APU 2275. 88

Timbre cues in monaural phase perception: distinguishing within-channel cues and between-channel cues

Roy D. Patterson

MRC Applied Psychology Unit
15 Chaucer Road, Cambridge, England

In: Basic Issues in Hearing. Proceedings of the 8th International Symposium on Hearing. Duifhuis, H., Horst, J.W. & Wit, H.P. (eds), London: Academic Press Ltd, pp.351-358. (1988).

Introduction

Recently, Patterson (1987a) proposed a model of peripheral auditory processing which suggested that global changes in the phase spectrum of a sound would cause timbre changes, in certain circumstances. Global phase changes shift the output of one auditory filter relative to another in time, and previously these shifts were not thought to be audible (de Boer, 1976). To test the hypothesis, Patterson (1987b) presented listeners with a filtered pulse train containing 31 equal-amplitude harmonics, all in cosine phase. Then he introduced a global phase change and increased its magnitude while checking for timbre changes. The experiments revealed that sensitivity to global phase changes existed over a wide range of stimulus conditions.

In fact, it is not actually possible to generate a global phase change without introducing at least a small, local phase change. Local phase changes alter the envelope of the output of the auditory filter in which they occur, and they are known to cause timbre changes when sufficiently large (Mathes and Miller, 1947). The current paper presents a pair of experiments designed to distinguish local and global phase sensitivity, and in so doing, to provide further evidence for global phase sensitivity.

The role of phase in timbre perception

The timbre experiments reported by Patterson (1987b) were performed with local as well as global phase changes. The operation of the model is illustrated in Fig. 1 using a sound (Fig. 1a) with local phase changes that are just audible when the fundamental is 125 Hz and the stimulus level is about 40 dB/component (dB/c). In this case, the odd harmonics are all in cosine phase and the even harmonics are all shifted 40 degrees from cosine phase. This is referred to as an alternating-phase, or APH, wave and the filtered pulse train is referred to as a cosine-phase, or CPH, wave.

In the first stage of the model, an auditory filterbank is used to simulate the frequency analysis performed by the cochlea. The filterbank converts the sound wave into a set of driving waves (Fig. 1b); in the latest version, the transfer function for the auditory filter is the gammatone suggested by Johannesma (1972). In the second stage, a bank of pulse generators converts the driving waves into streams of pulses that, in essence, record the times of the driving-wave peaks; the pulse generators are based on a hair cell simulation suggested by Meddis (1986). Together, the pulse streams form a pulse ribbon (Fig. 1c) which provides a representation of the overall neural firing pattern that the sound might be expected to produce in the auditory nerve. The vertical dimension of the ribbon is filter centre frequency on an ERB-rate scale (Moore and Glasberg, 1983). For a periodic sound like the APH wave, the pattern repeats on the ribbon and the repetition rate corresponds to the pitch of the sound. Timbre is assumed to correspond to the shape of the pulse pattern within a cycle. Thus, the pulse ribbon provides a

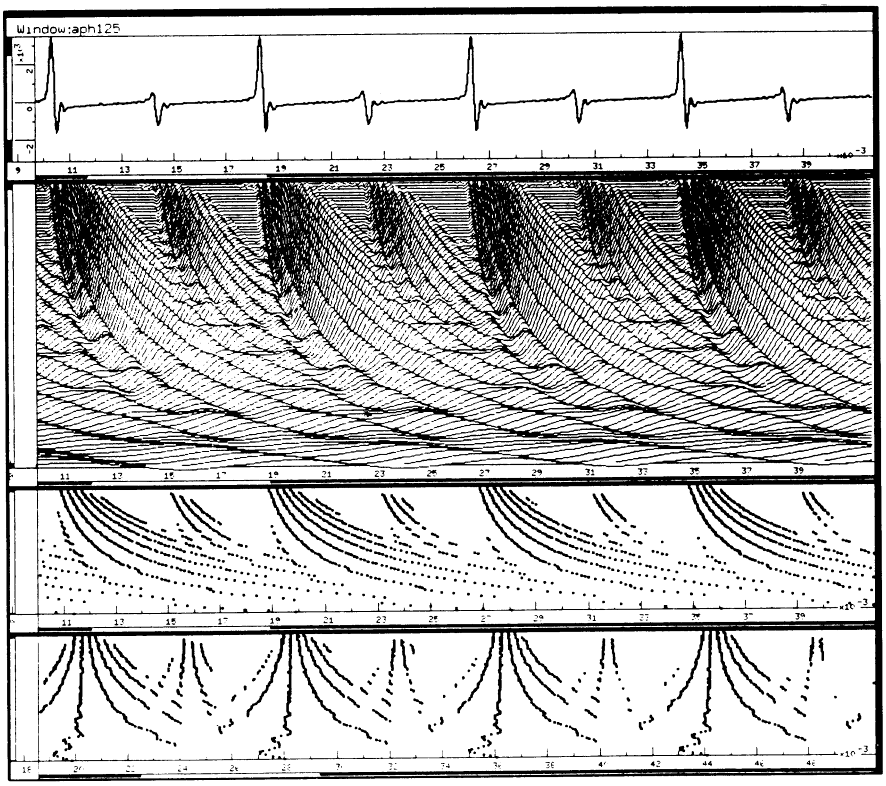

Figure 1. The processing of an APH wave by the pulse ribbon model (D = 40 degrees). The filterbank converts the wave (a) into driving waves (b) which the pulse generators convert into a pulse ribbon, either without (c) or with (d) phase compensation.

convenient means of separating pitch and timbre information, and relating phase manipulations to timbre perceptions in complex sounds.

Timbre cues within filter channels

The APH sound causes secondary maxima to appear in the envelopes of the higher-frequency driving waves midway through the cycle (Fig. 1b) -- maxima that do not occur in the CPH driving waves. Envelope changes of this sort are known to produce timbre changes when they are sufficiently large (Mathes and Miller, 1947). The secondary maxima of the APH driving waves generate a secondary column of pulses midway through the cycle of the pattern that appears on the pulse ribbon (Fig. 1c). In the model, it is these pulses that are assumed to mediate the timbre change. Since it is the interaction of adjacent harmonics within an auditory filter that causes the envelope change, the perceptual feature that accompanies the change is referred to as a within-channel timbre cue. Patterson used the data from the APH experiments to estimate sensitivity to envelope changes within filter channels.

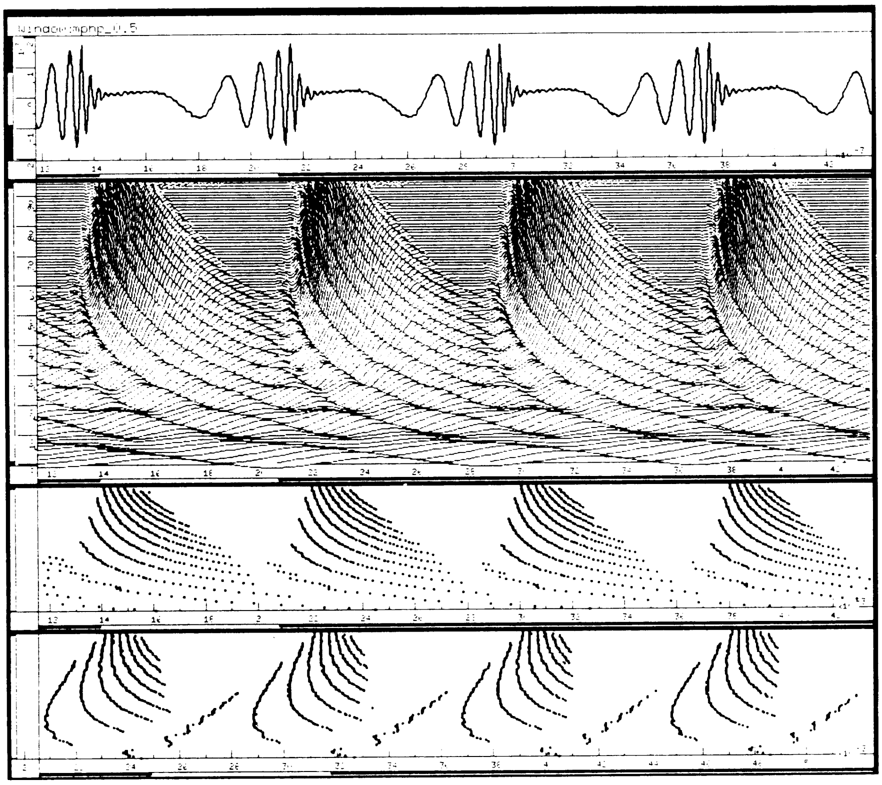

Figure 2. As in Fig. 1, but for an MPH wave (S = 1/2).

Note that the phase lag of the low-frequency channels in Figs. 1b and 1c is not produced by the APH stimulus. It is a property of the filterbank; the low-frequency filters have narrower bandwidths and greater phase lags. Figure 1d presents a version of the pulse ribbon in which the phase lag of the filterbank has been removed and it shows that the envelope maxima of APH driving waves are aligned across channels when measured relative to the phase lag of the filterbank.

Timbre cues between filter channels

Global phase changes were introduced into the CPH wave by applying an acceleration to the phase spectrum. Patterson (1987b) used decelerations as well as accelerations and so the stimuli are referred to as monotonic- phase, or MPH, waves. A decelerating MPH wave is shown in Fig. 2a. Some of the low- and high-frequency energy that would be concentrated in the single pulse of a CPH wave has been separated out in the MPH wave, with the low-frequency energy appearing just before the main pulse and the high-frequency energy appearing just after the main pulse. Figure 2b presents the driving waves of the MPH stimulus, and it shows that they have simple envelopes with no secondary maxima anywhere in the cycle. There is still a phase lag in the low-frequency channels, but a comparison of the magnitude of the time differences in Figs. 1b and 2b reveals that the global phase lag is reduced by the MPH wave. It had generally been assumed that global phase manipulations of this sort would not produce timbre changes (de Boer, 1976; Buunen, 1976).

The primary characteristics of the MPH manipulation are preserved in the initial pulse ribbon (Fig. 2c); there are no pulses midway through the cycle of the pattern and the global phase lag is reduced with respect to that in Fig. 1c. Figure 2d shows the MPH pulse ribbon after compensation for the global phase lag of the filterbank. It shows that the MPH manipulation results in a progressive shift of the envelope maxima relative to filterbank phase. Patterson (1987b) found that, contrary to previous expectations, MPH stimuli do produce timbre changes when compared with CPH stimuli. The MPH stimuli produce only minimal envelope changes, and so he argued that the timbre cue was probably the progressive delay of channels across the pulse ribbon. Since the cue involves a comparison across channels, it is referred to as a between-channel timbre cue.

Differential masking of timbre cues

In the model, the MPH and APH timbre cues differ in terms of the stability of the pulses on which they are based: The MPH cue is based on the most stable pulses in the ribbon - those arising from the largest driving-wave peaks. The cue is assumed to involve a comparison of the most stable pulses in different channels. The APH cue is based on some of the less stable pulses arising from secondary envelope maxima. The cue is assumed to involve a comparison of more and less stable pulses within individual channels. If these assumptions are correct, then we might expect to find that a rising background noise would disrupt the APH/CPH discrimination sooner than the MPH/CPH discrimination, because it would obscure the less stable pulses of the APH/CPH discrimination sooner than the more stable pulses of the MPH/CPH discrimination. A pair of experiments was performed to test this hypothesis.

Method

Four listeners with normal hearing were required to discriminate APH and CPH stimuli, or MPH and CPH stimuli, in the presence of a background noise. The signal level and the magnitude of the phase manipulation were both fixed during a run. The noise level was varied adaptively to determine the level at which the discrimination was just possible; that is, the point where performance was reduced to 71% correct in a two-alternative, forced-choice experiment. Between runs, the magnitude of the phase manipulation was varied to determine whether the MPH/CPH discrimination was, indeed, more resistant to noise masking than the APH/CPH discrimination.

There were two versions of the experiment: To begin with, since the peak factor of the stimulus decreases as the magnitude of the phase manipulation increases, it was important to demonstrate that the individual APH and MPH signals were all equally detectable in the masking noise. The noise masker had a bandwidth in excess of 3.0 kHz, and so in the first experiment, to keep the noise from becoming too loud, the signal level was set at 40 dB/c. The experiment revealed significant differences in the expected direction, but since it did not employ the full range of the phase effect, it did not provide the most sensitive measure of the differences. Accordingly, a second experiment was performed in which the signal level was raised 10 dB. In this case, it was not possible to measure masked threshold for the signals. Aside from the level difference, the two experiments were quite similar.

Stimuli

The signals were flat-spectrum periodic sounds containing 31 harmonics. The fundamental, fo, was 62.5 or 125 Hz in the first experiment and 62.5 Hz in the second. The signals were lowpass filtered at the 25th harmonic and bands of noise with steep cutoffs were used to restrict the listening region to a set of harmonics extending from the 4th to the 24th. A detailed description of the stimuli and the procedure is presented in Patterson (1987b).

The APH manipulation involved shifting every other harmonic away from cosine phase by a fixed number of degrees. The phase difference, D, was varied from 20 degrees -- where it is just large enough to support discrimination in the absence of the masking noise -- up to 80 degrees -- where the secondary maxima in the driving waves are almost as large as the primary maxima. The MPH manipulation involved shifting successive harmonics an ever decreasing amount to produce a phase spectrum in which greater shifts occur between lower harmonics and smaller shifts between higher harmonics. The aim was to produce between-channel phase lags that were roughly con-

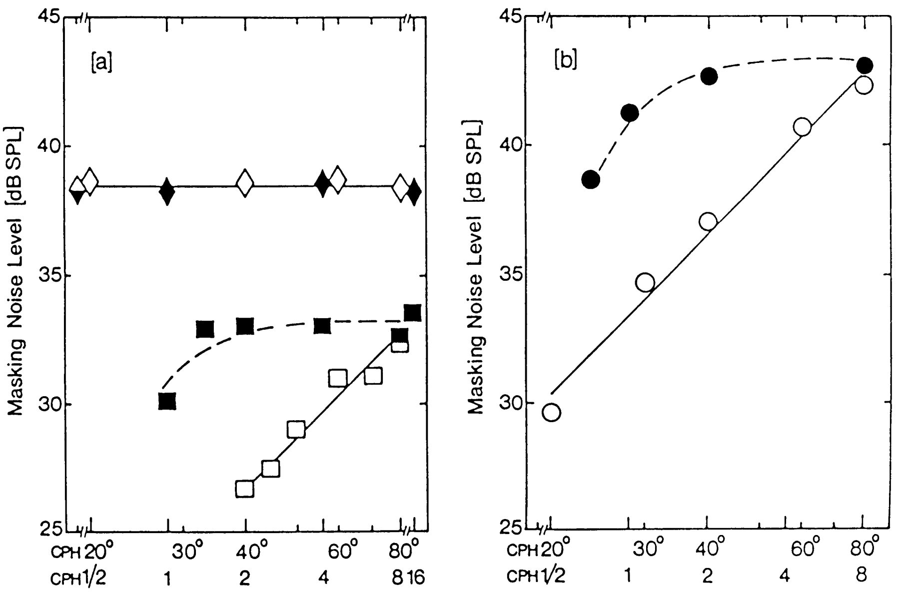

Figure 3. The noise spectrum level required to disrupt timbre discriminations for signals of 40 dB/c (squares) or 50 dB/c (circles); filled and open symbols are for MPH and APH conditions, respectively. The diamonds show masked threshold for 40 dB/c signals.

stant from channel to channel. The initial phase spectrum was created by calculating the number of auditory-filter bandwidths between adjacent harmonics and shifting the phase of the upper harmonic by 180 degrees per bandwidth. Then a scalar, with value S, was applied to the phase spectrum as a whole to produce the MPH manipulation.

The pattern of results was the same for the two fundamentals in the first experiment and so the data were combined. Furthermore, the pattern was quite similar for all four listeners, in both experiments, and so their results were averaged in each case. Each of the means presented in the next section includes the data from at least 32 individual runs.

Results

The average data for experiments one and two are presented in Fig. 3. The detection data (diamonds) in the upper part of Fig. 3a show that the APH and MPH stimuli were equally detectable, and that threshold does not vary with the magnitude of the phase manipulation in either case. In the absence of the noise masker, the MPH/CPH timbre difference is clearly audible provided S is greater than about unity. When the noise is introduced, we find that the discrimination requires a signal-to-noise (S/N) ratio of about 6 dB in the region where S is greater than unity - independent of the value of S (filled squares). Only when the discrimination becomes inherently difficult, does the masking noise have a differential effect on performance. In the absence of noise, at 40 dB/c, the APH/CPH timbre difference is clearly audible when D is greater than about 35 degrees. When the noise is introduced, we find that the APH/CPH discrimination is audible for a S/N ratio of just over 6 dB, provided D is large, say 80° (open squares). But as D is reduced, it soon becomes necessary to reduce the noise level to maintain performance, and when D is 40 degrees, the S/N ratio has to be in excess of 12 dB.

In the second experiment (Fig. 3b), the range of the APH measurements was extended down to 20 degrees by increasing the signal level and restricting fo to 62.5 Hz. The results for D = 20 degrees show that an aspect of the timbre perception that is capable of supporting APH/CPH discrimination in silence, is obscured by a noise whose level is almost 20 dB below that required to mask the presence of the sounds. There is no corresponding effect in the MPH/CPH data. The overall pattern of results is very similar to that found in the first experiment. Indeed, the curves drawn through the MPH/CPH data and the lines drawn through the APH/CPH data are the same in Figs. 3a and 3b, except that they have been shifted up 10 dB in Fig. 3b.

Discussion and Conclusions

The fact that the MPH/CPH discrimination is more resistant to noise masking than the APH/CPH discrimination supports Patterson's (1987a, 1987b) contention that the MPH and APH timbre cues are derived from different parts of the pulse ribbon. In the case of the APH sounds, it is argued that the timbre cue is the column of pulses that appears in the pulse ribbon midway through the cycle of the pattern -- pulses associated with secondary maxima in the envelopes of the driving waves. These maxima are relatively small and this accounts for the fact that the discrimination is disrupted by the introduction of a relatively low-level noise. The MPH cue is assumed to be the slant of the pulse ribbon, derived from pulses associated with the primary maxima in the driving waves. These maxima contain the largest driving-wave peaks and this accounts for the fact that it takes a relatively higher-level noise to disrupt this discrimination.

The contrast between the form of the MPH data and the APH data would appear to rule out the possibility that the MPH timbre cue might be based on small envelope changes occurring in the troughs of the driving waves. The current data do not mean that the MPH timbre cue is necessarily a between-channel cue; it is logically possible for the cue to be a change in the shape of the maximum of the driving-wave envelope. However, as Patterson has pointed out, the timbre cue is available when the shape change is very small, and so it seems more likely that it depends on a between-channel comparison.

Acknowledgements

The author would like to thank Robert Milroy for running the experiments and analysing the results.

References

de Boer, E. (**1976**). "On the residue and auditory pitch perception," in *Handbook of Sensory Physiology, Vol. V*, edited by W.D. Keidel & W.D. Neff (Springer-Verlag, Berlin), pp. 479-583.

Buunen, T.J.F. (**1976**). *On the perception of phase differences in acoustic signals.* Doctoral dissertation. University of Delft, Delft, The Netherlands.

Johannesma, P.I.M. (**1972**). "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in: *Proceedings Symposium on Hearing Theory*, pp. 58-69. IPO, Eindhoven, The Netherlands.

Mathes, R.C. and Miller, R.L. (**1947**). "Phase effects in monaural perception," J.Acoust.Soc.Am. **18**, 780-797.

Meddis, R. (**1986**). "Simulation of mechanical to neural transduction in the auditory receptor," J.Acoust.Soc.Am. **79**, 702-711.

Moore, B.C.J. and Glasberg, B.R. (**1983**). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns", J.Acoust.Soc.Am. **74**, 750-753.

Patterson, R.D. (**1987a**). "A pulse ribbon model of peripheral auditory processing," in *Auditory Processing of Complex Sounds*, edited by W.A. Yost and C.S. Watson (Erlbaum, New Jersey), pp.167-179.

Patterson, R.D. (**1987b**). "A pulse ribbon model of monaural phase perception," J.Acoust.Soc.Am., **82**, 1560-1586.