Published in final edited form as:

J Acoust Soc Am. 2005 January ; 117(1): 59–62. doi:10.1121/1.1830671.

Discrimination of first- and second-order regular intervals from random intervals as a function of high-pass filter cutoff frequency (L)

William A. Yost, Dan Mapes-Riordan^a), Raymond Dye, Stanley Sheft, and William Shofner
Parmly Hearing Institute, Loyola University Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

Abstract

This study extends the work of Kaernbach and Demany [J. Acoust. Soc. Am. **104**, 2998–2306 (1998)] in which regular interval stimuli (RIS) click trains with first-order intervals could be discriminated from random-interval click trains, but RIS with second-order intervals could not. Kaernbach and Demany concluded that their results cast doubt on autocorrelation as a method of analysis for such stimuli. The present study investigated the same stimuli, but for a variety of filter conditions. The results suggest that while RIS click trains with first-order intervals are more easily discriminated from random-interval stimuli than second-order interval RIS click trains, discrimination based on second-order intervals is possible except when the stimuli are high-pass filtered above 8 kHz, i.e., above the spectral region of phase locking.

I. Introduction

Kaernbach and Demany (1998) originally, and then Kaernbach and Bering (2001) and Pressnitzer *et al.* (2002, 2004), generated a set of regular interval stimuli (RIS) consisting of click trains such that the interclick interval (ICI) of the click trains contained either a constant interval between successive pulses in the click trains (first-order intervals) or constant intervals between every other pulse (second-order intervals). The long-term spectra and autocorrelation functions are similar for both first-order and second-order interval RIS click trains, especially in regard to the autocorrelation features that have been used to account for the perception of RIS sounds (Yost, 1996, 1997). By high-pass filtering the RIS click trains above 6 kHz, Kaernbach and Demany (1998) eliminated the ability of listeners to use resolved harmonics, forcing the listeners to use temporal cues such as could be modeled by autocorrelation. Their data showed that only first-order RIS click trains could be processed by the listeners. They argued that these results were inconsistent with an autocorrelation approach for processing such stimuli. Pressnitzer *et al.* (2002, 2004) showed that RIS click trains very similar to the ones used by Kaernbach and Demany (1998), high-pass filtered at 3 kHz with second-order intervals, appear to produce a pitch, and the pitch appears to be associated with the most common interspike interval of single units in the ventral cochlear nucleus. While their data do not fully support an autocorrelation approach for processing these RIS click stimuli, they did show in their 2004 paper that second-order interval stimuli can produce a pitch.

What has not been determined is how the discriminability of these RIS click trains, as originally studied by Kaernbach and Demany (1998), varies as a function of filtering. Pressnitzer *et al.* (2002, 2004) implied that the pitch strength of the second-order interval click trains was less

a)wyost@luc.edu.

salient than that of the first-order interval click trains. However, they did not measure pitch strength. Discrimination experiments like some of those performed by Kaernbach and Demany (1998) and Kaernbach and Bering (2001) have been used to indirectly measure pitch strength (Yost, 1997). The 6-kHz high-pass filter used by Kaernbach and Demany (1998) places the temporal information in RIS click trains above the spectral region where phase locking is assumed to operate. Kaernbach and Bering (2001) had one condition in which the click-train stimuli were high-pass filtered at 2 kHz. In this condition the second-order interval stimuli were discriminable from their noise foils at only about 64%, while the first-order intervals were discriminable at about 95%. The present study used similar conditions to those employed by Kaernbach and Demany (1998) to investigate RIS click-train discrimination in several spectral regions where the spectral ripples of the RIS click trains are likely unresolved, but where phase locking is probable (i.e., below approximately 6 kHz; see for instance Johnson, 1980). As such, the RIS click-train discrimination experiments can determine if the pitch strength of second-order interval click trains is different from that of first-order interval click trains in spectral regions in which phase locking is possible.

II. Stimuli

Click-train RIS sounds are a sequence of ICI values, each value representing the duration between the prior and current click. Random-interval click trains were created using the same method employed by Kaernbach and Demany (1998) by randomly assigning each ICI a duration from a uniform distribution between zero and twice the mean ICI (d ms) with a resolution of 25μ s. No more than two consecutive ICIs could be larger or smaller than d , in order to prevent the occurrence of long sequences of short or long duration ICIs which could result in a timbre cue.

First-order ICI sequences were created from random ICI sequences by replacing every third random ICI with a regular interval equal to d . These ICI sequences have been termed “kxx” by Kaernbach and Demany (1998), where k represents the regular interval of duration d and x represents the random intervals.

Second-order ICI sequences were created from random ICI sequences by replacing the first two of every three random ICIs with ICIs whose sum is $2d$. The duration of the first of the three intervals, denoted a , was randomly drawn from a uniform distribution between zero and $2d$. The duration of the second interval, denoted b , was equal to $2d - a$. Thus, the sequences contain second-order regular intervals of duration $a + b = 2d$. These sequences have been termed “abx” by Kaernbach and Demany. Figure 1 provides a schematic diagram of a random-order interval click train and its kxx and abx versions. Both the kxx and abx stimuli produce a very similar pitch sensation when the pitch can be detected, and the sensation has been labeled “rattle” pitch by Kaernbach and Demany (1998).

The kxx and abx stimuli have similar, but not identical, spectra and autocorrelation functions. The spectra have spectral peaks space at $1/d$ (for kxx) and $1/2d$ (for abx). The normalized autocorrelation functions have a major peak (AC1) at d (for kxx) or $2d$ (for abx), and the height of AC1 is 0.33 for both kxx and abx stimuli. These spectral and autocorrelation function features are those used by most spectral and temporal models of pitch processing. Autocorrelation models (see Yost, 1996, 1997) predict that the pitch of the RIS click trains is associated with the reciprocal of d or $2d$, and the strength of the pitch is determined by the height of AC1. These models would predict that the pitch strength of the kxx and abx RIS click trains would be the same. Pressnitzer *et al.* (2004) point out that the autocorrelation functions of their random-interval click trains also have a very small broad peak near lag d , and their data suggest that their random-interval click trains have a very weak pitch.

The amplitudes of the clicks were identical and were adjusted to create a click train with a spectrum level of 30 dB. Each stimulus was temporally windowed using a 5-ms squared cosine function. The duration of d was 2, 4, or 8 ms for kxx stimuli and 1, 2, and 4 ms for abx stimuli, resulting in fundamental frequencies, f_0 , of 500, 250, or 125 Hz (f_0 is the reciprocal of d for kxx stimuli and the reciprocal of $2d$ for abx stimuli). The total stimulus duration was kept constant at 480 ms. Stimuli were presented in a wideband condition (stimuli low-pass filtered at 15 kHz) and in several high-pass-filter conditions. The high-pass-filter cutoff frequencies were 8 and 16 times f_0 (all stimuli were low-pass filter at 15 kHz). In the high-pass-filter conditions, a low-pass-filtered Gaussian noise was added to mask distortion components that could be used as discrimination cues. The low-pass-filter cutoff frequency was one-half octave below the high-pass cutoff frequency applied to the click train. The spectrum level of the masking noise was 20 dB; 10 dB below the spectrum level of the click-train stimuli. The relationship between interval order, f_0 , d , and high-pass cutoff frequency is summarized in Table I.

III. Procedure and Listeners

Discrimination performance was measured using a cued, two-alternative, forced-choice (2AFC) procedure. Each trial consisted of a random-interval click train, followed by two other stimuli presented in random order, a random-interval click train and one of the regular-interval click trains. A 300-ms silent gap was inserted between each stimulus. Listeners were asked whether the second or third stimulus was different from the other two stimuli (i.e., to pick the regular-interval click train). Eighteen conditions were tested. A block of trials consisted of the three filter conditions (wideband, $8f_0$, $16f_0$) at a single order (first-kxx or second-abx) and at a single fundamental frequency (125, 250, or 500 Hz), that is, one row in Table I. For each f_0 and interval order, 40, wideband trials were followed by 40, $8f_0$ trials, which were then followed by 40, $16f_0$ trials for a total of 120 trials per block. This was done to provide listeners a clear impression of the pitch in the initial wideband stimulus trials in an attempt to maximize performance in the high-pass-filtered conditions. Not counting practice blocks, a block of 120 trials was repeated ten times throughout the course of the experiment, resulting in 400 trials per condition for each listener.

Four normal-hearing listeners participated in this experiment, including one of the authors (DM). The first block of trials for each listener was considered practice and was not included in the final results.

IV. Results

The mean results across the four listeners are shown in Fig. 2 as percent-correct responses per stimulus condition. With a couple of exceptions the mean results are representative of individual listener performance. One listener had difficulty in all high-pass-filter conditions except the first-order, $d=2$ -ms conditions. This poor performance is in contrast to this listener's relatively good wideband performance. Another listener (one of the authors, DM) was the only participant who consistently discriminated [$P(C)>75\%$] all of the high-pass filtered, second-order stimuli except the 8-kHz cutoff frequency condition where his performance was at chance consistent with the performance of the other listeners. As expected, wideband conditions provided an upper bound to discrimination performance with decreasing discrimination performance with increasing high-pass-filter cutoff frequency. Except for the 8-kHz high-pass conditions, the average listener could discriminate the kxx and the abx stimuli from their respective noise foils, but performance for the abx discriminations was always poorer than for the kxx conditions. Also, average performance decreased with decreasing fundamental frequency across all high-pass-filtered conditions.

V. Discussion and Conclusion

High-pass-filtered, second-order regular versus random-interval discrimination is possible, yet is more difficult than first-order discrimination except for the 8-kHz cutoff-frequency condition, where second-order discrimination is impossible. This implies that the pitch strength of first-order interval RIS click trains is more salient than that of second-order interval RIS click trains. While we used the exact method described by Kaernbach and Demany (1998) and Kaernbach and Bering (2001) to generate the click trains, we did not use the exact same psychophysical procedures, nor were all of the stimulus conditions exactly similar. Our procedure and stimulus conditions are most like the one used in experiments 3 in both of the papers by Kaernbach and colleagues. Even so, there are small differences among the studies including stimulus duration; listener training; number of random, x, intervals following the fixed, k, interval; d; and details of the low-pass noise used to mask distortion products.

Our results at 8 kHz are consistent with those of Kaernbach and Demany (1998), but the results at all other cutoff frequencies extend their findings in that performance was above chance for the discrimination of second-order intervals (abx) from their random-interval foils. Our listeners in the kxx conditions did not seem to be as sensitive as the subjects in either of the Kaernbach studies, but were slightly more sensitive in the comparable abx condition in the Kaernbach and Bering (2001) study. These small performance differences are probably accounted for by individual differences and/or by the small procedural and stimulus differences among the studies. The current results suggest that the auditory system may be sensitive to second-order intervals as long as they occur below 6 kHz, where phase locking is possible.

Kaernbach and Demany (1998) used their results as evidence against autocorrelation as a means of explaining auditory temporal processing. To the extent that listeners can process second-order interval statistics, the current results suggest that processes like autocorrelation may be useful for explaining temporal processing phenomena for stimuli that contain temporal information in spectral regions below about 6 kHz, where phase locking occurs. However, autocorrelation functions for the kxx and abx stimuli are not identical in all respects, and there is some weak temporal regularity in the random-interval click trains used in these experiments (Pressnitzer *et al.*, 2004). That is, while the AC1 peak is at d or 2d and its normalized height is one-third (these are the features of autocorrelation used to account for pitch and pitch strength of RIS sounds; Yost, 1996, 1997) for the kxx and abx stimuli, other more subtle features of the functions differ (see Pressnitzer *et al.*, 2004). Thus, the ability for listeners to discriminate abx from the random-interval click trains may be based on some other aspects of the stimulus fine structure than those typically used to account for RIS processing. A hint of this possibility can be seen for the wideband stimuli of Fig. 2. For most conditions performance for the abx stimuli is still poorer than for the kxx stimuli, even when many of the spectral components are resolved. Since these spectral components are almost the same for both kxx and abx stimuli, it might be that whatever temporal feature allows the abx stimuli to be discriminated from the random-interval click trains when the spectral features are unresolved is also used (at least to some extent) when the spectral features are resolved.

The current uses of autocorrelation models predict that the discriminability of the first- and second-order RIS click trains from their random-interval noise foils would be similar, and therefore that they would have similar pitch strengths. Since this did not occur, the current use of autocorrelation cannot be used to completely account for these data. Pressnitzer *et al.* (2004) attempted to account for the perceived pitch of these stimuli, but not pitch strength. It might be possible to use the strength of the interval statistics obtained in their work to derive predictions for pitch strength, but the current data and literature do not provide enough information to make quantitative predictions at this time. Thus, existing explanations of the pitch and pitch strength of RIS stimuli will need modification or new models will be required

in order to account for the results of these RIS click-train experiments, but such models will need to account for the pitch and pitch strength of second-order, as well as first-order interval RIS sounds.

Acknowledgments

This research was supported by a program project grant from the National Institutes of Deafness and other Communication Disorders, and was part of Dan Mapes-Riordan's Ph.D. dissertation.

References

- Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* 1980;68:1115–1122. [PubMed: 7419827]
- Kaernbach C, Bering C. Exploring the temporal mechanisms involved in the pitch of unresolved harmonics. *J Acoust Soc Am* 2001;110:1039–1048. [PubMed: 11519572]
- Kaernbach C, Demany L. Psychophysical evidence against the autocorrelation theory of auditory temporal processing. *J Acoust Soc Am* 1998;104:2298–2306. [PubMed: 10491694]
- Pressnitzer D, de Cheveigné A, Winter IM. Perceptual pitch shift for sounds with similar waveform autocorrelation. *ARLO* 2002;3:1–6.
- Pressnitzer D, de Cheveigné A, Winter IM. Physiological correlates of the perceptual pitch shift for sounds with similar autocorrelation. *ARLO* 2004;5:1–6.
- Yost WA. Pitch of iterated rippled noise. *J Acoust Soc Am* 1996;100:511–519. [PubMed: 8675844]
- Yost WA. Pitch strength of iterated rippled noise when the pitch is ambiguous. *J Acoust Soc Am* 1997;101:1644–1650. [PubMed: 9069631]

FIG. 1. RIS click trains; top: random intervals (x), middle: first-order intervals (kxx), bottom: second-order intervals (abx), d is mean duration.

FIG. 2.

Mean results (across four listeners) showing percent correct, $P(C)$, discrimination as a function of high-pass filter cutoff (in Hz) for random versus first-order (kxx) and second-order (abx) RIS click-train discrimination; bars are mean percent correct; error bars are standard errors. $P(C) = 50\%$ is chance performance.

Table 1

Summary of test conditions.

Interval order	Fundamental frequency, f_0 (Hz)	Mean interval duration, d (ms)	High-pass cutoff frequency (Hz)		
			Wideband	8f ₀	16f ₀
1-kxx	125	8	...	1000	2000
1-kxx	250	4	...	2000	4000
1-kxx	500	2	...	4000	8000
2-abx	125	4	...	1000	2000
2-abx	250	2	...	2000	4000
2-abx	500	1	...	4000	8000