THE CURIOUS HALF-OCTAVE SHIFT: EVIDENCE FOR A BASALWARD MIGRATION OF THE TRAVELING-WAVE ENVELOPE WITH INCREASING INTENSITY

Dennis McFadden
Department of Psychology
University of Texas
Austin, Texas, USA 78712

INTRODUCTION

Following exposure to an intense tonal stimulus, there may be no temporary threshold shift (TTS) at the exposure frequency, even though there is considerable hearing loss at a higher test frequency. This effect—commonly known as the half-octave shift in TTS—is among the oldest, best known, and most widely cited facts of psychoacoustics, yet it stands without a generally accepted explanation. This paper has three primary purposes: (1) to review the physiological findings apparently relevant to this upward shift of maximum effect; (2) to demonstrate that half octave-like shifts are not unique to TTS experiments, but rather, they can be found in data obtained in a wide array of psychophysical tasks not involving auditory fatigue; and (3) to present a possible explanation of the upward shift which accounts for many of the existing facts, and thus appears worthy of serious consideration by both theorists and experimentals.

By way of historical perspective, the first reports of the maximum temporary hearing loss being shifted upward in frequency from the exposure stimulus were apparently made by Rawdon-Smith [1] and Ewing and Littler [2] (see Perlman [3] for a review and additional data). The fact was then well-documented and popularized by Davis et al. [4] in their classic monograph on auditory fatigue. Specifically, the effect is this: when narrow-band or tonal waveforms of high intensity are used as exposure stimuli, the test frequencies showing the greatest TTS are typically 0.5-1.0 octave above the exposure [4-6]. Indeed, Davis et al. [4]—and many others since—sometimes measured essentially no hearing loss at the exposure frequency, even though sensitivity at higher frequencies was reduced 15-20 dB or more. Upward shifts of maximum hearing loss are evident in measures of permanent threshold shift (PTS) as well as in TTS. Looking across studies and paradigms, there are some indications that the mechanisms producing the upward shift of effect may differ somewhat in the frequency regions above and below about 1000 Hz. Below about 1000 Hz, shifts are sometimes absent, small, or in the opposite direction [4, 7-10], or they appear to behave in accord with rules different from those above 1000 Hz [11]. If confirmed, this difference may prove to be a reflection of the apparent change in the characteristics of the oocluar partition in that region (at least in humans) that also produces linear rather than logarithmic frequency spacing below about 1000 Hz [12].
Over the years, this upward displacement of maximum hearing loss has universally come to be called the half-octave shift in TTS; however, it must be emphasized that the term is merely descriptive, not explicit—the maximum TTS can occur at the half-octave frequency, at the octave frequency, or above or below either of these. But while there may be nothing particularly magical about the half-octave frequency itself, the effect the term denotes is truly curious, and it is surprising that the half-octave shift has received so little experimental and theoretical attention during the 50 years it has been known.

The student seeking an explanation of the upward shift in maximum hearing loss following exposure is naturally drawn to look for possible structural bases for the effect. In this regard, the auditory periphery appears to offer a wealth of possibilities. For example, there are two, mutually dependent populations of transducers—inner and outer hair cells—which (1) have different gross and ultrastructural morphologies and different patterns of distribution along the cochlear partition; (2) synapse with different populations of primary afferents; (3) have different patterns of innervation, both afferent and efferent; (4) are thought to be stimulated by different modes of shearing; (5) are known to be differentially sensitive to such insults as anoxia, reduction in blood supply, and drug toxicity; etc. In addition to differences between the two populations of hair cells, the shift of maximal effect to adjacent frequency regions might be attributed to factors such as possible changes in the mechanics or micromechanics of the cochlear partition [13,14], and surely there are various metabolic, biochemical, or other nonstructural features that might also be suggested as relevant to the half-octave shift.

But even though the cochlear partition offers numerous possible candidates for the basis of the upward shift in hearing loss following exposure, to my knowledge none of these has ever been proposed or investigated. Indeed, as far as I can determine, only one explanation of the half-octave shift has received repeated mention over the years, and even these mentions have often been somewhat off-handed; certainly no one has attempted to systematically examine the various implications of this explanation, or to test it directly. The essence of this recurring idea is that the peak of displacement of the traveling-wave (TW) envelope may not occur at the same point along the cochlear partition for all stimulus intensities, but rather the peak may gradually move basally as intensity is increased. If the peak of displacement of the TW envelope does move basally along the cochlear partition as intensity is increased, that is tantamount to having a multiplicity of intensity-dependent tonotopic maps of the cochlear partition—the map for each successively higher intensity being displaced slightly toward the base. Thus, the cochlear location maximally "fatigued" by a high-intensity exposure would be basal to the point at which the exposure tone produced its maximal displacement when presented at a low intensity. That is, at threshold intensities, the frequencies activating the maximally fatigued portion of the partition would be ones higher than the exposure frequency—which would itself maximally activate a slightly more apical region. This point is illustrated in the left panel of Fig. 1.

Place theorists of pitch perception long ago suggested such changes in position as an explanation of the claim that high tones get higher in pitch and low tones get lower in pitch with increasing intensity [15]. In recent times, the basic idea has been mentioned in various contexts [8, 16-26].

Davis [27] recently advanced a "conceptual model" of cochlear mechanics that suggests a possible mechanism for the basilarward migration of the peak of the displacement envelope with increasing stimulus intensity. Davis argued that the total envelope of membrane displacement produced by a tonal stimulus is always the point-for-point sum of two components. First,
Fig. 1. Schematic representation of migration of the traveling-wave envelope with increasing intensity. Shown at the left, on the dimension of cochlear position, are the top halves of the displacement envelopes for five tones at each of three intensities. For each intensity, a neural tuning curve is shown above the TW envelopes, and then is shown again in the right panel, now correctly located on the frequency dimension. The bar marked TTS at the bottom left is meant to indicate the region of "fatigue" produced by tone F2 at the highest intensity shown; at "threshold" intensities, F2 itself would show little or no aftereffect of the exposure, but frequencies above the exposure frequency would.

There is the mechanical TW of displacement that is the result of the physical properties of the cochlear partition, and which has a relatively broad peak of displacement; this is the classical TW envelope described by von Bekesy and which Davis regards to be basically a passive response to the input signal. Second, there is a steeply rising, horn-shaped pattern of displacement located at the apical foot of the classical TW that is the product of active elements residing in the cochlea which Davis refers to collectively as the cochlear amplifier (CA). In Davis' model, the CA is subject to a saturating non-linearity, so its relative contribution to the total envelope of membrane displacement is greatest at low intensities; thus, at low intensities the peak of the total envelope of displacement lies 2-3 mm apical to the peak of the classical, mechanical TW envelope. As intensity increases, the contribution of the apically located CA to the total envelope of displacement continuously diminishes, and the relative contribution of the basal, mechanical TW grows, with the result that the peak of the total envelope of displacement gradually migrates basally. In exposure situations, then, the "elements" maximally "fatigued" or "injured" are those located near the peak of the mechanical envelope of displacement—elements tuned to frequencies higher than the exposure tone when measured at low stimulus intensities.
While appealing in many ways, Davis' conceptual model does appear to contain a limiting condition that seems not to exist in the data. Namely, once the stimulus intensity reaches the point that the total displacement envelope coincides with the mechanical envelope, there can apparently be no further upward shift. Yet, there are numerous physiological and psychophysical demonstrations of the maximum effect being upward in frequency by an octave or more and apparently growing without bound [4,6,28].

As we shall see, there is considerable evidence in support of the idea of a basalward migration of the point of peak displacement with increasing intensity, but there are also some gaps and contradictions in the story. One point to remember during the following summary of the evidence is that even when some physiological or psychophysical measure does migrate in the correct direction for our purposes, the magnitude is seldom adequate to account for a half-octave shift—which corresponds to approximately 2 mm along a human cochlear partition [29]. (On the other hand, it is also rare for these measures to cover the 90-100 dB range typically covered in TTS paradigms.) In addition, it is important to remember that there is no intent here to suggest that migration is the sole explanation for any fact considered; surely most of these effects are at least partially the result of other aspects of cochlear behavior. Thus, in the following we will be emphasizing qualitative, not quantitative, agreements.

PHYSIOLOGICAL AND ELECTROPHYSIOLOGICAL EVIDENCE

By now, most auditory scientists are sufficiently familiar with the twin concepts of the spatial pattern of the TW envelope and the tuning characteristic of a single point along the cochlear partition that they can easily alternate their thinking between them. However, when one adds the factor of basalward migration of the TW envelope, the potential for confusion grows. One fact for the reader to remember is that if the maximum of the TW envelope does migrate basally with increasing intensity, then any measurements obtained from a fixed position along the membrane (e.g., tuning curves obtained from the basilar membrane, a hair cell, or a primary fiber) should appear displaced downward in frequency as intensity is raised—reference to Fig. 1 should make this point clear. (Note that the same downward shift is also expected in these various measurements following any manipulation that reduces or eliminates the cochlear amplifier segment of the TW envelope, but that type of shift is different from an intensity-dependent one.) It is sometimes helpful to think about this downward shift in tuning in terms of the input/output functions underlying it. If the input/output functions are parallel for all individual frequencies within the passband of a fixed point on the partition, then the tuning curves associated with that point would have both the same nominal characteristic frequency (CF) and essentially the same shape at all intensities. But if, for any reason, the input/output functions are not parallel, the nominal CF will be different at different intensities. Now, migration of the peak of the TW envelope ought to produce non-parallel input/output functions for frequencies located in different regions within the passband of the cochlear position being studied. The reason is that input/output functions for frequencies on the high side of the "actual" CF (the CF measured at near-threshold intensities) should be relatively shallow because the peaks of membrane displacement for those frequencies are gradually sliding basalward, "out from under" the membrane location being measured; in contrast, the input/output functions for frequencies on the low side of the actual CF ought to be steeper since their points of maximal displacement will remain apical to the membrane location being studied. (Again, this is not to argue that factors other than migration might not also contribute to these different input/output functions.)

298
Early data on basilar membrane mechanics did show some of the expected characteristics described above [30,31], but since early data have become suspect in light of findings about the effects of inadvertent damage to the delicate structures in the cochlea [32], we shall concentrate upon data obtained under procedures designed to minimize cochlear damage. Using Mossbauer techniques, Sellick et al. [33] obtained non-parallel input/output functions of the sort suggested above. For data they felt were among their best (their Fig. 5a), the CF for the location of their Mossbauer source was 19 kHz at low displacement amplitudes (actually velocities); it dropped to 18 kHz at higher amplitudes and to 16 kHz at the highest displacement amplitudes. Rough estimates indicate that the rate of migration was about 0.07 octave/10 dB in this 16-19 kHz region. The basilar membrane data of LePage and Johnstone [34] also indicated shallow input/output functions at frequencies above CF. Thus, an expectation that follows directly from the idea of basalward migration of the peak of the TW envelope is confirmed—the apparent CF for a fixed basilar membrane location decreases with increasing stimulus intensity.

Similar effects on input/output functions and tuning curves were observed by Russell and Sellick [35] in their studies of high-frequency inner hair cells, but while the magnitudes of the effects appear large, estimates of migration rate are not possible from their published figures. The inner hair cell data of Dallos [10] do not show these effects as strongly as the above data do, but interestingly, he was studying hair cells located in the 800-Hz region of the cochlear partition, where half-octave effects are sometimes problematic.

While cochlear microphonic (CM) data do provide weak support for a basalward migration of the displacement peak with increasing intensity [36], Dallos [18] has clearly shown that the non-linear input/output functions of CM can produce the appearance of a basalward migration without there actually being a translation of the displacement peak along the cochlear partition. Thus, we shall ignore CM data.

Not surprisingly perhaps, the primary fibers show the same downward shift in tuning that Russell and Sellick [35] observed in the inner hair cells. At high intensities, the peak firing rate often moves to a lower frequency [7], as does the CF of the tuning curve [37-39]. Again, the 1000-Hz region appears to be a dividing line; Evans [37], Liberman and Mulroy [38], and Liberman and Kiang [39] all reported downward shifts for cells with CFs above 1000 Hz and upward shifts for cells with CFs below 1000 Hz. Sachs and Abbas [40] showed that the slopes of the functions relating primary-fiber discharge rate to stimulus intensity are increasingly less steep as the test frequency is raised above CF, and Abbas and Sachs [41] showed that two-tone suppression is less at high levels of excitor and suppressor than at low levels when the frequency of the suppressor is above CF. Both of these outcomes are in accord with a basalward migration of the TW envelope with increasing intensity.

Kim and Molnar [42] attempted to gain insight into the pattern of excitation along the cochlear partition for different intensities by stimulating about 1400 primary fibers with identical sets of tonal stimuli and then plotting the population response to each stimulus. Unfortunately, this interesting procedure is hindered by the problem of saturation in neural firing rates, so that as the intensity of the tone was increased, the peak of activity in the population response first broadened and then simply flattened, and while Kim and Molnar [42] did report hints of a basalward shift, the evidence is not unequivocal. Also relevant, perhaps, is that the stimulating tone used by Kim and Molnar was 1000 Hz.
Physiological Aftereffects of Intense Stimulation

Of the various physiological measurements that have been made during and following exposure to intense sounds, only a few are relevant to the idea of a basalward migration of peak displacement with increasing stimulus intensity. Lonsbury-Martin and Meikle [8] obtained post-exposure measures in primary fibers after exposures of one minute to moderately intense tones of different frequencies—at CF, one-half octave below CF, and one-half octave above CF; they monitored both driven and non-driven firing rates. For our purposes, the principal result was that exposure to a tone one-half octave below CF produced the greatest, and the longest lasting, post-exposure depression in firing rate (measured at CF). The second most effective exposure was the CF stimulus itself, and the least effective was the stimulus one-half octave above CF. Cody and Johnstone [43] repeated the Lonsbury-Martin and Meikle [8] experiment using more exposure frequencies at more intensities, with the same general outcome. Cody and Johnstone alluded to mechanical explanations of the effect, without specifying or embracing a particular one. Since the exposure duration was only one minute in both the Lonsbury-Martin and Meikle and the Cody and Johnstone experiments, both can be thought of as analogous to TTS experiments. These results clearly square with the explanation being considered here; exposures with frequencies at and above the CF of the fiber produced maximal displacement of the cochlear partition, and thus maximal "fatigue," at points basal to the one served by the fiber being monitored, whereas the point being served by that fiber was maximally displaced, and fatigued, by exposures to frequencies somewhat below CF.

Results parallel to those of Lonsbury-Martin and Meikle and to Cody and Johnstone have been obtained with the whole-nerve action potential (AP). The AP shows greater aftereffects above the exposure frequency than at it [44-46], and there is a tendency for the maximum effect to shift to higher frequencies with increasing exposure intensity [45-48]. In a related paradigm, Abbas [49] used the AP to obtain tuning curves before, and soon after, presentation of adapter tones of varying intensity and duration. In accord with the idea of a migrating peak of displacement in the TW envelope, he found greater post-adaptation decrements in the AP at frequencies above the adapter than at it, and this effect was greater for the more intense adapters.

Other physiological findings obtained following exposure to intense sounds will be mentioned for the sake of completeness, even though they reveal little about the possibility of a basalward migration with increasing intensity. To my knowledge, no experiment on basilar membrane mechanics has ever systematically manipulated the exposure variable, but as noted above, it is now believed that many early experimenters—including von Bekesy himself— inadvertently studied noise-damaged cochleas [32]. The effects of noise damage, and of numerous other sorts of damage, are less-sharp tuning and a downward shift in the measured CF [28, 32, 33, 50]. It is comforting to those seeking parsimony that essentially identical changes are believed to occur in hair cells and primary fibers following damage and/or exposure [26, 51-53]. Cody and Johnstone [47] monitored single primary fibers during prolonged exposures to intense tones one-half octave below CF. Beginning about one minute after the onset of the exposure, there was a downward shift in the nominal CF that continued and progressed for about the first hour of exposure. In that time, sensitivity had declined by about 60 dB on the average, and nominal CF was approximately half an octave below original CF. As the exposure continued, the tuning curves became increasingly flatter and less sensitive, and somewhere between 2-3 hours of continuous exposure, all cells failed to respond to test tones of 110 dB. This experiment clearly involved both TTS and PTS, but there is no information as to the dividing line between them. Liberman
[26] worked with a PTS-like paradigm and also found a downward shift in apparent CF (as much as 0.66 octave) in primary fibers following noise exposures that induced varying degrees of localized damage to cochlear structures. Averaged across neurons of differing sorts (his Fig. 2), the downward shift in CF was at a rate of about 0.08 octave/10 dB of PTS, but the variability was high. Finally, Van Heusden [54] has obtained downward shifts in nominal CF of nearly two octaves in the tuning curves of cells in the anteroventral cochlear nucleus following exposure to intense noise.

COMMENT

The preceding section revealed that there are considerable physiological data in qualitative agreement with the idea of a migration of the peak of displacement of the TW envelope with increasing stimulus intensity. In the following section, a comparable array of psychophysical evidence is presented in support of the idea. However, most of the psychophysical demonstrations we will examine differ in an important way from most of the physiological demonstrations just considered, and the difference points up an important feature of the migration effect not yet mentioned.

The physiological experiments discussed above typically involved measurements made using single tones as stimuli. In contrast—with the exception of TTS paradigms—psychophysical manipulations capable of establishing whether or not a slightly different segment of cochlear partition is being stimulated by different intensities always require the presence of a second tone, or set of tones; logically, one cannot hope to measure a tuning curve or filter characteristic psychophysically with single-tone stimuli. This necessity raises an interesting and potentially difficult problem. Clearly, if both the signal and the auxiliary tones were subject to migration, and their rates of migration were the same, no psychophysical experiment could possibly reveal evidence of that migration, even if it were substantial. As the following sections will show, a number of psychophysical tasks do provide evidence for migration of the TW envelope with increasing intensity, implying that—under some conditions at least—the signal and auxiliary tones are either not both migrating, or not both migrating at the same rate. Let us consider two general possibilities for the form of the migration. First, below some "critical intensity" there may be no migration of the displacement peak with increasing stimulus intensity, and above that critical intensity, migration may proceed in the same linear manner for all frequencies (illustrated at the top in Fig. 2). For the sake of the example, let us consider the implications of this alternative for a tone-on-tone masking situation. Increases in the intensity of a relatively weak signal—necessitated by increases in the intensity of the masker—cannot erase the evidence for a migration in the masker tone as long as the signal stays below some critical value of intensity. Once the intensity of the signal also exceeds the critical value, however, migration will appear to cease, and the masker will appear to have migrated less than it actually has.

Alternatively, it might be that migration occurs over the entire range of intensity, but that the rate of migration is greater at high intensities than at low (illustrated at the bottom in Fig. 2). According to this version, a relatively weak signal is itself subject to some migration, but as long as there is a reasonable difference in the intensities of the signal and the masker(s), migration will be revealed. Note that in this case the magnitude of migration will be underestimated by psychophysical measures, but that, unlike the first version, the apparent magnitude will (correctly) increase continuously over the entire range of intensities studied. From this consideration of these two possibilities, it should be clear that—unlike physiological measures—psychophysical procedures can, at best,
Fig. 2. Schematic representations of two simple versions of the process of migration of the traveling-wave envelope. The curves shown describe the loci of the peaks of the displacement envelopes for each of three tonal stimuli. In the top panel, the peak of displacement does not migrate along the partition until the intensity reaches some "critical" value; above that value, the peak migrates basally at a fixed rate. In the bottom panel, the peak of displacement migrates over the entire range of intensity, but, for a unit change in intensity, the magnitude of the basal-ward shift is greater at high intensities than at low. Psychophysical measures necessarily reveal only differences in rate of migration.

measure differential rates of migration of stimuli, and thus, it would be remarkable indeed to find good quantitative agreement between physiological and psychophysical results.

PSYCHOPHYSICAL EVIDENCE

Masking

If different intensities of the same frequency do maximally activate different portions of the cochlear partition, one might expect this fact to
be revealed in tone-on-tone masking experiments in which maskers of different intensity and test tones of varying frequency are used. The expectation would be that at low and moderate masker intensities, the masking pattern would peak at the masker frequency, whereas at high masker intensities, the peak of masking would be displaced toward higher frequencies. The classical data of Wegel and Lane [55] and of Egan and Hake [56] naturally come to mind in this context, and while there are no marked shifts of the masking peak in those data—just a general broadening of the extent of masking with increasing masker intensity—Fig. 3 of Egan and Hake does show more masking at 1000 and 1400 Hz than near the 400-Hz masker at high masker intensity [57]. While this evidence is obviously not compelling, simultaneous masking procedures using two tones are known to be plagued by the problem of beats. When this problem is circumvented, more satisfying evidence for migration is obtained.

Zwislocki developed a simultaneous contralateral masking procedure for use in his studies of central masking. Zwislocki et al. [58] used a fixed signal frequency in one ear and a slowly varying masker frequency of fixed intensity in the other; the subject adjusted signal intensity for threshold. If there were migration, this fixed-signal procedure should yield a downward shift in the maximally effective masking frequencies, and the peaks of the masking functions obtained by Zwislocki et al. (their Figs. 9-11) do show a displacement downward in frequency with increasing masker intensity. The rate was about 0.04 octave/10 dB.

Vogten [22] minimized the problem of beats in simultaneous presentations of tonal signal and masker through the use of a phase-locking technique that obviates uncontrolled intensity fluctuations. He has reported experiments conducted in two ways: with the signal level held constant and the masker level adjusted at a number of masker frequencies (yielding what is commonly known as a psychophysical tuning curve or PTC; see his Fig. 6) and—like Zwislocki et al. above—with the masker level held constant and the signal level adjusted at a number of masker frequencies (see his Fig. 7). Both methods reveal a monotonic decrease in the "peak" frequency of the masking function with increasing intensity. In a second article, Vogten [59] argued that the shift in the peak seen at low intensities is due to a contribution from a suppression mechanism, but he accepted the possibility that the shifts at higher intensities may be related to a migration in the peak of displacement along the cochlear partition. He estimated the migration to be about 0.04-0.07 octave/10 dB in the 1000-Hz region. In related research, a rate of migration similar to that reported by Vogten [59] was obtained by Florentine and Houtsma [60] in a single impaired ear; the simultaneous PTCs of Carney and Nelson [61] also showed a downward shift of the "peak" at higher signal intensities; and while the simultaneous masking data of Zwicker and Jaroszewski [62] contain no obvious shifts, perhaps this is owing to the use of weak maskers only.

Temporal masking is the most commonly used procedure for avoiding the unwanted interactions between signal and masker that occur in tone-on-tone experiments. If the frequency of a forward masker were fixed, the migration idea would lead us to expect that at moderate and high masker intensities, the signals most masked would be slightly higher in frequency than the masker. The classical data of Munson and Gardner [63] confirm this expectation (their Figs. 8 and 18)—the peaks of the forward-masking patterns they obtained migrate from their 1000-Hz masker frequency toward 1500 Hz as masker intensity is increased. The results of a forward-masking experiment by Zwislocki and Pirotta [64] were similar; for two of their three subjects, the maximum of the masking function moved upward in frequency by about 1000 Hz as the intensity of their 3150-Hz masker was increased from 60 to 100 dB SPL (about 0.11 octave/10 dB). In the intervening years, numerous investigators have published temporal masking
patterns or PTCs that contain shifts which are in accord with the idea of a basalward migration of the peak of the TW envelope with increasing intensity, and which are often substantial [65-71]. Abstracting across studies, it does appear that shifts of reasonable magnitude are most reliably obtained in the frequency regions above 1000 Hz. McFadden and Yama [9] obtained shifts of about 0.10-0.15 octave/10 dB, and they suggested that shifts are more likely to be seen when the signal duration and masker-to-signal interval are longer than typically used nowadays (at least 50 msec).

We saw, in our earlier consideration of tuning curves for fixed basilar membrane locations, that there is a necessary relation between tuning curves and the input/output functions measured for individual frequencies within the passband of the point of interest, and this is no less true for PTCs than for basilar membrane tuning curves. The migration idea predicts relatively steeper growth-of-masking functions for masker frequencies below CF, and relatively flatter ones at and above CF, because in the latter case the maskers are "moving away" from the signal as masker intensity is increased, necessitating greater increases in level to achieve a fixed increment in masking. Results of this sort have been reported by Widen and Vierling [69], among others. In related, unpublished measurements, Pasanen used forward masking to determine the psychometric function for a 20-msec, 6500-Hz tone of 12 dB SL masked by preceding 200-msec tones of higher or lower frequency and variable intensity. In accord with the idea at hand, the psychometric function obtained with the higher frequency masker was less steep than that with the lower frequency masker--approximately 9 dB of change in the masker was required to cover the range from 60% to 90% correct with the high-frequency masker, and only about 3 dB with the low-frequency masker.

**Lateral Suppression**

Were the migration idea under examination here correct, one might expect to see manifestations of it in lateral suppression data. Specific predictions are made difficult by the complex of possible interactions between suppression, masking, and migration, but there are indications that effects compatible with the migration idea do occur [22]: suppressors on the high-frequency side of the signal become decreasingly effective with increasing level [21]; the asymmetry of lateral suppression toward the high-frequency side of the signal which is seen at relatively low suppressor intensities appears to reverse at high suppressor intensities [72]; and, as noted previously, the effects on two-tone suppression in primary fibers are different when suppressor level is varied for frequencies above, versus below, cell CF [41].

**Pitch Shifts**

The initial reaction of nearly everyone meeting the migration idea for the first time is, "But if that were right, pitch ought to shift upward with increasing intensity, and it doesn't." In fact, there now appears to be general agreement that the pitch of high-frequency tones does get higher (and the pitch of low-frequency tones gets lower) with increasing intensity [73-75], but it is necessary to qualify this fact as it pertains to the migration idea. First, the pitch shifts observed typically amount to only a small percentage change in frequency (although Kim reported data for several subjects who showed upward shifts of 10-15%, about 0.04-0.06 octave/10 dB, at 5000 Hz [19]). Second, the middle frequencies (about 1000-4000 Hz) typically show little or no pitch shift with intensity; the effect is restricted to more extreme frequencies. These qualifications are generally unnecessary for the various other psychophysical and physiological effects described above that are in accord with the proposition at hand. Thus, the relative constancy of pitch unquestionably stands as the
most troublesome fact to rationalize with the other evidence pertaining to migration. To paraphrase an old saw, "pitch is a witch." (Perhaps the relative invariance of pitch across wide ranges of intensity is evidence that pitch constancy has had great evolutionary value and—since the "problem" of migration has presumably existed for a long time—that the nervous system has developed ways to achieve constancy despite migration.)

Aftereffects of Intense Stimulation

Given that the half-octave shift was first noticed in TTS research, it is natural to expect to find considerable detailing of the effect in the TTS literature. Surprisingly, and unfortunately, this is not the case; the half-octave shift has itself not been the object of much research. In the vast majority of TTS experiments, hearing loss and/or recovery are measured at only one or a few test frequencies, and these are chosen with the presumption that there will be a half-octave shift rather than to test for its presence. But while the TTS literature is not packed with crucial information relevant to the possibility of migration of the peak of the TW envelope with increasing intensity, some facts are worth mentioning.

As already noted, the data of Davis et al. [4], Ward [5], Cohen and Baumann [6], and others, do show more TTS above the exposure frequency than at it. Also, the finding by Reger and Lierle [76], Hirsh and Bilger [77], Selters [78], and Young and Sachs [79] that a low-intensity exposure tone produces more TTS at the exposure frequency than does a high-intensity exposure comes into accord with the migration idea once it is appreciated that this decline is accompanied by increasing TTS at higher test frequencies as exposure intensity increases [19,20,77,80]. Note that at very high exposure levels—above about 115-120 dB—there is evidence of a different sort of decline in TTS with increasing exposure intensity [5]; this decline can exist across a wide array of test frequencies, and thus it appears to be due to mechanisms quite different from migration of the TW envelope. For the sake of completeness, we note that there is at least one report of the peak of TTS moving toward the exposure band with increasing exposure intensity [81].

Following exposure to an intense tone, McFadden and Plattsmier [25] measured changes in both detection threshold and in suprathreshold loudness at both the exposure and the half-octave frequencies. In accord with past research, there was little or no TTS at the exposure frequency, even though there was as much as 15-20 dB of TTS one-half octave higher. But interestingly, there was as much as 15 dB of loudness change at the exposure frequency, even though its threshold was unchanged. An outcome of this sort is expected if the peak of basilar membrane displacement migrates basalward with increases in intensity, because absolute sensitivity for the exposure frequency would be mediated via a relatively apical and unaffected segment of the partition, whereas the loudness measures—involving as they do a match to a stimulus of suprathreshold intensity—would be mediated over a slightly more basal segment of the cochlear partition, one including the (more basal) region of maximal "fatigue" (Fig. 1).

If there were a gradual basalward migration of the peak of displacement along the cochlear partition with increasing intensity, the TTS induced by a relatively weak exposure stimulus should all be at, or close to, the exposure frequency, whereas the TTS induced by more intense stimuli should gradually move toward increasingly higher frequencies as exposure intensity increases. As noted previously, the TTS data of Hirsh and Bilger [77] support this idea, but they only monitored post-exposure performance at two frequencies— the exposure and the half-octave frequencies. Hood [80] and Kim [19,20] did test at several frequencies other than the exposure frequency, and both did find evidence for a gradual upward shift of the
peak TTS frequency, but both used exposures of only 30 sec, and as a consequence, obtained only small amounts of TTS (less than 4 dB in most cases). Since both experimenters used tracking procedures to estimate post-exposure sensitivity, such small values of TTS are problematic. Ward [5] reported TTS for a series of exposure intensities, but the weakest was sufficiently intense (115 dB) that migration was already well-established. McFadden and Plattsmier [24] described experiments in which frequency patterns of TTS were obtained for a number of intensities of the exposure tone; the intensities were less extreme than Ward's, and longer and more intense than Hood's and Kim's, and data were collected for more test frequencies than is typical. The expectation was that the peak in the TTS pattern would gradually migrate from the exposure frequency at low intensities toward higher frequencies with increasing exposure intensity, but the results were not this simple. These detailed TTS patterns did not have single, well-localized peaks that just shifted upward in frequency with increasing exposure intensity. Instead, the patterns had at least two, apparently reliable, local maxima whose relative sizes changed with exposure intensity (calculated "centers of balance" for these TTS patterns shifted upward at a rate of about 0.03-0.12 octave/10 dB). Indications of a two-peaked pattern of effect can also be seen in the data of Hood [80], Munson and Gardner [63], Zwislocki and Pirodda [64], Zwislocki et al. [58], Kim [19,20], and Pirodda and Ceroni [67]. Perhaps these complex patterns are reflections of the transition envisioned by Davis [27] from an overall TW envelope dominated by the cochlear amplifier component to one dominated by the mechanical TW component. If so, detailed TTS patterns may be providing us with more information about cochlear mechanics than has previously been appreciated.

It is well-known that there can be substantial shifts in pitch induced by exposure to intense sound [82]. Davis et al. [4] did not even begin their pitch measurements until 60 minutes following the exposure, yet they often measured shifts of 40-80%. For our purposes here, perhaps the most important fact about these pitch shifts is that they were minimal or non-existent for test frequencies below and at the exposure frequency, and they were maximal for frequencies a half to a full octave above the exposure frequency. This locus of maximum effect is exactly what the migration idea predicts, of course, especially when it is recalled that Davis et al. used quite intense exposures. Also of interest is that the direction of the pitch shifts was always upward in the Davis et al. experiment, which is in accord with what is now known about the neural aftereffects of exposure. Following exposure, a primary fiber loses its sensitive tip and is thereby "re-tuned" to a lower frequency. Thus, presenting a test tone in the region of maximal post-exposure "fatigue" will lead to greatest activity in fibers whose actual CF is higher than the test tone, and since those fibers are presumably still "labeled" by the higher neural centers as being higher in pitch than the test tone, a higher frequency is needed contralaterally to achieve a pitch match [46]. The pitch of the tonal tinnitus that is often induced by narrow-band exposures is also located well above the exposure frequency [83].

Form of the Migration Process

Now that the psychophysical data have been presented, we can return to the questions raised previously about the form of the migration process—specifically, whether there exists a "critical intensity" at which migration begins, and whether migration proceeds at the same rate for all intensities once it has begun.

In the data of Hirsh and Bilger [77], the upward shift in TTS appears to begin somewhere between 60 and 80 dB SL. In the Kim [19,20] data, there are hints of an upward shift of TTS beginning at 45 dB SL for some subjects. (In the McFadden and Plattsmier data [25], there appears to be an
upward displacement of the TTS pattern even for the lowest exposure intensity used—82 dB SPL—so they are uninformative on the issue of a critical intensity.) In the forward-masking data of Zwislocki and Pirodda [64], the earliest evidence of an upward shift in the masking pattern is for maskers of about 80 dB SPL; the same is true for Pirodda and Ceroni [67]. For Munson and Gardner [63], a shift first appears between 70 and 90 phons. The observations of Selters [78] further muddy the waters; Selters monitored the growth of TTS at 1000 Hz as a function of the intensity of a 1000-Hz exposure tone (of 10-sec duration, so adaptation may be a better term than TTS—but no matter). After growing continuously for exposures of 5-30 dB SL, TTS plateaued from 30-60 dB SL, and from 60-80 dB SL, TTS declined from its value at lower exposure intensities before rising again at exposure intensities of 90-100 dB SL. The plateau and the decline are to be expected if there is a "critical intensity" for migration, but the final rise is not. A similar plateau is also present in the forward-masking data of Munson and Gardner [63]. In contrast, at the physiological level, the basilar membrane input/output functions for frequencies above CF are shallow, with no obvious inversion point, from the lowest intensities tested [33]. It is obviously difficult to generalize across these disparate studies, but—recalling the earlier observation that psychophysical measures have the potential to obscure the details of the migration process—it may be prudent to side with the still-scarce physiological data, and tentatively conclude that migration exists over the entire range of audible intensities.

The psychophysical data also provide a blurred picture of how migration proceeds once it does begin. Recall that, were the parallel version at the top of Fig. 2 correct, migration should appear to cease once the weaker of the two test tones begins to migrate, but if the non-parallel version were correct, migration should (correctly) appear to grow continuously with increasing intensity. There is precious little experimental evidence available on this point, and what little does exist unfortunately goes in both directions. There are some indications that maximum TTS moves continuously upward in frequency with increasing exposure intensity [4,25], but in other cases, maximum TTS appears to remain in the vicinity of the half-octave frequency in the face of large increases in exposure intensity [5]. At this time, the available physiological data are of little assistance with this question.

SUMMARY

Over the years, numerous investigators have suggested that the displacement peak of the TW envelope may migrate basalward with increasing intensity and that this migration is the basis of the so-called half-octave shift in TTS, as well as of other effects. Recently, the physiological evidence relevant to this issue has become less contradictory, and it now stands in qualitative agreement with the migration idea. Considerable psychophysical data have been in qualitative accord with it for years, so the migration idea carries the strong attraction of being a parsimonious and unifying concept covering a broad spectrum of effects. To be sure, no single piece of evidence examined can yet be viewed as conclusive, and the magnitudes of the various effects are not in perfect agreement, but taken together, the evidence is impressive. Not to be ignored is the fact that there is presently no alternative explanation with anything like the breadth of coverage possessed by the migration explanation. Perhaps an implication of this argument requires emphasis—half octave-like shifts appear to be normal consequences of the basilar membrane's response to increasing intensity; they are not just the result of over-driving and fatiguing the system as in TTS paradigms. At the very least, the migration explanation deserves to finally be put under the microscope of experimental
investigation, so that its future acceptance or rejection can be based on more than hints and trends extracted by looking across experiments, paradigms, and decades.

ACKNOWLEDGMENTS

Versions of this paper have existed since 1981. In this time I have been supported in part by grants from the National Institute of Neurological and Communicative Disorders and Stroke (NS 08754 and NS 15895) and the National Institute for Environmental Health Sciences (ES 03539). Special thanks are due Wilson S. Geisler and Craig C. Wier for their numerous valuable comments on a preliminary draft, and to E. G. Pasanen and H. S. Plattsmeier for assistance of various sorts.

REFERENCES


309
52. D. Robertson, Effects of acoustic trauma on stereocilia structure and spiral ganglion cell tuning properties in the guinea pig cochlea, Hearing Research, 7:55-74 (1982).
64. J. J. Zwislocki and E. Pirodda, On the adaptation, fatigue and acoustic trauma of the ear, Experientia, 8:279 (1952).
NOTE ADDED TO WORKSHOP

In the above paper, much of the physiological evidence cited in support of a migrating TW envelope was necessarily indirect. In the absence of direct measurements of basilar membrane motion made over a wide range of intensity, and made with the required fine grain in frequency. I was forced to make inferences about the behavior of the cochlear partition from tuning curves and input/output functions taken from hair cells and primary fibers. While I hope this evidence was adequate to convince most readers, new direct evidence was presented at this Workshop that should help to convince skeptics as well. In his Fig. 3, Dr. Patuzzi shows measurements supporting Davis' suggestion that the broad, and more basally located, mechanical component of the TW envelope should come to dominate the more apical cochlear-amplifier component at high stimulus intensities, with the result that peak displacement will migrate basally with increases in intensity. Now that this missing link in the argument has finally been supplied, I hope that both physiological and psychophysical investigators will be motivated to study more fully the numerous interesting concomitants of this migration.

DISCUSSION

Salvi: Some of your psychophysical data suggests that at high levels there was a dramatic jump in the data, whereas some of the explanations you have given sounds like the shifts are very gradual. Is it a shift or is it a jump?

McFadden: There are two options to explain the migration data. The critical intensity idea, or there is the continuously varying idea. On the basis of the available data, it is difficult to decide. The psychophysical data looked like there is very little or no migration at first over a reasonable range of intensity and until some break occurs. But the electrophysiological shows that the differences begin at the lowest intensities. So it appears that the migration physiologically starts right at the very beginning.