ACUSTICA
S. HIRZEL VERLAG · STUTTGART
Vol. 19 1967/68 No. 2

Repetition Pitch Mediated by Temporal Fine Structure at Dominant Spectral Regions

In a previous paper (BILSEN [1]) one of the authors described the subjective pitch effect due to the interaction of a sound with its repetition, called Repetition Pitch (RP). When a sound and the repetition of the same sound after a delay τ are presented together, a pitch is evoked corresponding to the reciprocal value $1/\tau$. This may be expressed in the following notation: $RP_0^0 = 1/\tau$. The sound may be music, noise, or merely a pulse (repeated periodically or at random).

However, after adding the repetition with negative sign (in fact this means that all the frequency components of the repeated sound have been shifted in phase by 180°) a bivalent pitch is perceived, viz. $RP_{180}^0 = 1.14/\tau$ and $RP_{180}^0 = 0.88/\tau$; this implies an up- and downward pitch shift of a little more than a full tone, independent of the value of τ. By shifting the phase of all the frequency components of the repeated sound by ±90° pitches are heard lying between the values for RP_0^0 and RP_{180}^0, viz. $RP_9^{00} = 1.07/\tau$ and $RP_{270}^0 = 0.94/\tau$.

So far this pitch behavior could not be explained, simply, in terms of time or frequency analysis. In this letter, however, we will show that an explanation is possible if one assumes:

1. Fine-structure detection. The hypothesis of fine structure detection has been tested extensively for amplitude-modulated sinusoids and bandpass filtered periodic pulse trains (SCHOUTEN, RITSMA, CARDOZO [2], RITSMA, ENGEL [3]). For these types of signals it has been shown that to a first approximation the distance between two positive peaks in the fine structure of the acoustical waveform near two successive crests in the time envelope is an adequate parameter of the pitch perceived.

This type of signals with narrow frequency spectrum produces a so called "ready made" envelope at the corresponding place on the basilar membrane and so we may be sure that the displacement waveform at that place is about the same as the acoustical waveform. Therefore the hypothesis of fine-structure detection can be extended to a relation between the frequency of the low-pitch perceived and the displacement waveform at a specific place along the basilar membrane.

Due to the spectral analysing function of the basilar membrane, signals with a flat frequency spectrum will give displacement waveforms which differ markedly at different places along the membrane. For a place which is most sensitive for the frequency f_0, the displacement waveform is similar to the waveform of the acoustical signal passing a bandpass filter with center frequency f_0.

The response of an ideal bandpass filter with center frequency f_0 and bandwidth Δf to a DIRAC pulse $\delta_0(t)$ can be shown to be equal to

$$\delta_0'(t) = \frac{1}{\pi} \int\limits_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} \cos\omega t \, d\omega = \frac{2}{\pi t} \cos\omega_0 t \sin\frac{\Delta\omega}{2} t . \tag{1}$$

In the same way we find for the response to a 90°-phase shifted DIRAC pulse

$$\delta'_{90}(t) = -\frac{1}{\pi} \int\limits_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} \sin\omega t \, d\omega = -\frac{2}{\pi t} \sin\omega_0 t \sin\frac{\Delta\omega}{2} t . \tag{2}$$

For $t=0$ the envelope $\left(2\sin\frac{\Delta\omega}{2}t\right)\Big/(\pi t)$ is at maximum, thus it is in this vicinity the fine-structure has its greatest peaks. In particular, $\delta_0'(t)$ has its major positive peak for $t=0$, $\delta'_{90}(t)$ for $t=-1/(4f_0)$, $\delta'_{270}(t)$ for $t=1/(4f_0)$ and $\delta'_{180}(t)$ has two major peaks, one for $t=1/(2f_0)$ and the other for $t=-1/(2f_0)$.

Now, if RP, due to the interaction of two subsequent DIRAC pulses (pulse distance τ), is indeed correlated with the reciprocal value of the time distance between the major, positive, peaks of the fine structure, the RP-behavior for that place on the basilar membrane which corresponds to the frequency f_0, must fulfill the following relations

$$[\mathrm{RP}_0^0]_{f_0} = 1/\tau ,$$
$$[\mathrm{RP}_{90}^0]_{f_0} = 1/(\tau - 1/(4f_0)) ,$$
$$[\mathrm{RP}_{180}^0]_{f_0} = 1/(\tau \pm 1/(2f_0)) ,$$
$$[\mathrm{RP}_{270}^0]_{f_0} = 1/(\tau + 1/(4f_0)) .$$

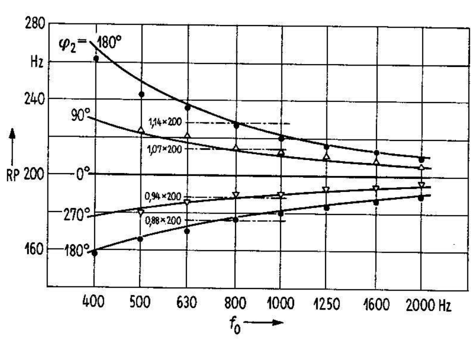

Fig. 1. RP, due to a third-octave bandpass filtered 0°- and φ_2-pulse, as a function of the third-octave bandpass center frequency f_0; $\tau=5$ ms.

These relations are represented in Fig. 1 by the solid lines, for $\tau=5$ ms.

The hypothesis is confirmed by the results of listening tests with third-octave bandpass filtered pulse pairs. Each measured point is the average of ten pitch matchings by two subjects. The test signal and the matching signal were of the same character. The test signal built up of a 0°-pulse and a 90°-, 180°- or 270°-pulse; the matching signal built up of two 0°-pulses. The pulse pair was repeated each 100 ms. (Other tests proved that the way of repeating the pulse pair, periodically or at random, is not of significant influence on the perceived pitch.)

2. Dominance of one spectral region with respect to pitch. This hypothesis has been tested for periodic white pulse trains with equal and with alternating polarity (RITSMA [4]). It has been found that the temporal information taken at a specific place along the membrane is dominant in the perception of pitch over the temporal information present at other places along the membrane. The most effective spectral area influencing pitch is more or less subject bound. In general the most effective spectral areas cover the region of the third, fourth, and fifth harmonics for fundamental frequencies in the range of 100 Hz to 400 Hz.

Now, with our knowledge about both hypotheses in mind, we may direct our attention to the broad band values of RP mentioned at the beginning of this letter. It will be clear from the foregoing about dominant spectral regions that we will find the center frequencies of these regions by substituting the broad band values in eqs. (3). Solving for f_0 and averaging we get the average dominant center frequency

$$[f_0]_{\text{dominant}} = (3.9 \pm 0.2)\,\mathrm{RP}_0^0 . \tag{4}$$

In Fig. 1 the broad band values are denoted by the dotted lines; these lines intersect with the graphs at a center frequency of about 800 Hz, this being four times the value of RP_0^0 (= 200 Hz).

The value of eq. (4) is in full accordance with the results from the test with periodic pulses where the region of the third, fourth, and fifth harmonics proved to be dominant.

Further evidence on this subject will be presented later.

(Received July 7th, 1967.)

F. A. BILSEN
Technological University
Delft, Netherlands

and

R. J. RITSMA
Instituut voor Perceptie Onderzoek
Eindhoven, Netherlands

References

[1] BILSEN, F. A., Repetition Pitch: Monaural interaction of a sound with the repetition of the same, but phase shifted, sound. Acustica **17** [1966], 295.

[2] SCHOUTEN, J. F., RITSMA, R. J., and LOPES CARDOZO, B., Pitch of the Residue. J. Acoust. Soc. Amer. **34** [1962], 1418.

[3] RITSMA, R. J. and ENGEL, F. L., Pitch of Frequency-Modulated Signals. J. Acoust. Soc. Amer. **36** [1964], 1637.

[4] RITSMA, R. J., Frequencies dominant in the perception of the pitch of complex sounds. J. Acoust. Soc. Amer. **42** [1967], July issue (at press).