PHASE AND THE HEARING-IMPAIRED*

Stuart Rosen
Department of Phonetics & Linguistics, University College London,
4 Stephenson Way, London NW1 2HE, England

INTRODUCTION

Although Helmholtz, on the basis of experiments with 8-component harmonic complexes of fundamental frequencies near 119 and 238 Hz, claimed to "have never experienced the slightest difference in the quality of tone" with changes in relative phase among the components (Helmholtz, 1954), more recent studies have modified his conclusions (e.g., Mathes and Miller, 1947; Goldstein, 1967). It is now apparent that the primary determinant of the perceptibility of a given phase change is the frequency spacing between the sound's constituent sinusoidal components. When relative phase changes are made in components that are "close enough" together, they are perceptible; when they are made to widely spaced components, they are not. Phase sensitivity is thus understood to reflect the <u>failure</u> of frequency resolution - only when a sound's constituent sinusoids interact (i.e., lie sufficiently within a single critical band, or auditory filter) will a phase change be detectable. (For a discussion of other factors, see Rosen, 1986).

Especially relevant for estimating the importance of phase on the perception of speech (in particular, for vowel-like sounds) are studies like those of Licklider (1957) and Schroeder (1959), who restricted their attention to harmonic complexes, noting that changes in timbre and pitch were readily produced by phase manipulations. What seemed to have been the final word along these lines was an impressive multi-dimensional scaling study by Plomp and Steeneken (1969), who concluded that the effect of phase on timbre (in the limited sense of the perceptual attribute which distinguishes periodic sounds of identical pitch and loudness), although real, was small compared to the effect of the relative amplitude of the components.

All these studies, though, used normal listeners. There has been almost no investigation of the role of phase in determining the percepts of the hearing-impaired (the notable exceptions being Hoekstra and Ritsma [1977] and Hoekstra [1979]). Given that phase sensitivity is supposed to be constrained by auditory frequency selectivity, and that many impaired listeners have impaired selectivity,

---

*This work has been supported by the Medical Research Council of the U.K. Many thanks to V. Ball, C. Bootle, C.M.Green, V. Hazan, and H. Wall for their extensive participation as listeners, and P. Howell for proofreading.

it seems likely that phase will play a larger role for them than for normal-hearing listeners (Rosen, 1984; Rosen and Fourcin, 1986).

METHODS

Test stimuli were synthesized digitally by a DEC PDP-12 computer running a 10-bit DAC at a sampling frequency of 10 kHz. The phase and amplitude of each stimulus component was corrected (except where noted) for the phase and amplitude distortion produced by the headphones (a Connevans CE8, chosen for its relatively low nonlinear distortion at low frequencies and high levels). This was determined with a small electret microphone mounted on the grid protecting the headphone diaphragm, thus allowing monitoring of the sound pressure while the listener wears the headphones (Dominitz, 1975; Rosen and Nevard, in press). Preliminary measurements on a KEMAR manikin indicate that the sound pressure measured by the headphone-mounted microphone will be within 6 dB and 10° of that at the listener's tympanic membrane at 1.8 kHz (the maximum frequency in the following studies), with improving accuracy as frequency is lowered. The same phase and amplitude corrections (a mean of 8 ears) were applied for all listeners. The headphone output was intermittently monitored, using a real-time spectrum analyzer, to set levels and check the waveform. Typically, amplitudes were within ±1.5 dB, and phase within -5 to +10° of those specified.

All sounds had a steady-state duration of 400 ms, with 50-55 ms raised-cosine rises and decays added. They were presented, after low-pass filtering and amplification, to a single earphone in a sound-treated room. Spurious spectral components in the sounds, measured at the headphones, were at least 40 dB down from the smallest component of the complex. Masking noise, when present, was band-pass (20 Hz to 2-3 kHz) at about 30-35 dB SPL/Hz.

A 3-interval 3-alternative forced-choice (3I-3AFC) task was used for testing the discrimination of phase shifts, while a 2I-2AFC task was used for assessing abilities to discriminate changes in fundamental frequency. Inter-stimulus intervals were about 580 ms. Feedback as to the correctness of response was given. During a particular session, the two sounds whose discriminability was being tested remained constant.

At the start of each 30-trial session, listeners were given the opportunity of unlimited practice with the pair of sounds to be tested. At this time, they effectively controlled the presentation of the stimuli. The experimenter, too, often trained the listeners using this facility. This initial practice was crucial, especially when phase discrimination was being tested. It was frequently reported in such tests that all three sounds were identical at the start of practice, but a difference could often be found after listening for a little while.

HYPER-SENSITIVITY TO PHASE CHANGES IN THE HEARING-IMPAIRED

Rosen (1984) argued that if phase sensitivity reflects the failure of frequency resolution, then at least some impaired listeners (with widened auditory filters and relatively intact temporal analyzing capabilities) should be, under appropriate circumstances, more sensitive to phase changes than normal listeners. This possibility was tested using stimuli that have seen extensive use in phase perception studies, so-called 100%-SAM (for sinusoidally amplitude modulated) and QFM (for quasi- frequency modulated) sounds (Mathes and Miller, 1947). Both sounds have an identical amplitude spectrum: a central sinusoidal component, and two sinusoidal side-bands, 6 dB lower in amplitude than, and equally-spaced in linear frequency from (by an amount given by the modulation rate) the central component. They differ only in their phase spectrum, a 90° change in the central component.

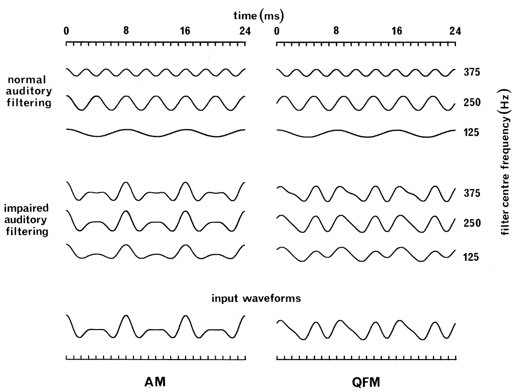

FIGURE 1, The outputs of hypothetical normal and impaired auditory filter banks to two three-component harmonic complexes which differ only in the relative phase of their central component (SAM and QFM sounds with a carrier frequency of 250 Hz and a modulation rate of 125 Hz). The auditory filters are centered at the frequencies of the harmonic components. Normal auditory filtering uses the rounded exponential model and bandwidths given by Moore and Glasberg (1983), while impaired auditory filtering assumes the same trend of bandwidth with frequency, but with absolute values ten times larger than in the normal case. From Rosen and Fourcin (1986).

Figure 1 shows graphically why impaired auditory filtering might lead to better discrimination of SAM from QFM sounds. With normal auditory filtering (when the spectral components for this particular sound are essentially resolved), in order to distinguish SAM from QFM sounds there needs to be some way for comparing the time of events across auditory filters (which available evidence suggests is not possible). In an impaired auditory filter bank, the same phase change is expressed as a within channel change.

Figure 2 shows an instance in which an impaired listener (XG) did, indeed, evidence increased sensitivity to phase changes in a SAM/QFM discrimination task centered at 400 Hz. XG was a young (late twenties), successful hearing-aid user with a relatively flat loss of 30-50 dB across the frequency range 0.125-8 kHz. Her degree of frequency selectivity was assessed at 500 Hz using the "notched-noise" technique of Patterson et al. (1982). The difference between the threshold of a 500 Hz tone obtained in a broadband noise, and one with a "notch" in its spectrum (400 Hz wide) centered linearly on the tone frequency, was determined. The bigger this difference in thresholds is, the narrower are the auditory filters. For the noise level of 60 dB SPL/Hz used, normal listeners obtain about a 20 dB change between the two conditions (Rosen and Stock, in preparation). XG showed only a 10 dB change in threshold, with an estimated equivalent rectangular bandwidth (ERB) approximately twice that of normal listeners. She also showed better than normal performance for SAM/QFM sounds centered at 500 Hz.

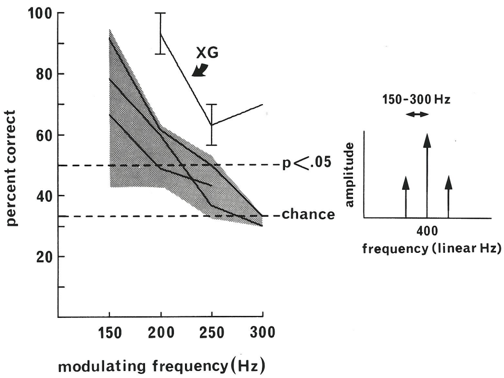

FIGURE 2, The performance of three normal-hearing listeners compared to that of a single hearing-impaired listener in a SAM/QFM discrimination task, all with approximately equal exposure to the task. The carrier frequency of the stimuli was always 400 Hz, with modulation frequency varying from 150 to 300 Hz. The level of the carrier was equal for all listeners, but varied somewhat (from about 93-98 dB SPL) with modulation rate. The inset at right shows the amplitude spectrum for the stimuli. Although the amplitude of the three stimulus components was adjusted to give the correct amplitude relationships at the output of the headphones, all stimuli were in sine phase at their input. Thus the true phase relationships among the sounds varied with modulation rate, although the QFM stimulus always differed from the SAM stimulus by a + 90° phase shift at 400 Hz. Inspection of the output of the headphones to the SAM stimulus complex of 200, 400 and 600 Hz showed that the three components were, coincidentally, very nearly in -cosine phase. The shaded area shows the range of performances obtained across all the normal listeners, while the solid lines within the shaded area show their individual results, averaged across sessions. Note how overall performance decreases with increasing modulation rate. The solid line at the top of the figure shows the mean results obtained from a hearing-impaired listener, while the bars show the range of performance that was exhibited. XG was shown to have a loss of frequency selectivity at 500 Hz.

That the degradation in frequency selectivity is the important factor in accounting for this "hyper-sensitivity" is supported by results from another hearing-impaired listener. His audiogram shows a loss sloping from 15 dB at 125 Hz to 40-50 dB at 1-8 kHz. Even with a loss of 35 dB at 500 Hz, he shows normal frequency selectivity there, and is also within the normal range for detecting phase changes in stimulus complexes centered at that frequency. At 1 kHz, however, where his selectivity is degraded (the ERB about 50% bigger than normal), he is better than any normal listener tested at distinguishing SAM from QFM at modulation rates from 600 to 800 Hz. At the same frequency, he is able, under certain circumstances, to distinguish complexes which are added together in cosine phase from those added together in sine phase (Rosen, 1986). Normal listeners were unable to perform this particular discrimination.

Of course, not every hearing-impaired listener will show better discrimination performance than normal listeners. For one thing, the detection of phase changes clearly relies on sufficiently good temporal analyzing ability. Although an impaired listener may theoretically gain an edge by broadened auditory filters, s/he may just as well lose it through impaired temporal processing. What is striking, though, is that only very rarely are impaired listeners, even when they are impaired to a profound degree, much inferior to normal listeners in this task. As the impaired often become less sensitive to changes in the amplitude spectrum, the relative role of phase is still likely to be greater than that found, for example, by Plomp and Steeneken (1969) for normal listeners, even if they are not more acute in absolute terms.

THE PERCEPTUAL IMPORTANCE OF INCREASED SENSITIVITY TO PHASE

Given that the frequently broadened auditory filters of the hearing-impaired will allow a greater interaction between spectral components, and hence a greater role for phase, there are likely to be two main ways in which this will influence the perception of speech.

Firstly, in so far as temporal information is important in the perception of spectral shape, as has been proposed, for example, by Sachs and Young (1979) and Young and Sachs (1979), impaired listeners will hear changes in phase as changes in vowel quality. In fact, many listeners, both normal and impaired, report phase changes in harmonic complexes as changes in vowel quality. Darwin and Gardner (1986) have shown changes in vowel labelling performance with changes in phase in normal listeners, and those effects are likely to be stronger in impaired listeners.

Secondly, in so far as pitch perception relies on a temporal analysis of waveforms after a preliminary frequency analysis (as in the models of Moore and Glasberg [1986] and van Noorden [1982]), the perception of voice pitch is likely to depend on the relative phases of the constituent components of the sounds in a much stronger way than is found in normal listeners. Such effects have been shown by Hoekstra and Ritsma (1977) and Hoekstra (1979), albeit for sounds that are only remotely related to speech. They used SAM and QFM complexes with a centre frequency of 2 kHz and modulation rates near 200 Hz. Instead of requiring listeners to discriminate between SAM and QFM sounds at the same modulating frequency (as in the experiments reported in the previous section), they were asked to discriminate changes in modulation rate with sounds that were both SAM or QFM. This is roughly equivalent to perceiving changes in the fundamental frequency of a speech sound from three upper harmonics (the so-called "residue"). Hoekstra (1979) reported that three of five hearing-impaired listeners were significantly better at discriminating changes in modulating frequency for the SAM complex ("in phase" components), than for the QFM complex. Normal listeners (and the other two impaired listeners) showed no difference between the two conditions. It seems likely that in the impaired listeners who showed this difference, widened auditory filters allowed spectral components to interact, thus affecting the waveform presented to the temporal analyzers by the auditory filters. We might well suppose that "in phase" spectral components would reflect the modulating frequency of the signal in a clearer way than "out of phase" components (figure 1).

A similar result has been obtained with rather more speechlike sounds. Rosen and Fourcin (1983) extensively investigated the auditory capabilities of one profoundly impaired listener who lost his hearing in his mid-forties as the result of a skull fracture from a fall. One ear was made totally deaf, while the other had a so-called "left-hand corner" audiogram (70 dB HL at 125 Hz, falling off to 115 dB HL at 1 and 2 kHz, with thresholds greater than 120 dB HL for 4 and 8 kHz). They found that his discrimination of changes in fundamental frequency in the voice frequency range was better when the stimuli were sinusoids, than when they were pulse trains or speech. This is the opposite pattern to that found in normal listeners, where discrimination of fundamental frequency in sounds with multiple harmonic components is generally better than that found for sinusoids at the fundamental (e.g., Henning and Grosberg, 1968). Rosen and Fourcin argued that the almost certainly impaired frequency selectivity of the listener (difficult to measure in such a profound loss), was allowing harmonic components to interact to a much greater extent than in the normal case. Thus the temporal analyzers were presented with more complex waveforms in the case of pulse trains than in the case of sinusoids, and this was the cause of the worsened discriminability.

Confirmation of this hypothesis is found in further studies of the same patient, who was tested on his ability to discriminate changes in fundamental frequency for three-harmonic complexes with fundamental frequencies near 240 Hz. The fundamentals of the two stimuli to be compared were fixed at 228.6 and 252 Hz (a change of 10.25%), and could be either SAM or QFM (equivalent to what would be obtained by a modulation frequency of exactly half the carrier frequency). The phase relationships of all the stimuli within a session were always the same (i.e., appropriate either for SAM or QFM). Here, a 2I-2AFC task was used, in which the listener was required to label the direction of the pitch change. To prevent the use of any possible loudness differences between the stimuli (which seems unlikely, anyway), each sound was jittered by a different random amount each presentation, over a range of $\pm$ 2 dB. Table 1 shows that the change in fundamental frequency between the two stimuli is more salient for the complexes in SAM phase. Note though that phase corrections were not applied to these sounds. At the input to the headphones, all components were in sine phase, but this led to SAM complexes that were in approximate -cosine phase.

Table 1, Percent correct in discriminating a fixed change in fundamental frequency in a three-component harmonic complex when the relative phases of the components in the pair of stimuli to be compared are varied.

phase relationship SAM	QFM	statistical significance of the difference
65.6% (of 90)	47.8% (of 90)	$p<0.05$

"In phase" components do not always lead to more salient pitches. Table 2 shows the results in a similar task by the same listener. The sounds are now four-component harmonic complexes with all components at the same amplitude. The fundamental frequencies of the two stimuli to be compared again differed by 10.25% (238.1 and 262.5 Hz), and the components could be either in sine phase (at the headphone output, as phase corrections were applied), or with alternating sine and cosine terms. No amplitude jitter was present. These results show that the detailed temporal structure of the sound needs to be taken into account, along with the phase distortions imposed by the listener's auditory system.

Table 2, Percent correct in discriminating a fixed change in fundamental frequency in a four-component harmonic complex when the relative phases of the components in the pair of stimuli to be compared are varied.

phase relationship		statistical significance of the difference
sine	alternating	
60.0% (of 110)	100.0% (of 110)	$p<0.001$

FINAL REMARKS

In considering the perceptual consequences of hearing impairment, we tend to assume that abilities to make perceptual distinctions, if not uniformly degraded, certainly do not become any better. Phase may be one instance in which a feature that is of relatively little consequence in determining the percepts of normal listeners (although possibly of great use in investigating temporally-based models of spectral feature extraction), becomes much more important in the hearing impaired. For example, increased sensitivity to phase may well be part of the reason why impaired listeners are often disturbed by reverberation (or, as one profoundly impaired listener put it to me: "The echoes, to people like us, are a bit disconcerting"), even though the randomization of phase relationships caused by reverberation has little perceptual effect in normal listeners (Plomp and Steeneken, 1973).

Not all phase effects may be negative. As impaired listeners can hear changes in phase as changes in vowel quality, even when they are poor at distinguishing vowels on the basis of their amplitude spectra, phase manipulations could provide a way of signalling useful information. The problem, though, is that if phase changes also affect the saliency (and indeed even the perceived value) of the voice pitch, how is one to independently manipulate the two perceptual features?

REFERENCES

1. Darwin, C.J. and Gardner, R.B. (1986). Mistuning a harmonic of a vowel: grouping and phase effects on vowel quality. J. Acoust. Soc. Am., 79, 838-845.
2. Dominitz, R.H. (1975). Headphone monitoring system for binaural experiments below 1 kHz, J. Acoust. Soc. Am., 58, 510-511.
3. Goldstein, J.L. (1967). Auditory spectral filtering and monaural phase perception, J. Acoust. Soc. Am., 41, 458-479.

4. Helmholtz, H. (1954). On the Sensations of Tone. New York.
5. Henning, G.B. and Grosberg, S.L. (1967). Effect of harmonic components on frequency discrimination. J. Acoust. Soc. Am., 44, 1386-1389.
6. Hoekstra, A. (1979). Frequency discrimination and frequency analysis in hearing. Ph.D. Thesis, University of Groningen.
7. Hoekstra, A. and Ritsma, R.J. (1977). Perceptive hearing loss and frequency selectivity. In: E.F. Evans and J.P. Wilson (Eds.), Psychophysics and Physiology of Hearing. Academic, London.
8. Licklider, J.C.R. (1957). Effects of changes in the phase pattern upon the sound of a 16-harmonic tone. J. Acoust. Soc. Am., 29, 780 (abstract).
9. Mathes, R.C. and Miller, R.L. (1947). Phase effects in monaural perception. J. Acoust. Soc. Am., 19. 780-797.
10. 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.
11. Moore, B.C.J. and Glasberg, B.R. (1986). The role of frequency selectivity in the perception of loudness, pitch and time. In: B.C.J. Moore (Ed.), Frequency Selectivity in Hearing. Academic, London.
12. Noorden, L. van (1982). Two channel pitch perception. In: M. Clynes (Ed.), Music, Mind and Brain. Plenum, New York.
13. Patterson, R.D., Nimmo-Smith, I., Weber, D.L., and Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. J. Acoust. Soc. Am., 72, 1788-1803.
14. Plomp, R. and Steeneken, H.J.M. (1969). Effect of phase on timbre of complex tones. J. Acoust. Soc. Am., 46, 409-421.
15. Plomp, R. and Steeneken, H.J.M. (1973). Place dependence of timbre in reverberant sound fields. Acustica, 28, 50-58.
16. Rosen, S. (1986). Monaural phase sensitivity: Frequency selectivity and temporal processes. In: B.C.J. Moore and R.D. Patterson, (Eds.) Auditory Frequency Selectivity, Plenum, New York.
17. Rosen, S. (1984). Hyperacute monaural phase sensitivity in the hearing-impaired, Brit. J. Audiol., 18, 257-258 (abstract).
18. Rosen, S. and Fourcin, A.J. (1986). Frequency selectivity and the perception of speech. In: B.C.J. Moore (Ed.), Frequency Selectivity in Hearing. Academic, London.
19. Rosen, S. and Fourcin, A.J. (1983). When less is more further work. Speech, Hearing and Language: Work in Progress, I, 3-27 (Phonetics and Linguistics, University College London).
20. Rosen, S. and Nevard, S. (in press). A headphone monitoring system for low-frequency psychoacoustics. Brit. J. Audiol., (abstract).
21. Rosen, S. and Stock, D. (in preparation). Auditory filter bandwidths as a function of level at low (125 Hz-1 kHz) frequencies.
22. Sachs, M.B. and Young, E.D. (1979). Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J. Acoust. Soc. Am., 66, 470-479.
23. Schroeder, M.R. (1959). New results concerning monaural phase sensitivity. J. Acoust. Soc. Am., 34, 1579 (abstract).
24. Young, E.D. and Sachs, M.B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J. Acoust. Soc. Am., 66, 1381-1403.

The Psychophysics of Speech Perception

edited by:

M.E.H. Schouten

Institute of Phonetics
University of Utrecht
Utrecht
The Netherlands

1987 **Martinus Nijhoff Publishers**
Dordrecht / Boston / Lancaster
Published in cooperation with NATO Scientific Affairs Division