 36.3$, $p < 0.0005$] and of duration [$F(2,6) = 91.3$, and $p < 0.00005$]. Again, the interaction between frequency duration was not significant.

Two paired-samples t -tests were used to compare the ratio between the FDL at 400 and 200 ms at 300 Hz to the ratio between the FDL at 200 and 100 ms at 600 Hz, separately for the HP and the NB stimuli. This tests whether the change in the FDL was the same for equal changes in number of cycles. For the HP stimuli, the geometric mean ratios were 1.68 at 300 Hz and 4.07 at 600 Hz. For the NB stimuli, the geometric mean ratios were 1.74 at 300 Hz and 3.60 at 600 Hz. The t -test was significant for the HP stimuli ($p = 0.0016$) but was not significant for the NB stimuli. Hence, for the HP stimuli at least, the duration effect was not determined by the number of waveform periods.

IV. DISCUSSION

The pilot experiment suggests that frequency discrimination for HP and matched NB stimuli is very poor at durations of 50 ms or less. Subjectively, the pitch percept is weak to non-existent at these durations. Although there are differences between the mean HP and NB thresholds in Fig. 1(A), this was not consistent across listeners, and the main effect of stimulus type was not statistically significant. The variability may reflect in part the limited data collected in the pilot experiment and the use of a single level for the NB stimuli in phase 2 rather than one chosen on an individual basis.

There are two principal findings from the main experiment. First, duration effects at 300 and 600 Hz are similar for HP and NB stimuli. The results are consistent with hypothesis 1 that the pitches of monaural and binaural narrow-band stimuli are processed in a similar way, presumably by a mechanism at or above the level of the SOC (see also Gockel *et al.*, 2009). The results are also consistent with the triplex model proposed by Licklider (1959) in which the temporal integrator occurs after binaural and pitch processing, rather than before pitch (and after binaural) processing as suggested by Krumbholz *et al.* (2009). Second, there is no interaction between frequency and duration in the FDLs for the HP and NB stimuli. The present data do not support hypothesis 2 that pitch processing for low-frequency stimuli *in general* is limited by the number of waveform periods (Plack and Carlyon, 1995; Micheyl *et al.*, 1998).

However, there are two notes of caution. The duration effects shown here for both NB and HP stimuli are larger than for pure tones with similar frequencies (Moore, 1973b; Freyman and Nelson, 1986). It is possible that the effects of binaural sluggishness for the HP stimuli were obscured by the pitch-related duration effect. Another possibility is that performance was influenced by detectability. Patterson (1976) reported a detection threshold of 56 dB SPL for a 400-ms, 500-Hz pure tone in a 40-dB spectrum level noise. The average NB level at 600 Hz in phase 2 of the present experiment was 21.6 dB greater than the spectrum level of the noise, which should have been clearly audible. However, thresholds for the detection of pure tones decrease at approximately 10-dB/decade decrease in duration out to at least 500 ms at 250 Hz (Florentine *et al.*, 1988). At the shortest duration, therefore, the NB may have been close to the detection threshold, and this could have limited performance. Henning and Wartini (1990) measured the effect of duration on the frequency discrimination for 250-Hz pure tones in continuous noise, in which the sensation level was *fixed* at 7 dB across durations. Interestingly, they included a dichotic condition in which the signal was inverted between the two ears. The duration effect was broadly similar for diotic and dichotic stimulation and similar to that for pure tones at high sensation levels, although the variability was large. In addition, the dichotic condition they used is not equivalent to the HP stimuli used in the present study, not least because the noise used by Henning and Wartini was continuous. Our data on frequency discrimination for HP do not address the possible effects of sensation level nor the variation in sensation level with duration; these questions remain for future research.

Acknowledgments

The research was supported by EPSRC (United Kingdom) Grant No. EP/D501571 and Wellcome Trust Grant No. 088263. We thank the two anonymous reviewers for constructive comments on earlier versions of the manuscript.

References

- Akeroyd MA, Moore BCJ, Moore GA. Melody recognition using three types of dichotic-pitch stimulus. *J. Acoust. Soc. Am.* 2001; 110:1498–1504. [PubMed: 11572360]

- Akeroyd MA, Summerfield AQ. A fully temporal account of the perception of dichotic pitches. *British Society of Audiology Short Papers Meeting on Experimental Studies of Hearing and Deafness* (London). 1999:106–107.
- Cramer EM, Huggins WH. Creation of pitch through binaural interaction. *J. Acoust. Soc. Am.* 1958; 30:413–417.
- Culling J, Summerfield AQ. Measurements of the binaural temporal window using a detection task. *J. Acoust. Soc. Am.* 1998; 103:3540–3553.
- Florentine M, Fastl H, Buus S. Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. *J. Acoust. Soc. Am.* 1988; 84:195–203. [PubMed: 3411048]
- Freyman RL, Nelson DA. Frequency discrimination as a function of tonal duration and excitation-pattern slopes in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 1986; 79:1034–1044. [PubMed: 3700858]
- Gockel HE, Carlyon RP, Plack CJ. Pitch discrimination interference between binaural and monaural or diotic pitches. *J. Acoust. Soc. Am.* 2009; 126:281–290. [PubMed: 19603884]
- Grantham W, Wightman FL. Detectability of varying interaural temporal differences. *J. Acoust. Soc. Am.* 1978; 63:511–523. [PubMed: 670548]
- Hall JW, Wood EJ. Stimulus duration and frequency discrimination for normal-hearing and hearing-impaired subjects. *J. Speech Hear. Res.* 1984; 27:252–256. [PubMed: 6738037]
- Hartmann WM. On the origin of the enlarged melodic octave. *J. Acoust. Soc. Am.* 1993; 93:3400–3409. [PubMed: 8326066]
- Hartmann WM, Zhang PX. Binaural models and the strength of dichotic pitches. *J. Acoust. Soc. Am.* 2003; 114:3317–3326. [PubMed: 14714812]
- Henning GB, Wartini S. The effect of signal duration on frequency discrimination at low signal-to-noise ratios in different conditions of interaural phase. *Hear. Res.* 1990; 48:201–208. [PubMed: 2272929]
- Krumbholz K, Magezi DA, Moore RC, Patterson RD. Binaural sluggishness precludes temporal pitch processing based on envelope cues in conditions of binaural unmasking. *J. Acoust. Soc. Am.* 2009; 125:1067–1074. [PubMed: 19206881]
- Licklider, JCR. Three auditory theories. In: Koch, S., editor. *Psychology: A Study of Science*. McGraw-Hill; New York: 1959. p. 41–144.
- Micheyl C, Moore BCJ, Carlyon RP. The role of excitation-pattern cues and temporal cues in the frequency and modulation-rate discrimination of amplitude-modulated tones. *J. Acoust. Soc. Am.* 1998; 104:1039–1050. [PubMed: 9714923]
- Moore BCJ. Frequency difference limens for narrow bands of noise. *J. Acoust. Soc. Am.* 1973a; 54:888–896. [PubMed: 4757459]
- Moore BCJ. Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.* 1973b; 54:610–619. [PubMed: 4754385]
- Patterson RD. Auditory filter shapes derived with noise stimuli. *J. Acoust. Soc. Am.* 1976; 59:640–654. [PubMed: 1254791]
- Plack CJ, Carlyon RP. Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics. *J. Acoust. Soc. Am.* 1995; 98:1355–1364.
- Santurette S, Dau T. Binaural pitch perception in normal-hearing and hearing-impaired listeners. *Hear. Res.* 2007; 223:29–47. [PubMed: 17107767]
- Yin, TCT.; Kuwada, S. Binaural localization cues. In: Rees, A.; Palmer, AR., editors. *The Auditory Brain*. Oxford University Press; Oxford, United Kingdom: 2010. p. 271–302.

FIG. 1.

(A) The geometric mean results of phase 2 of the pilot experiment, showing the variation of the FDL with stimulus duration at 300 Hz. Error bars show standard errors across listeners. (B) The geometric mean results of phase 1 of the main experiment. FDLs for the 400-ms NB stimuli are plotted against the relative level of the NB. The symbols have been displaced slightly for clarity. The horizontal lines show FDLs for the HP stimuli. (C) The geometric mean results of phase 2 of the main experiment. Error bars show normalized standard errors, calculated after dividing the individual log FDLs by the geometric mean values for each set of three durations for each listener. This removes the variance due to overall listener performance.