

# Chapter 17

## Dispersive and synchronization hearing impairments

### 17.1 Introduction

Hearing and vision are the two foremost senses in humans that gather information about their environment. Impairment in either sense can lead to a substantial loss in quality of living. In comparison with prescription glasses, contact lenses, and corrective surgeries that passively correct the most prevalent refractive errors type of eye disorders, sensorineural hearing impairments are not as amenable to treatment using today's technology that is centered on restoring audibility using active circuitry. However, the comparison between the most prevalent visual and auditory disorders may be fundamentally misleading, because the visual ones are refractive where visibility is not a problem (except for cataract), whereas the auditory ones are often dominated by audibility issues. But the present work begs the question whether dispersive and synchronization disorders of hearing exist, in a manner that degrades the auditory image and its perception, similarly to refractive problems in vision. In other words, it is argued that there can be conditions under which listeners experience abnormal loss of sharpness (or blur) of auditory images.

The comparison between auditory and visual impairments may also appear somewhat naive and even unfair, because of the heightened intricacy of structure, functional complexity, and enclosed anatomy of the hearing organ compared to the eye. The acoustic signal reaching the outer ear is transduced multiple times before it is neurally coded, whereas the optical signal is carried electromagnetically throughout the periphery and its transduction is confined to the photoreceptors and a subset of nonimaging photosensitive retinal ganglion cells (Burns and Lamb, 2003; Do and Yau, 2010). Furthermore, as this work has attempted to show, the analogous image processing that takes place within the visual periphery is realized centrally within the auditory system. Other technical factors challenge hearing theory in unique ways compared to vision: a much larger total relative bandwidth, overlapping modulation and carrier spectra, strong nonlinearity, folded dimensionality, and highly concealed and encoded image. These factors undoubtedly hamper the attempts to fully account for some of the more complex hearing impairments to a degree that enables engineering a fully satisfactory treatment for them.

This chapter aims to explore what can go wrong in auditory temporal imaging and whether it can be a root cause for any hearing disorders. There are three approaches that such an analysis can follow. A bottom-up approach may be pursued to systematically examine every component in the imaging system, test its likelihood to break down, and analyze the effects it may have on the image. In general, this should include drift in the main parameter values of the system (e.g., cochlear dispersion, time-lens curvature), excessive aberrations, and impairments in the accommodative function. A second

top-down approach may be pursued to survey known hearing impairments and find out whether they can be reinterpreted using the new theory. In the third approach, known impairments in vision can be analogized to hearing in search for known disorders. Ideally, these approaches should converge and help demystify some of the most persistent and lesser-understood aspects of certain hearing disorders. The exploration space is therefore very large, so only select cases will be examined, where the degree of speculation is also deemed relatively modest. In that, an ulterior intention of this chapter is to open the door to other interpretations of hearing disorders in light of an imaging theory—most likely by incorporating new data and methods.

The analyses presented in the chapter are organized according to the main dispersive system components, where available evidence for changes in each parameter is considered—often with respect to specific disorders. Therefore, several conditions, such as age-related hearing loss, are featured in several sections throughout the chapter. We systematically explore possible effects and evidence to changes in to the cochlear and neural group-delay dispersion, the time-lens curvature, problems in phase locking, enhanced aberrations, and disorders of accommodation. The result is not a complete map of possible dispersive disorders, but it is deemed sufficient to encourage further explorations that will hopefully be less speculative in nature.

## 17.2 The missing dimensions of hearing impairments

While there is no specific test at present that assesses auditory sharpness, it is clear that information about it cannot be possibly obtained in audibility tests such as the pure-tone audiogram. Indeed, the limitation of the audiogram as a diagnostic tool for sensorineural losses has been occasionally noted, as it does not necessarily correlate with suprathreshold listening performance where audibility is not an issue (e.g., [Plomp, 1978](#); [Papakonstantinou et al., 2011](#); [Parker, 2020](#)). Several extra factors have been proposed that may constitute the missing dimensions that describe the space of hearing impairments more completely—often involving the outer-hair cell (OHC) function and the auditory nerve (e.g., [Parker, 2020](#)). Most recently, the degeneration of the ribbon synapse of the auditory nerve, or **cochlear synaptopathy**, has been implicated with a host of “**hidden hearing loss**” problems ([Schaeffe and McAlpine, 2011](#)). Definitively demonstrating it in humans and rigorously associating it with hearing impairments have been elusive ([Bramhall et al., 2019](#); [Ripley et al., 2022](#)), but at least on one occasion it was suggested that a loss of spectrotemporal details in hidden hearing loss may be analogous to visual blur of object edge details ([Shinn-Cunningham et al., 2017](#)). Another prominent line of research attempting to identify the missing dimensions in hearing loss has focused on the association between degraded sensitivity either to temporal-fine structure (TFS) or to envelope cues. Often these studies also estimate the broadening of the auditory filters, or the distortion product otoacoustic emissions (DPOAE) to monitor the function of the OHCs. Either one of these factors could be considered an additional dimension to hearing loss beyond audibility.

As has been contended repeatedly in this work, in order to be able to make sense of envelope and TFS cues, it is essential to consider the degree of coherence of the incoming signal (and masker), as well as its downstream incarnations. Perfectly coherent and incoherent signals may be relatively easy to analyze, but in reality, sounds (like speech) are partially coherent, which requires a more involved analysis. This is not only because partial coherence is more complex mathematically, but also because the degree of coherence is often unknown and fluctuates in time, which has to be considered in light of multiple time constants throughout the ascending auditory system. This means that the available part of the signal that contains usable TFS information—that may be extracted using coherent detection—fluctuates both extrinsically and intrinsically to the auditory processing. Furthermore, once the signal is transduced, the auditory system is most certainly dynamically engaged in optimizing and accommodating to the specific signals (at least during selective attention), using a complex

network of feedback processes. The very property of phase locking was speculated to be one such process, which we hypothesized arises in the auditory phase-locked loop (PLL) of the OHCs and is propagated downstream from there. The normal operation of the PLL itself may come across as a narrowband filter, so phase-locking and frequency selectivity can appear to correlate using certain measurement methods. Finally, the usefulness of phase locking itself depends on its retention downstream, so that different sites of lesion can degrade its quality between the OHCs and the auditory cortex, and then contribute to a measurable loss of coherence.

Keeping in mind these complex interrelationships between the different components of the auditory system, the potential role of dispersive losses as a possible cause for hearing impairment will be explored below. Also, phase locking as a function that depends on the health of the OHCs—and hence on the auditory PLL—will also be discussed.

## 17.3 Evidence for dispersive shifts in hearing impairments

This section systematically reviews the different elements of the auditory imaging system using the bottom-up approach spelled out in the beginning of the chapter. It highlights data from literature that may point to significant drifts in the values of the dispersive variables in hearing-impaired listeners. The methods considered are closely related to those that were originally used to estimate the normal-hearing dispersive values of the system (§11).

### 17.3.1 Changes in the total group-velocity dispersion

Temporal auditory imaging includes three gross dispersive elements: the cochlear group-delay dispersion  $u$ , the time-lens curvature of the outer hair cells (OHCs)  $s$ , and the neural group-delay dispersion  $v$  from the inner hair cells (IHCs) through to the auditory nerve and all the way to the inferior colliculus (IC). The cascade of the three is observable as a characteristic chirping effect, which constitutes a defocus that differentially affects coherent and incoherent signals. Earlier, psychoacoustic results of the phase curvature measurements by Oxenham and Dau (2001a) were used to test the temporal imaging theory and derive the temporal aperture of the system (§12.4). Additional studies are reviewed below that indicate whether there are significant differences between normal-hearing and hearing-impaired subjects with regards to the total curvature.

#### Schroeder phase curvature drift

Only few psychoacoustic studies used the Schroeder phase method to specifically test hearing-impaired subjects (Summers and Leek, 1998; Summers, 2000; Oxenham and Dau, 2004). In general, there are differences between sensorineural hearing-impaired subjects and normal-hearing controls mainly in the level of masking achieved and, occasionally, in the curvature as well. Oxenham and Dau (2004) provided detailed curvature data at two frequency bands with the same stimuli as in Oxenham and Dau (2001a)—a 250 Hz tone masked with Schroeder complex with  $f_0 = 12.5$  Hz and a 1000 Hz tone masked with  $f_0 = 100$  Hz (as well as  $f_0 = 12.5$  Hz, which was not used in §12.4). The two frequencies that were measured produced mixed results, though (Oxenham and Dau, 2004, Figure 1). While all 12 hearing-impaired subjects showed a significantly reduced masking effect, the observed curvature at 250 Hz shifted from  $C = 0$  (see Eq. 12.29) only in two subjects (HI10 to  $C = 0.125$  and HI12 to  $C = 0.5$ ), who also had the largest audibility loss at that frequency (55 dB HL). At 1000 Hz ( $f_0 = 100$  Hz), the three normal-hearing subjects had curvatures of  $0.5 \leq C \leq 1$ . At the same frequency, the hearing-impaired curvature shifted to  $C \approx 0$  in two cases (HI2 to  $C = 0.125$  and HI4 to  $C = 0.25$ ; both with no exceptional pure-tone audiograms),

and was about the same as for the normal hearing in two other cases (HI10 and HI6), but it was unmeasurable in the other eight cases with no clear peaks. As [Oxenham and Dau \(2004\)](#) noted, a key assumption in this test method is that the curvature is locally constant (i.e., within the paratonal approximation). However, curvature constancy may not be true in the case of hearing loss. Also, a complex frequency-dependent change to the impaired cochlear phase response may have caused some shifts. The authors tried to correlate the findings with filter broadening and found a significant but weak effect, which can be taken as indicative of a loss of OHC function that causes a loss of compression. Similar results were obtained at 1, 2, and 4 kHz of by [Summers and Leek \(1998\)](#), but with curvature values limited to  $C = 0, \pm 1$ .

### Derived-band auditory brainstem response changes

An alternative method to evaluate the possible effect of hearing impairment on the total group-delay dispersion relies on the derived-band evoked auditory brainstem response (ABR) ([Don and Eggermont, 1978](#)). This method employs unfiltered broadband clicks that are partially masked by high-pass filtered pink noise. By successively subtracting the adjacent-band ABR responses, it is possible to recover octave-wide narrowband ABR responses. The narrowband responses are primarily attributed to changes in the cochlear function, but include all dispersive contributions downstream ([Don et al., 1998](#)). Changes to the total group-delay dispersion may be appreciated from changes in the frequency-dependent slope of the ABR latencies. [Strelcyk et al. \(2009\)](#) compared the derived-band ABR in five frequency bands of five normal-hearing and 12 hearing-impaired subjects with mild-to-moderate losses. Wave-V latency curves were obtained for all subjects, but had larger variability in the hearing-impaired data. Using psychoacoustic frequency selectivity measures at 2 kHz, the authors concluded that the difference in latencies between the two groups is most likely due to auditory filter broadening, and not due to synaptic or central differences. This is in accordance with linear filter theory, which predicts slower filter response (larger group delay) the narrower the bandwidth gets. The results agree with previous ABR data where filter bandwidth was not measured directly ([Don et al., 1998](#)).

As in other studies (§11.7.3), the latency data in [Strelcyk et al. \(2009\)](#) were fitted according to a power law of a similar form to Eq. 11.20

$$\tau(f, i) = a + bc^{0.93-i} f^{-d} \quad (17.1)$$

where  $b$  is a subject specific parameter,  $a$  is a constant delay,  $i$  is the click intensity normalized to 100 dB, and the frequency  $f$  is in kHz. The equation did not have a significantly different frequency-dependent slope  $d$  for the two subject groups, but had a significantly steeper intensity-dependent slope  $c$  for the hearing-impaired group.

In order to estimate the range of variation between impaired and normal hearing, two corner ABR responses (at 93 dB SPL) from [Strelcyk et al. \(2009, Figure 3\)](#) are replotted in Figure 17.1, left. The data points were iteratively fitted to Eq. 17.1 from the initial mean values provided in [Strelcyk et al. \(2009\)](#), where  $b$  was matched individually and  $d$  only for the hearing-impaired subject. The group-delay dispersion can be readily calculated from the power law, by differentiating  $\tau$  with respect to  $\omega$  and dividing by 2. The fits were then compared to the defocus parameter, as was obtained earlier from Eq. 12.34, where it was defined as  $x = u + sv/(s + v)$ . The normal-hearing subject response from [Strelcyk et al. \(2009\)](#) is relatively close to the  $x$  value we obtained earlier (smaller by a factor of 0.8, approximately), whereas the hearing-impaired response is half as large, on average (Figure 17.1, right).

The decrease in latency and, consequently, in total group-delay dispersion data analyzed here suggests that the defocus  $x$  must decrease accordingly, through an effect on  $u$ ,  $v$ , or  $s$ . As the magnification is close to unity,  $M \approx 1$ , then  $x$  can be rewritten as  $x = u + v/M \approx u + v$ .

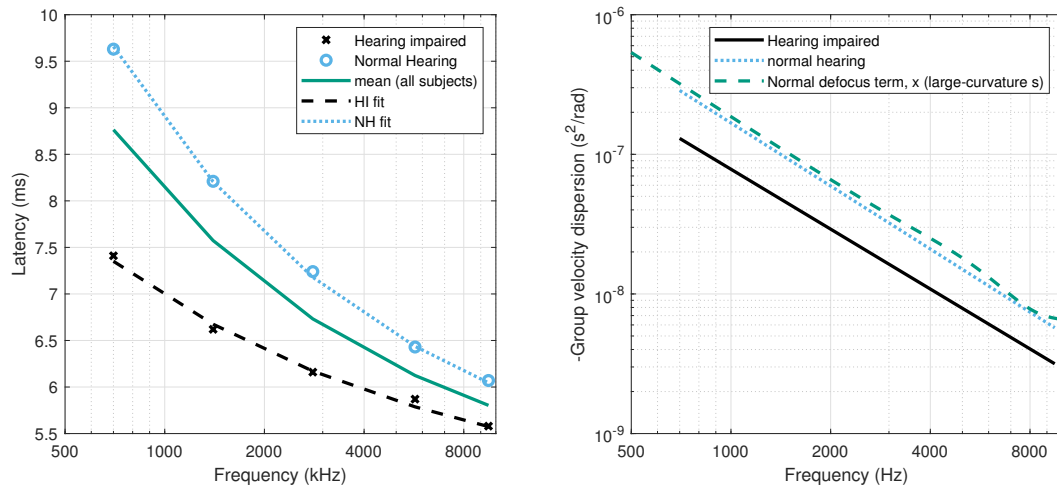


Figure 17.1: Comparison of the derived-ABR latency and group-velocity dispersion of a normal-hearing and a hearing-impaired subject from [Strelcyk et al. \(2009\)](#), Figure 3. The original normal-hearing data is plotted in black asterisks and the hearing-impaired subject in blue circles. **Left:** The latency measurements and individual fits to Eq. 17.1, as well as the average across all subjects (hearing impaired and normal hearing). **Right:** The corresponding group-velocity dispersion, compared to the defocus term  $x$  obtained earlier in §12.4, obtained for the large-curvature time lens (both broad- and narrow-filter estimates) that we have used in most chapters (§11.6.4), which produces a close prediction to the normal ABR.

Therefore, the causes for the measurable change may stem from a change in  $u$  or in  $v$ . However, the two parameters are close in magnitude (Figure 12.2, right), so a decrease of  $x$  by half, as was found above, cannot be achieved by changing a single parameter. Nonetheless, a filter broadening explanation may be stretched to account also for the interdependence of  $u$  and  $v$  that leads to covariation when the filter is broadened. Hence, on the cochlear side, the mechanical activity of the OHCs extends to the basilar membrane region also basal to the characteristic frequency (CF), which impacts the cochlear group delay that belongs to  $u$ . On the neural side, the group delay of the auditory filter and its associated IHC belong to  $v$  and depend on its broadened bandwidth (Figure 11.1). This functional division makes the differentiation between  $u$  and  $v$  challenging (see also [Eggermont, 1979](#)). We will look into contributions of the specific components below (§17.3.2 and §17.3.4), using data that pertain to where differentiating between them is more clear-cut.

### Filter broadening and group-delay dispersion changes

Regardless of the exact cause of the differences between the two listener groups (filter broadening or other), the implications of the drop in the total group-delay dispersion entail a decrease in the maximum amount of defocusing that the system can provide (not taking into account accommodative effects on the time lens that may dynamically vary the defocus). It means that the ability to segregate signals based on their degree of coherence—achieving optimal sharpness and depth of field—may be impaired. This may translate to difficulties in hearing speech in noise, which are common in hearing-impaired populations. However, a direct connection between loss of defocus and speech intelligibility cannot be established at present. Still, it is instructive to briefly review some of the effects associated with filter broadening on hearing.

In literature, the effect of increased bandwidth on hearing is considered a function of degraded frequency selectivity, which diminishes the spectral contrasts of speech in noise (e.g. [Moore, 2007](#), pp. 214–219). In cochlear hearing-impaired listeners, the role of frequency selectivity may be difficult to test independently of other factors, and the results have been mixed. For example, a

principal component analysis was used to account for the performance of hearing-impaired subjects in an extensive test battery (Festen and Plomp, 1983). Two components accounted for 65% of the variance, and frequency selectivity was clustered along with speech-in-noise performance, whereas hearing threshold was clustered with speech in quiet. Similar test battery correlations were found in Horst (1987) and Dreschler and Plomp (1985), where temporal as well as spectral resolution was found to be strongly correlated with performance. In contrast, smearing the spectral envelope of speech for both normal- and impaired-hearing listeners was only weakly correlated with speech reception thresholds, indicating that spectral resolution per se may not be critically important for speech recognition (Ter Keurs et al., 1993b). Similarly, in another test battery reported by Strelcyk and Dau (2009), no significant correlations were found between frequency selectivity and speech reception, while temporal measures proved to be more strongly correlated. Interestingly, Strelcyk and Dau (2009) included two subjects with normal audiograms, who had poor speech-in-noise performance (“obscure dysfunction”), whose frequency selectivity at 750 Hz was comparable to that of hearing-impaired subjects. The two subjects’ binaural masked detection performance was also poorer than normal-hearing subjects.

One method to tease apart the effect of frequency selectivity from elevated threshold has been through simulation. In normal hearing, lower speech intelligibility was found in noise, but not in quiet, in simulated filter broadening experiments (Ter Keurs et al., 1992; Baer and Moore, 1993). The broadening or spectral smearing was achieved by convolving a Gaussian filter of the desired bandwidth with the envelope of the short-time Fourier transform of the signal, and then resynthesizing it (Ter Keurs et al., 1992). The performance worsened with poorer signal-to-noise ratio (SNR) and broader filters. However, with competing speech as masker a poorer SNR was required to degrade the intelligibility to that of broadband noise with better SNR (Ter Keurs et al., 1993a; Baer and Moore, 1994; see also, Duquesnoy, 1983). This was the case for normal bandwidths and was further exacerbated with broadened filters. The difference was attributed to the increased capability to listen in the dips at equal SNRs with speech masker that is naturally amplitude modulated. Higher-level stream segregation may have a role in this difference as well.

Any filter broadening effects that are captured in these findings appear to be too general to be able to tie directly to dispersion. However, a more mundane explanation for the deterioration in performance has to do with the shift in the dual spectrum of speech (§ 6.6.2), as components that are normally resolved in adjacent filters when they are narrow have to be processed temporally within the broadened channels. However, the temporal modulation transfer function (TMTF) cutoff frequency is constrained by the sampling rate of the system, which might not keep up with the increased modulation frequencies in the shifted dual-spectrum. Effects of changes to the TMTF are discussed in § 17.3.4 and § 17.7.1.

### 17.3.2 Changes in cochlear dispersion as a result of impairment

Cochlear insults that impair the OHC mechanics lead to changes in the response of the basilar membrane (BM), primarily around the CF. By selectively disrupting the OHC motility using furosemide in the chinchilla’s cochlear base around CF= 9 kHz, Ruggero and Rich (1991) found that the phase response is both frequency- and intensity-dependent in an area confined to the CF range, but is unaffected at lower frequencies than the CF (1 and 5 kHz). In another physiological study of noise-induced hearing-impaired chinchillas, the relative delay between channel CFs separated by up to half an octave was found to change relatively to normal hearing animals, in a frequency-dependent manner (Heinz et al., 2010). These and similar findings suggest that it may be impossible to completely disentangle changes in the OHC curvature (associated mainly with  $s$ ) from the cochlear dispersion

One global cochlear change that relates more specifically to the BM mechanics is an abnormal rise in fluid pressure in the scala media. This condition is referred to as **endolymphatic hydrops** and is usually associated with Menière's disease. Endolymphatic hydrops cause the BM stiffness to increase and, consequently, for the cochlear traveling wave velocity to increase as well (Donaldson and Ruth, 1996). This velocity does not change with cochlear hearing loss without hydrops (but see Eggermont, 1979). Derived-band ABR may be used to compare the traveling wave velocity, using the inverse of the latency function (Donaldson and Ruth, 1993). This procedure assumes a negligible contribution of the OHCs (Don et al., 1998), which may not be entirely justifiable given that a subset of Menière's disease patients exhibit OHC pathology as well (Cianfrone et al., 2000). Also, OHC motility depends on cochlear fluid changes that tend to cause hearing fluctuation—a characteristic of the disease (Dulon et al., 1987). However, the presence of endolymphatic hydrops has been attributed to changes in the traveling-wave velocity, based on derived-band ABR measurements (e.g., Kim et al., 1994; Donaldson and Ruth, 1996; Claes et al., 2008). Abnormal data from of Menière's patients suggest that even when the ABR latency increases, its frequency dependence remains close to normal, as the traveling wave velocity slope is unchanged on a log scale (Donaldson and Ruth, 1996, subjects GM, VF, and TH). All other seven patients in that study had normal latency curves, as did the noise-induced hearing-impaired subjects in the study. This means that the cochlear group-delay dispersion  $u$  does not always change with the velocity in Menière's patients, but if it does, then it increases by a constant factor. As a consequence of Menière's disease, patients normally have poorer speech intelligibility in quiet than normal-hearing listeners (Bosmana and Smoorenburg, 1995; Tonndorf, 1976). In contrast, they exhibit no unusual sensitivity to noise compared to normal-hearing subjects—contrary to listeners with noise-induced and acquired impairments (Bosmana and Smoorenburg, 1995). Once again, it is not clear that the dispersive shift per se should have a significant effect that can lead to this peculiar pattern. In principle, it may be that speech in quiet relies more on coherent imaging, where dispersion can be more dominant to the different coherent and incoherent components of speech, so it might be impacted. In contrast, speech in noise relies more heavily on incoherent imaging, where phase locking makes relatively little difference, so that changes to  $u$  also do not matter much.

Changes in the apparent cochlear dispersion can be theoretically caused by conductive losses as well. However, these losses tend to be flat and using broadband evoked ABR measurements reveals only small mean differences in latency, with high variability, between impaired and normal ears (Mackersie and Stapells, 1994; Ferguson et al., 1998). These findings are not suggestive of great dispersive changes to the cochlear group-delay dispersion  $u$ , but derived-band measures will be needed to obtain more refined association.

### 17.3.3 Static changes in the OHC (time lens) curvature

We hypothesized two new roles for the OHCs in this work: a phase locked loop (PLL) and a time lens. The speculative nature of these ideas notwithstanding, disentangling possible impairments in them may not always be possible, especially considering the traditional functions of the OHCs—to provide amplification and dynamic range compression, sharpening of the filters, generation of distortion products, and therefore the production of otoacoustic emissions (OAEs). Below, we focus on the time lensing feature of the OHC and return to the PLL in 17.4.

We have argued that the time lens with its estimated curvature has a role in setting the degree of defocus in the image and its accompanying depth of field. Having a time lens also enables the temporal image to retain its orientation and not become time-reversed as would be the case with two negative group-delay dispersion segments—more in line with current knowledge of hearing operation. We look at two of its functions that were explored earlier. The olivocochlear reflex

effects that may be related to the time-lens impairment are examined in §17.7.

### Forward masking and depth of field

It was argued in § 15.11.3 that the pattern of forward-masking variation for different degrees of coherence between temporally adjacent stimuli can serve as the auditory analog for the spatial depth of field. As nonsimultaneous masking interacts with cochlear hearing impairment, it is interesting to find out if any effects may be attributed directly to the focal time of the lens, which by way of analogy is thought to control the depth of field, at least in part. Another significant part of the depth of field is most likely controlled centrally, so the exact cause may be difficult to pin down in most of the observations below.

In general, the release from forward masking changes with the hearing impairment, due to the elevated absolute threshold. When the masking thresholds are compared at equal sound pressure level, then the baseline simultaneous level is the same in both groups, but the forward masking decays much more slowly for the hearing-impaired than the normal-hearing listeners (Glasberg et al., 1987). This effect can be accounted for using the loss of compression in the impaired cochlea (Oxenham et al., 1997). If instead the comparison is done at equal sensation level, so that the relative levels above the normal and elevated absolute thresholds are compared, then the masking decay is better matched between the two groups, although still somewhat slower for the hearing-impaired listeners. These results hold for both on-frequency and off-frequency masking (Nelson and Pavlov, 1989). The ability to benefit from release from forward masking may not be always equal between the groups, especially at high frequencies, where hearing-impaired listeners may benefit less than normal-hearing listeners from various manipulations in the masker (Svec et al., 2016). As a consequence, old normal-hearing and hearing-impaired listeners tend to score more poorly than young normal-hearing controls in vowel-consonant or consonant-vowel pair identification, as a function of noise and the delay between the phonemes, which may potentially translate to poorer speech intelligibility in complex conditions (Fogerty et al., 2017).

It is difficult to draw direct conclusions from these data regarding the time lens. It is not impossible that the nonlinearity of the cochlea and loss of compression also entails time-lens curvature reduction, as they are all thought to depend on the healthy operation of the OHCs. However, at present we have no direct information to support this, based on forward masking patterns alone. Given that the forward masking decay is slow and long, elongation of its decay pattern may be a result of central changes.

Regardless of the specific cause for it, there appears to be substantial changes to the extent of depth of field as experienced by hearing impaired listeners, primarily due to the change in the decay function cause by masking.

### Binaural diplacusis

As the lens curvature dominates the expression of magnification, it enabled a rough prediction of the stretched octave effect, which is qualitatively similar to published data. If this effect is indeed related to the time lens, then it implies that abnormal pitch height perception may correspond to static shifts in the time-lens curvature.

**Binaural diplacusis** is a hearing disorder in which the same stimulus is perceived as having a different pitch in the two ears (Shambaugh, 1907). It may also be accompanied by a loss of tone quality and increase of distortion (Ward, 1955). Slight diplacusis ( $< 1.5/16$  octave difference, which is 112 cents—a slightly sharpened semitone) is not uncommon in normal hearing people, but it is quite common with unilateral hearing-impaired loss, where differences may be up to a few tones in extreme cases (Colin et al., 2016). Because of its small effect, it is usually not a dominant source of



dissatisfaction even among professional musicians (Jansen et al., 2009). In sloping high-frequency hearing loss, the pitch difference tends to be larger in the impaired range compared to the unimpaired range, and the pitch is usually perceived as higher than the reference pitch of the better ear.

Binaural diplacusis is particularly challenging for pitch theory, because the direction of pitch change is inconsistent among studies and listeners, even when they have similar losses (i.e., pure-tone audiograms) *f*. Still, the most common explanation so far has been due to place theory, as impaired cochlea results in a shift in the frequency-position mapping of the BM, so that places that are normally maximally excited by a frequency *f* move apically toward less impaired areas, which are normally excited by a lower frequency  $f' < f$  (Müller et al., 2010; Colin et al., 2016). However, this explanation is at odds with the few cases that show pitch shifts in the other direction, sometimes combined with intensity-dependent effects that change the shift direction too (Burns and Turner, 1986). Standard temporal theory of pitch predicts correspondence to the stimulus period independently of place, but it also does not offer an obvious explanation for the inconsistencies. These considerations suggest that a combined place-time pitch perception may be the most adequate to account for the various observations (Colin et al., 2016; Moore, 2007).

The temporal imaging theory may be able to resolve the inconsistency of pitch perception in binaural diplacusis since it scales the frequency by the magnification factor. The scaling operation is both place- and time-dependent and depends almost exclusively on the OHC curvature *s* and on the neural dispersion *v*. Any shift in cochlear mapping or impairment to the degree of phase modulation of the healthy OHCs are therefore expected to affect the perceived pitch. When the time-lens curvature is unimpaired and the frequency-position remapping is the only physiological change, the frequency scaling will change along the magnification curve of Figure 15.10 (right). If the pitch that is normally associated with frequency is scaled, or stretched according to  $f \rightarrow M_f f$ , then the impaired pitch, under cochlear remapping, would be associated with the place and magnification of  $f' < f$ , so  $f \rightarrow M_{f'} f$ . However, as was concluded in § 17.3.1, the cochlear impairment also causes a decrease in *v*, which leads to an increase in magnification  $M^* > M$ . The compounded effect of remapping and magnification shift may cancel out, as  $f \rightarrow M_{f'}^* f$ , where  $M_{f'}^*$  may be either larger or smaller than  $M_f$ . As magnification is a function of frequency, the effect will be noticeable mostly where the magnification function bends, above 1 kHz, approximately (Figure 15.10). In case that the OHC dysfunction changes the value of *s*, deleterious changes in pitch may take place as well.

These hypothetical explanations of pitch changes also have to take into account the existence of dead regions in the cochlea, where severe diplacusis has been occasionally reported (Huss and Moore, 2005a). Such cases may be a result of severe aberration that distorts the signal that is coded off-frequency in active channels at the edge of the dead region (see § 17.6.3).

#### 17.3.4 Impairments in neural group-delay dispersion

The neural group-delay dispersion includes all auditory pathways past the OHCs: the peripheral IHCs and the auditory nerve, as well as the different subcortical circuits, at least down to the IC. As such, there are several sites that can be impaired (Rance and Starr, 2015) and may cause the signal to be delayed in a frequency-dependent manner. Accordingly, multiple disorders have been identified, hypothesized, and associated with the peripheral and central parts of these pathways—most of which are not well understood at present. The dispersive point-of-view developed here assumes only a single time lens per auditory channel, so all pathways downstream from the auditory nerve are readily combined into a single frequency-dependent variable *v*, although nonuniform resampling artifacts and different noise regiments may apply. Changed parameters may include the neural group-delay dispersion (e.g., frequency-dependent latency), changes in the internal coherence (phase locking) properties, and changes in sampling rate. In some situations, the latter two may be interpreted as

the addition of noise to the signal (see § 16.4.7). Additionally, changes in the pupil function, or the role of aperture stop, may be considered as well, although it is likely dependent on the other dispersive changes. Evidence for some of these changes is presented below.

### Direct evidence for neural group dispersive changes

Based on data by [Strelcyk et al. \(2009\)](#), it was concluded in § 17.3.1 that changes in neural group-delay dispersion may be inevitable in listeners whose ABR latency dropped to about half its normal value. Otherwise, it was impossible to arrive at the dispersed ABR figures of some hearing-impaired individuals, given the baseline values used throughout this work. [Strelcyk et al. \(2009\)](#) concluded that the drop in ABR latency is very likely a direct result of auditory filter broadening, which causes a reduction in the filter group delay. However, it is not difficult to come up with other mechanisms in the periphery or in the brainstem that may cause frequency-dependent delay. When such hypothetical changes cannot be represented as a linear function of frequency, then the changes to the group delay are also frequency-dependent, which entails a change in group-delay dispersion  $v$ .

Theoretically, channel-specific damage to any of its neural dispersive elements can be either a result of within-channel changes to the filter group delay, or of dominant off-frequency response in place of damaged channels. The latter case is a form of aberration and was discussed in § 15.9 in the context of dead regions and is further analyzed in § 17.6.3. The former case is more difficult to establish directly based on available data. For example, the IHC synaptic delay is considered to be independent of the CF (e.g., [Palmer and Russell, 1986](#)). In the elusive condition of cochlear synaptopathy, the loss of IHC synapses is thought to impair the auditory function, but the minimum amount of loss necessary to cause perceptible impairment in humans is unknown at present and appears inconsistent between studies ([Bramhall et al., 2019](#)). Additionally, it has not been reported in this context, to the best knowledge of the author, whether synaptic impairment that does not completely incapacitate the synaptic activity can result in a shift to normal synaptic delay values. At the same time, in-vitro experiments indicate that the IHC synaptic delay is plastic under some conditions ([Goutman and Glowatzki, 2011](#); [Cho and Von Gersdorff, 2014](#))—a fact whose impact on hearing is also unknown at present.

Neural conductive abnormalities also characterize acoustic tumors (**acoustic neuroma**), which can be indirectly detected using derived-band methods ([Eggermont and Don, 1986](#); [Don et al., 1997](#); [Philibert et al., 2003](#)), but may be also gathered from an increase in ABR latency and decrease in amplitude (rather than complete absence of ABR; [Telian et al., 1989](#)). Regardless of the specific causes, if any such changes to neural conduction are selectively applied in some channels but not in others, then a respective change in neural group-delay dispersion will occur as well. Depending on the magnitude of the change, it may not be observable in broadband ABR measurements.

### Effects on the temporal modulation transfer function (TMTF)

Effects of a degraded contrast sensitivity function (CSF) (either in a limited band or broadband range of the spatial frequency spectrum) in vision are well-known to reflect a blurred image perception (e.g. [Bour and Apkarian, 1996](#); [Woods et al., 2000](#)). Thus, continuing the space-time analogy between vision and hearing, it is naturally to seek for various correlates between degraded TMTFs and various indicators of hearing loss.

A significant change in neural group-delay dispersion  $v$  can have different effects on the perceived image quality. If the system were completely continuous, then a smaller  $v$  should lead to a larger bandwidth of both the coherent and the incoherent TMTFs. This is also obtained by considering an increased bandwidth from the broadened filters, if they function as the aperture stop (something we suspect normally occurs only at low frequencies, § 13.4.1). In this case, the effect of the modulation

bandwidth increase will be compounded by the elongation of the temporal aperture as well. In reality, it has been repeatedly found that the TMTF bandwidth under cochlear hearing loss, once controlled for audibility, is about the same as in normal-hearing subjects, with rare cases that exhibit broader than normal TMTFs (Bacon and Gleitman, 1992; Moore et al., 1992; Moore and Glasberg, 2001; Füllgrabe et al., 2003, but see Grant et al., 1998).

Physiological data from the auditory nerve of chinchillas with noise-induced loss confirmed the psychoacoustical data, as relatively few fibers had an extended TMTF bandwidth, despite the broadened tuning of the carrier frequency (Kale and Heinz, 2012). This was explained using related data that found no increase in amplitude modulation (AM) synchronization strength as a result of noise-induced loss in chinchillas (Kale and Heinz, 2010), so the availability of extra high-frequency modulation information could not be coded by the auditory nerve (Kale and Heinz, 2012). A followup study that employed nonlinear analysis (Wiener kernel) of firing patterns of the chinchilla auditory nerve fibers with noise-induced cochlear impairment revealed changes in the TMTF shape and a 20–30% increase in frequency cutoff (Henry et al., 2014b). At the same time, the auditory tuning curves broadened by 100–200%. Therefore, in line with previous conclusions (Kale and Heinz, 2010, 2012), the increased modulation bandwidth was not fully matched by a corresponding firing rate, perhaps due to temporal limitations stemming from adaptation or refractoriness in the auditory nerve (Henry et al., 2014b).

It is possible to reframe the above findings in sampling terminology: high modulation frequency from the broadened cochlear filters may be neurally undersampled. This could have resulted in what appears like occasional aliasing (Kale and Heinz, 2012, Figures 2D, 5D and 5F; see §14.7 for a further discussion about aliasing with additional examples). However, as neural sampling is nonuniform, aliasing might be only a spurious concern, as the extra modulation information will be converted to noise, unless it is removed by an anti-aliasing filter (§14.7). Whether undersampling is perceived as noise or aliasing, its effect may be detrimental to hearing.

If certain species or situations exist in which undersampling due to broadened filters is not a concern, increased TMTF bandwidth still entails risk of losing defocus—as the contrast between coherent and incoherent stimuli erodes when it can only differentiate very high—possibly irrelevant—modulation frequencies<sup>171</sup>. A somewhat similar mechanism has been hypothesized without reference to imaging, based on extra gain that was seen to be internally applied to the TMTF of unilaterally hearing-impaired human listeners (Moore et al., 1996)<sup>172</sup> and in chinchillas (Kale and Heinz, 2010). It was suggested by Kale and Heinz (2010) that this gain may be excessive to the extent that hearing speech in fluctuating masker (e.g., competing speech) gets more difficult with cochlear hearing loss, because the received modulation depth of both target and masker is enhanced.

Theoretically, as a limiting case, a broadened filter causes the temporal aperture to significantly shorten and approach the dispersion limit, not only at low frequencies (below 660 Hz), where the filters are thought to function as the aperture stop (§13.4.1). It was seen in §12.7.5 that the system is biased toward geometrical blur, rather than being dispersion-limited. However, with much shorter aperture stop this bias can change and can lead to dispersive blur that might affect coherent objects too. This possibility is currently uncharted.

While the loss of defocus does not imply enhanced gain, it entails increase in contrast and loss of blur for high-frequency modulations, as fluctuating noise and speech, for example, may be perceived as equally sharp with respect to their modulation spectrum. This may be a component in the loss of

<sup>171</sup>For example, contrast the amplitude-modulation spectra surveyed in Singh and Theunissen, 2003 and the spectrum necessary for full speech intelligibility in Drullman et al. (1994a) and Drullman et al. (1994b).

<sup>172</sup>Anecdotally, Moore et al. (1996) referred to the enhanced modulation depth as “magnification”, but with no relation to imaging and no scaling implied for the imaged modulation frequencies. Modulation depth is exactly analogous to contrast, though, which is only indirectly dependent on magnification in standard imaging.

speech intelligibility with hearing impairments, as it can impact the depth of field. However, more precise estimation of the significance of this effect alone is lacking.

### Auditory neuropathy

Other hearing impairments may result in an increase in  $v$  due to abnormal latency patterns, which are not directly related to the cochlear filter bandwidth. In such cases,  $v$  may change independently of the temporal aperture and cochlear dispersion. **Auditory neuropathy** can be diagnosed when listeners have normal OAEs (undamaged OHCs) and cochlear microphonic response, but missing or abnormal ABR (Berlin et al., 2010). Documented neuropathy cases can have almost any possible audiogram from normal to profound, and the effect on their hearing also varies considerably between cases. Auditory neuropathy can have different etiologies that manifest in loss of IHCs, in demyelination or deafferentation of the auditory nerve, and/or the dysfunction of the synapse between the IHC and the auditory nerve (synaptopathy). Central lesions (e.g., acoustic neuroma) are excluded from this family of disorders. Auditory neuropathy can either cause reduced ability to phase lock and/or lead to poorer conductivity of information because of deafferentation (Rance and Starr, 2015). Common temporal symptoms of auditory neuropathy are a decrease in TMTF cutoff frequency and an increase in gap detection (Zeng et al., 1999, 2005). Additionally, if not altogether absent, there is an increase in click-evoked ABR wave-I to wave-V latency of a couple of milliseconds (Starr et al., 2001; McMahan et al., 2008). These features are fundamentally different from temporal processing in cochlear hearing loss, where adequate audibility can largely bring the performance to near-normal levels and the ABR latency decreases.

There is no published derived-band or evoked-burst-tone ABR that can clarify which channels dominate the latency increase in click-evoked ABR of auditory neuropathy, which can be indicative of the direction of change of  $v$ . Indirectly, it may be gathered from some frequency-dependent measures (e.g., frequency discrimination and beat detection) that the largest differences between auditory-neuropathy and normal-hearing patients manifest at low frequencies, whereas performance is near normal at high frequencies (Zeng et al., 2005). However, there are documented cases of auditory neuropathy with high-frequency loss and near normal speech audiometry, but no brainstem response beyond the cochlear microphonics (e.g., Berlin et al., 2010, Patient C). An excessive rise in  $v$  should lead to blur and loss of intelligibility in quiet and in noise for signals with significant high-modulation content, because of poor modulation response. This may have been the case for the few auditory neuropathy patients that were able to complete a speech intelligibility test (25 out of 95 tested; Berlin et al., 2010, Table 5). A more general option is that both coherent and noncoherent detection circuits are impaired (e.g., both the DCN and the VCN) in these cases, and create distorted or blurry images that do not converge properly at the level of the auditory retina (the IC).

## 17.4 Phase locking and coherent detection impairments

Throughout this work, we have emphasized the dichotomy between coherent and incoherent imaging and, in parallel, the one between coherent and noncoherent detection. Coherent detection (and imaging) ultimately depends on the auditory system being able to track the phase of the incoming signal, which is achieved through synchronization that is observable as phase locking. Therefore, a system with impaired phase locking can be expected to have impaired coherent detection and possibly become disproportionately dependent on noncoherent detection—on the slowly-varying temporal envelope. The effect of this kind of degradation may become noticeable in signals that are naturally coherent, such as sustained periodic sounds as are found in music. It may become also critical in

localization based on interaural time difference.

### 17.4.1 Loss of balance between coherent and noncoherent detection

The auditory defocus was introduced using the imaging analysis, but it is useful also in exaggerating the differences between coherent and noncoherent detection. Given a signal that is not completely incoherent, strong phase locking between two system elements suggests that coherence is conserved and sharp imaging is possible, whereas poor phase locking entails that some coherent stimuli may become decohered and effectively blurred. Where the normal-heard image comprises a certain balance of coherent and incoherent contributions, this balance may shift when the coherent contribution is impaired. Therefore, if impaired conditions can be identified where phase locking is degraded, they may correspond to increased blur of coherent or partially coherent signals, which can make them difficult to tell apart from incoherent maskers. For example, this can manifest as an increase of signal-in-noise threshold. Opposite conditions, where phase locking is excessive, may generate an auditory perception of pseudo-coherent images originating from incoherent stimuli, noise, or partially coherent narrowband sounds. This can give rise to coherent noise in some conditions rather than enhance the image sharpness. Overall, it may also entail a loss of the defocus effectiveness, as it is no longer possible to harness the signal coherence as a cue to differentially process signals according to their phase function.

Degraded TFS sensitivity—a proxy for phase locking—has been correlated with poor speech reception in modulated and unmodulated noise in sensorineural hearing-impaired listeners (e.g., [Buss et al., 2004](#); [Lorenzi et al., 2006](#); [Strelcyk and Dau, 2009](#); [Hopkins and Moore, 2011](#); see [Moore, 2014, 2019](#) for recent reviews). However, the degradation in TFS processing is correlated with old age as well, independently of the key factors of filter broadening and audibility ([Lorenzi et al., 2006](#); [Hopkins and Moore, 2011](#); [Füllgrabe et al., 2015](#); see also § 17.7.1 for a brief review of presbycusis). Using a nonlinear Wiener-kernel technique<sup>173</sup> to analyze single auditory-nerve units, [Henry et al. \(2019\)](#) obtained the phase locking and envelope coding amplitudes in normal-hearing and noise-induced hearing-impaired chinchillas. Additionally, the normal-hearing chinchillas were treated with furosemide, which induces a temporary and reversible metabolic hearing loss that serves as a model for the endocochlear potential reduction effect that is characteristic in aging. Both impaired groups had a downward-shifted tuning of the TFS, which peaked at lower than normal CFs, in a manner that corresponded to a broadening of the tuning curves in the respective channels. The noise-induced animal group had the mistuning appearing at lower levels than the metabolic-loss group, and it also exhibited a mistuned envelope response, which was normal in the metabolic impairment group. These single-unit results also carried over to frequency-following response (FFR) measurements (see § 9.11), where the balance between envelope and TFS power was distorted in noise-induced hearing-impaired relatively to normal-hearing chinchillas ([Parida and Heinz, 2021](#)). The TFS power was higher in hearing-impaired animals as a result of the low-frequency energy that entered through the mistuned channels. It was suggested that such a response pattern in humans may cause an upward spread of masking that may make speech intelligibility worse. In general, however, it has been difficult to establish a consistent trend of the effect of hearing impairment on the envelope and TFS FFR ([Jacxsens et al., 2024](#)).

Listening in reverberation is a particularly interesting situation for the dual detection (coherent/noncoherent) strategy, since, realistically, only direct sounds can be coherently detected using TFS (§ 9.11). In the limit of a diffuse sound field (§ 8.4.2), the reverberant sound is incoherent and has to be detected noncoherently using envelope cues. In realistic reverberant fields, there may be an advantage in partially-coherent detection, which can target the direct signal and blur the indirect

<sup>173</sup>See reservations about the Wiener-kernel method in footnote 101.

field. In another FFR study, listeners had to identify reverberant spoken digits from the front that were being masked by similarly spoken digits arriving from the sides, which forced the listeners to use interaural time difference (ITD) cues (Ruggles et al., 2012). Envelope cues, as were quantified by  $FFR_{ENV}$ , were more robust to reverberation than TFS cues, but they were more effectively used by younger listeners. Older listeners tended to rely more on TFS cues that were degraded, which lowered their relative digit identification performance. In a different study, reverberant speech intelligibility performance of older listeners correlated only with  $FFR_{TFS}$  of several harmonics around the first formant, whereas envelope cues were not significantly correlated with performance (Fujihira and Shiraishi, 2015).

These two studies underline the importance of TFS cues that are implicated by their erosion with age and, predictably, reverberation. However, they are curious in that envelope detection—the most immediate and less costly strategy for detection in reverberant fields—is weaker in older listeners. This may entail a decrement in the DCN, or other subcortical pathways that are normally responsible for noncoherent detection. This indeed has been found in aged rats that had lower synchronization to sinusoidal AM stimuli and loss of tuning to modulation frequencies (in units with bandpass response) than young rats (Schatterman et al., 2008). Comparable findings about the VCN are not as suggestive, though—they show some age-related decrement, whose consequences are not necessarily detrimental—at least not at low CFs (e.g., Wang and Manis, 2006; Wang et al., 2019). These findings may provide some credence to the hypotheses that the DCN is responsible for noncoherent detection and that it may become impaired with age, whereas the VCN that is specialized in coherent detection is more resistant to aging-related deficits.

### 17.4.2 OHC impairment

Earlier in this work, we tied the phase locking that is measurable on the auditory nerve and higher levels to a hypothetical auditory PLL. It was argued that different parts of the OHC comprise the necessary elements for a generic PLL: a phase detector that is responsible for the  $f_2 - f_1$  distortion product generation (the mechanoelectric transduction channels), a (low-pass) loop filter (the OHC cell membrane), and an oscillator (the hair bundle) (§9.8). The somatic electromotility then provides the loop gain for the closed feedback loop of the PLL. Therefore, it is interesting to find out whether an impairment in the OHCs may lead to degraded phase locking. However, establishing a direct link between normal operation of phase locking and the OHCs has not been done at a level that provides an unequivocal support for the auditory PLL hypothesis, but several interesting findings are indicative that such a connection does exist.

Physiological evidence of degradation in phase locking with hearing loss has been mixed and usually not directly associated with the OHCs, sometimes due to questionable experimental methods or control. Harrison and Evans (1979) were the first ones to test this connection using pure tones as stimuli. They found no impact of selective kanamycin-induced OHC degeneration in guinea pigs on the synchronization as measured in the auditory nerve. These results were contested by Woolf et al. (1981), who pointed at the unprecedented high synchronization index ( $> 0.45$ ) obtained in the tested animals at frequencies of 5–7 kHz (Harrison and Evans, 1979, Figure 2). In findings from similarly treated chinchillas, the phase locking in channels with damaged OHCs showed a decreased frequency range in which phase locking could be observed in measurements of ventral cochlear nucleus (VCN) neurons (Woolf et al., 1981). In that study, though, it is not clear how well the presentation levels in the impaired animals were matched to the normal-hearing controls (Moore, 2014, p. 20). Another study on noise-induced hearing-impaired cats tested the phase locking to tones and to the vowel / $\epsilon$ / embedded in a synthesized syllable (Miller et al., 1997). Using a somewhat nonstandard “synchronization coefficient” measure, the phase locking to tones had about

the same average in the impaired animals, but with a larger corresponding variance, in comparison to normal-hearing controls. In contrast, there was a substantial degradation in the coding of the F2 (1700 Hz) and F3 (2500 Hz) formants of the vowel in the impaired animals. This was the case as the channels became broader and let through a larger number of spectral components that the auditory nerve had to synchronize to, but at a lower power. However, as the tuning curves and frequency selectivity data revealed, the hearing loss had affected both the IHCs and the OHCs, which made it impossible to implicate the drop in phase locking with one hair cell type or the other.

Better methods are found in more recent studies that indirectly tested the relation between phase locking and OHC function. In noise-induced hearing-impaired chinchillas, the maximum phase locking frequency dropped by about 500 Hz, so the envelope became more dominant at lower frequencies (Kale and Heinz, 2010). The authors observed that the phase locking capability was independent of the auditory fibers themselves, but since no decrease was measured in frequency selectivity, it was suggested that the shift in the phase locking range was caused by damage to the IHCs and not to the OHCs. In another noise-induced hearing-impaired chinchilla study, phase locking dropped only for tones in noise but not in quiet (Henry and Heinz, 2012). The amount of decrease depended on the SNR and in some cases brought the impaired synchronization to negligible levels<sup>174</sup>. Auditory filter broadening was shown to negatively correlate with the degree of synchronization strength, which indicates that phase locking is at least partially determined by the periphery, and possibly by the OHCs rather than to the IHCs.

The interpretation of the results that point to a peripheral source of phase locking is made complicated because of the tonotopic mapping distortion and the change in envelope to TFS power balance that has been observed in animals with hearing loss (Henry et al., 2016, 2019; Parida and Heinz, 2021). Nevertheless, the shifts in phase locking were thought to be related to loss of OHC gain in one study (Henry et al., 2019) and in another study to OHC function—as can be seen by a weak DPOAE response (Parida and Heinz, 2021). In human subjects too, the TFS FFR was weakly correlated with DPOAEs in young and old normal-hearing listeners (Anderson et al., 2021). In contrast, auditory nerve afferent function was quantified using the ABR wave-I amplitude, but it was not significantly correlated with phase locking and did not interact with the DPOAE (Anderson et al., 2021; in line with Parker, 2020). Contrary to these findings, neither phase-locking nor envelope synchronization degradation was found in aged gerbils, but rather an increase in spontaneous rates, possibly due to auditory nerve fiber or synapse loss, independent of filter broadening (Heeringa et al., 2020).

Integrating these results is somewhat puzzling, because of their inconsistency. According to the PLL model we indeed expect to see a drop in phase locking with a drop in loop gain that is taken to be caused by OHC function decrement. This may appear as loss of cochlear gain, irrespective of phase locking, though. Additionally, loss of loop gain in the PLL should cause a decrease in its pull-in range, so that off-frequencies relative to the CF may not lock in as effectively and they can lose some of their ability to reject noise. These predictions are roughly in accord with a subset of the studies (Woolf et al., 1981; Miller et al., 1997; Henry and Heinz, 2012). Loss of phase-detection gain is not necessarily dependent on the DPOAE, which is quantified using the cubic distortion product  $2f_2 - f_1$ , rather than the quadratic  $f_2 - f_1$ —the critical distortion product for the PLL. However, if the two product types co-vary, then the weakening of the phase detection product can be deduced from both (Parida and Heinz, 2021; Anderson et al., 2021). Qualitatively, this appears to be the case in normal-hearing listeners from data in Baiduc and Dhar (2018), but no corresponding hearing-impaired data are available to corroborate this reasoning. The results in the gerbil from

---

<sup>174</sup>The drop in synchronization strength is less than 0.1 at -20 dB SNR, according to Figure 2 in Henry and Heinz (2012). However, it is displayed as a constant average drop, which in relative terms can be anything from 10% to 100% of the baseline synchronization.

[Heeringa et al. \(2020\)](#) could not be explained using the auditory PLL framework, unless the aged animals had intact OHC gain that did not disrupt the PLL function, unlike the conclusion that we drew from human data by [Anderson et al. \(2021\)](#). Finally, we generally assume that the type of animal tested is a relatively insignificant factor, although we have to allow for this assertion to be wrong if future data will say otherwise.

### 17.4.3 Neural impairment

In addition to the cochlear PLL, we hypothesized that neural PLLs might exist more centrally than the cochlea. While it is in line with a few modality-general neural models, little evidence could be produced to illustrate it in hearing (§9.10).

When looking at hearing impairments, the lack of substantiating data is even more stark, and only one study was found that specifically tested synchronization impairment in a brainstem tissue as a function of impairment. The robustness of phase locking between the auditory nerve and the bushy cells of the anteroventral cochlear nucleus (AVCN) was studied in brain slice preparations of two strains of young and aged mice with and without age-related hearing impairment ([Wang and Manis, 2006](#)). It was found that while the auditory nerve excitatory postsynaptic potential jitter pattern was unaffected by aging, the spike threshold in the endbulb of Held of the bushy cells was elevated at high CFs in the (aged) hearing-impaired mice. While this finding might have implications primarily for temporal-based localization tasks, it might also affect other coherent detection processing tasks, which rely on the excellent temporal precision of the AVCN for processing. This can theoretically point to loss of phase locking of central origin. All considered, though, the VCN does not appear to be strongly involved in hearing impairment effects, as was also suggested in §17.4.1.

### 17.4.4 Excessive phase locking

In principle, phase locking precision may also be higher than is required for hearing a particular stimulus. As phase locking (or coherent detection more generally) employs a local oscillator, this may be the case if the output of the oscillator is too dominant in the signal chain due to excessive coupling, if it is locked in a self-oscillating mode, or if it is applied to incoherent inputs that do not benefit from coherent detection. These hypothetical dysfunctions are most readily associated with tinnitus.

Heightened phase locking that is associated with tinnitus has been studied primarily in cortical areas (e.g., [Norena et al., 2003](#); [Weisz et al., 2007](#)), whereas subcortical theories of tinnitus relate its existence to hyperactivity—possibly combined with central gain—that usually corresponds to the edge frequencies of cochlear impairment (e.g., [Henry et al., 2014a](#)). In and of themselves, these cortical and subcortical mechanisms are not contradictory, but they seem to constitute a systemic abnormal behavior, rather than a specific impairment of the coherent processing of the system.

An alternative point of view using auditory accommodation is presented in §17.7.3.

## 17.5 Sampling rate and clocking impairments

The maximum sampling rate of the system—auditory nerve spiking that is initially encoded in the cochlear nucleus (CN)—has to correspond to the highest modulation frequency in the stimulus, which is itself determined by the smallest aperture in the system (§14). Too low a sampling rate may cause high modulation components to be undersampled and potentially aliased, leading to blur. Too high a sampling rate can lead to over-detection of high-modulation frequencies that cannot be further recoded and processed by the system, and may be perceived as noise. Alternatively, if a



high sampling rate can be detected, it may lead to sensory information overload, which might lead to non-auditory-specific adverse psychological and cognitive effects (Weisser, 2018, p. 180). The overall sampling rate may indirectly decrease as a result of cochlear impairment, synaptic changes, or deafferentation, as they may all entail a lower-energy transduced product, and therefore a reduced spiking probability either peripherally or centrally. The author is not aware of any documented evidence to a global change in sampling rates in the system—perhaps in accord with the fact that the auditory system is not centrally clocked—as most changes appear to be instantaneous and stimulus driven. Nevertheless, this hypothetical impairment is presented here because long-term global changes to in the spiking rate most are most certainly expected to have perceptible impact on the quality of auditory imaging and, perhaps, on other cognitive variables.

## 17.6 Excessive aberrations

Optical aberrations are the primary cause for some of the most prevalent eye disorders, which in this context are referred to as refractive errors. Notably, myopia (short-sightedness) and hyperopia (far-sightedness) are caused by defocus and can be treated by wearing eyeglasses, contact lenses, or sometimes by going through a corrective surgery. In contrast, hearing disorders are primarily driven by loss of audibility, whose closest, most prevalent, analogous visual impairment may be **cataract**—a leading cause of blindness worldwide—the opacification (clouding; attenuation of light) of the crystalline lens that results in loss of image quality (Chou, 2010). Several types of aberrations were discussed in §15.9 and §15.10 as part of the normal operation of the ear, which may have relevance to impaired operation as well, although probably not as critical as aberrations in vision.

### 17.6.1 Excessive transverse chromatic aberration

Transverse chromatic aberration was indirectly discussed in the context of binaural diplacusis and abnormal curvature of the OHCs (§17.4.2). Different magnification of the two ears combined with a remapped cochlea due to impairment can lead to a different pitch perception of the same stimulus frequency. This explanation may be also applied to a single ear, where pitch is stretched in normal-hearing people (§15.10.1). Thus, an abnormal magnification may skew the pitch stretching curve. If the loss is symmetric, then the effect will be bilateral and will not give rise to an obvious diplacusis. While the stretching effect is generally much more pronounced for pure tones than for complex tones (Jaatinen et al., 2019), such an aberration may potentially impair the perception of musical intervals, melody, and harmony, as well as music enjoyment as a whole.

There is only limited evidence that can attest to an abnormal transverse chromatic aberration in hearing impaired listeners. Two trained musicians with high-frequency moderate-severe SNHL performed a monotic (same-ear) octave matching test for the same ear with pure tones of a fundamental frequency of 125–2000 Hz (Arehart and Burns, 1999, Table 1). The matching became more off-tune closer to the spectral region of loss and at 2 kHz it was about 1.5 semitones up for one subject and 1.5 semitones down for the other. In another study, Huss and Moore (2005b) tested monotic octave matching in listeners with dead regions. The performance was mixed and often erratic, but in general, tones whose frequencies fell within the dead region range were perceived to be of higher pitch than normal. This can be translated to an enhanced stretching effect, and may suggest lower magnification, as may be implied by filter broadening (that leads to a drop in neural group-delay dispersion  $v$  and, consequently, in the magnification  $M$ ). However, in another octave-matching study where the psychophysical tuning curves were evaluated (dead regions not reported), it was noted that there was no clear correspondence between changes in the location of the channel tuning curve peaks and the pitch shifts (Turner et al., 1983).

Overall, the limited available data suggest that pitch perception may be impaired with cochlear hearing loss, in a way that can well correspond to the magnification model, perhaps following from a curvature change of the OHCs, or a change in the filter bandwidth. However the involved changes may not be particularly debilitating.

### 17.6.2 Neural synchrony: Excessive temporal chromatic aberration

If conditions exist in which the neural group-delay dispersion is excessively long, it may result in an audible temporal chromatic aberration, beyond the natural depth of field that was assumed tolerable earlier (§ 15.11.2). Such may be the case when the evoked-ABR wave V appears completely absent. In theory, if the neural signal is very dispersed, then the different channels cannot synchronize to produce what appears as a single peak<sup>175</sup>. Instead, the energy from the different channels will be smeared in time. For example, this appears to be the case in the “poor morphology” patients in [McMahon et al. \(2008\)](#), whose wave-V responses were severely smeared. This effect is the neural dyssynchrony component of the auditory neuropathy/dyssynchrony disorder ([McMahon et al., 2008](#); [Berlin et al., 2010](#)). However, the effects of neural dyssynchrony and poor conductance are often conjoined, so there are no available accounts of the perceptual effects of the dyssynchrony component alone in the experience of listeners who suffer from it.

Other auditory disorders may be less sensitive to neural dyssynchrony. On one extreme of the latency changes, the evoked-ABR of Menière’s disease patients exhibits significantly longer latency than normal ([Donaldson and Ruth, 1996](#)), which may also involve dispersive changes (§ 17.3.2). On the other extreme, the effect may be much smaller in moderate cochlear hearing loss, where the ABR tends to have shorter latency than normal, judging from the derived-band ABR in [Strelcyk et al. \(2009\)](#) that was reviewed in § 17.3.1. If across-channel wave-V latency is on the order of 1–2 ms, then in normal-hearing people it is a difference that may be just noticeable (§ 15.11.2), but probably of little consequence ([Møller et al., 2007](#)). [Stone and Moore \(2003\)](#) tested the effects of frequency-dependent delays, which were applied uniformly to the low-mid frequency bands, while frequencies above 2 kHz were not delayed. Subjects with flat or sloping moderate sensorineural hearing loss were tested aided (with amplification) with closed-fitting and a linear amplification program. When a delay of 4 ms was applied, subjects were only slightly more annoyed with their own-voice sound, and experienced negligible consonant/vowel recognition drop. The effects of delay became more adverse the longer it was, but overall were never dramatically bad.

### 17.6.3 Higher-order aberrations and adjacent-filter phase mismatch

Spherical aberration and coma were the two major higher-order aberrations that were hypothesized to be relevant in hearing. They both entail changes in the phase function away from the center frequency of the channel. Dead regions that characterize profound cochlear loss have already been discussed in § 15.9, where their effects on pitch perception were used as possible evidence for the existence of these aberrations. However, as far as the impairment itself is concerned, the reverse causation is probably more useful: severe aberrations at the flanks of the auditory filter cause some of the abnormal sensations to pure tones and other stimuli that were reported by listeners with dead regions ([Huss and Moore, 2005a,b](#)).

A corollary of off-frequency aberrations may have general relevance to hearing-impaired listeners with broadened filters. In many of the analyses above and in all accounts of cochlear hearing impairments, the effect of filter broadening due to OHC dysfunction is pervasive. However, having narrow

<sup>175</sup>Channel synchrony—a term that is frequency used in this context—should not be confused with a lack of synchronized phase locking within channels, which has a very different effect (see § 17.4).

auditory filters in the normal hearing system can be justified from the system design perspective to minimize higher-order aberrations, and thereby to retain operation within the paratonal approximation. In general, the broader the channel is (relative to its center frequency), the more dominant its aberrating flanks will be with respect to arbitrary stimuli. This should have two main effects on imaging. First, signals that are normally resolved are processed in the same filter, creating modulated patterns that are perceptually unlike the unresolved image. For example, two previously resolved tones may beat together in a broadened channel. In this case, though, if the modulation frequency is not prohibitively high, then the information about the existence of the two tones is conserved, but transformed to a qualitatively different percept (§6.6.2). Therefore, the filter broadening effect on envelope cues (amplitude effects only) is not necessarily detrimental to hearing, if the listener can adapt to the transformed images. This effect is not related directly to off-frequency aberrations.

The second effect on imaging is a direct result of the aberration, once we consider complex envelope cues that include temporal fine-structure. Here, signals whose instantaneous frequency is well within the broadened filter bandwidth may undergo severer aberration than is encountered in the normal cochlea. As was argued in §15.9, the effect of aberration is to distort the phase function of the image and therefore produce blur. Furthermore, if the signal is spectrally positioned so it is analyzed by two filters simultaneously (see the audio examples given in §15.9), then there is high likelihood that the two filters do not have matching phase response that extends to the broadened flanks, and further degradation of the phase information will occur upon combining the images between channels (e.g., during a sweep). This means that with broadened auditory filters, the original information contained in the instantaneous phase spectrum may no longer be available to the listener due to the effect of higher-order aberrations.

Although it is mathematically impossible to truly isolate the effects of envelope and temporal fine structure in arbitrary signals (§6.5.1), there is some data that can support the two aberrating effects above. Lorenzi et al. (2006) showed that young and elderly listeners with moderate flat cochlear hearing loss performed poorly (< 20%) in a vocoded consonant identification task when they had to rely on TFS cues only, unlike normal hearing subjects (> 90%). In contrast, all subjects had no difficulty to use envelope cues only (> 90%). All stimuli were processed in a similar manner, by filtering the speech signals into broad filters, and using the Hilbert transform to extract the envelope magnitude and phase. Using the TFS cues without the disambiguating envelope cues may be severely affected by higher-order aberrations. If the hearing-impaired listeners have broader filters, then their received TFS image may be distorted across channels. Note that these patterns appear to be opposite to findings that older listeners tend to rely on TFS cues more than envelope cues in reverberation, where envelope cues should be more accessible (Fujihira and Shiraishi, 2015 and §17.4).

The relation between TFS sensitivity and frequency selectivity appears to be more complex. For example, young and old normal-hearing and hearing-impaired subjects were tested for their frequency selectivity, TFS, and speech processing (Hopkins and Moore, 2011). When the effect of elevated threshold was removed, there was no significant correlation between the TFS performance and frequency selectivity, as determined by the filters bandwidth. This led the authors to the conclusion that the impaired TFS response is driven by other factors such as loss of fibers, or weakening of phase locking capability. In the same test, monaural TFS scores were correlated with the speech scores in modulated noise masker, while frequency selectivity was correlated with speech scores in unmodulated noise. These results are similar to those found in Strelcyk and Dau (2009), who found correlation between TFS measures and speech-on-speech performance, but not in other masker conditions, including modulated noise. Additionally, no correlation was found between TFS and frequency selectivity. These results may be explained by better ability of subjects to listen in the dips when the masker is modulated, which requires using temporal cues. If aberration is key in

any of these tests, then there may be large individual variability between subjects, which may be independent of hearing status, but could be exacerbated by age. In analogy to vision, this would be entirely normal, as higher-order ocular aberrations show large individual variations that hardly affect the image quality when the pupil is relatively small, but have a much larger effect when the pupil is open and lets in higher modulation frequencies (Liang and Williams, 1997)

Physiological data using Wiener-kernel methods on chinchillas with noise-induced loss reveal that TFS cochlear mapping becomes dissociated from the normal envelope mapping at high frequencies, so their tuning does not match (Henry et al., 2016, 2019; Parida and Heinz, 2021). It suggests that basal place coding becomes unreliable. However, at low frequencies, the tonotopic mapping remains normal, despite broadened filters. More recent data associated this tonotopic distortion with more energy being admitted to the channels off-frequency in animals with noise-induced loss, compared to normal-hearing ones—an effect which was not captured by the traditional filter broadening measure ( $Q_{10}$ ) (Parida and Heinz, 2022). While these data are not presented as phase response, it is possible that they can be reinterpreted as an asymmetrical coma-like aberration in basal sites.

This small sample of studies reveals a complex interrelationship between the basic elements of hearing that are not well understood today. While higher-order aberrations may seem as a far-fetched hypothesis at the moment, it is entirely consistent even with standard filter theory, if phase distortion is considered with respect to TFS. Still, more work will be needed to elucidate these issues and, hopefully, to come up with more succinct explanations of these hearing impairments and their effects.

## 17.7 Impairments of auditory accommodation

The final class of impairments that is considered here is based on the existence of auditory accommodation and its dysfunction. The role of accommodation and possible mechanisms that can facilitate it were discussed in §16. One effect that was hypothesized in §16.3 as a possible behavioral correlate of accommodation is compensation for the level of reverberation when the acoustical context of the environment is given before the target, which was measured in speech intelligibility tasks. When this experiment was repeated with a mixed-age group of listeners with mild–moderate sensorineural hearing loss, the compensatory effect seemed to be intact, despite poorer speech intelligibility performance overall (Zahorik and Brandewie, 2011). However, in general, old hearing-impaired listeners are less sensitive to different reverberation durations (Reinhart and Souza, 2018), and even old normal-hearing listeners show degraded listening in reverberation—even when audibility is controlled for (Nábělek and Robinson, 1982; Helfer and Wilber, 1990; Harris and Reitz, 1985; Gordon-Salant and Fitzgibbons, 1993, 1995; Ruggles et al., 2012; Fujihira and Shiraishi, 2015). We reviewed earlier the possibility that these results have a peripheral cause in the OHCs, but a central cause may be possible too, in line with animal studies that found subcortical correlates in reverberation encoding in the IC (§16.3 and Heeringa et al., 2020).

Two different varieties of accommodation impairments are explored below. The discussion is more inspired by analogy to vision than in previous sections, although some evidence from hearing research is presented as well. The motivation for this is to underscore that ocular accommodation, while a relatively obvious feature of the eye in contemporary perspective, is not at all a trivial process given that it involves optical imaging considerations, a biomechanical variable lens system, a neural control and feedback system, constraints to the pupil and binocular vergence (accommodation) reflex, as well as non-visual factors such as attention. The starting point for understanding accommodation, though, must be image sharpness, so it is difficult to imagine making sense of such a system in the absence of the very notion of the imaging function of the eye. When considering accommodative impairments, this logical chain of processes can only become more opaque if it has serious holes in its

understanding. Therefore, the main purpose of this section is to use the concept of auditory imaging developed in this work and try to unravel potential elements in the rest of the accommodative system that could lead to impairments. Despite the speculative nature of this analysis, it may be more valuable to pursue it than to overlook it completely, in hope that it can advance the understanding of these impairments in the future.

It is noted again that gain accommodation has been largely neglected in this work on purpose, although its hypothetical existence may be well-incorporated into the auditory accommodation framework. Compression is known to be reduced with hearing loss and it is tightly associated with the OHCs, which we used for the time lens and PLL models that are not necessarily independent of compression. Thus, dynamic gain manipulation in the central auditory system, or as is mediated peripherally by the OHCs, will eventually have to be analyzed as a possible mechanism whose range decreases in presbycusis.

### 17.7.1 Neural synchrony: Loss of defocus amplitude

#### Presbyopia and prebycusis

The general findings about age-related effects in reverberation and noise present a notional analogy between prebycusis and presbyopia, which goes beyond the inevitable age-related decline in both auditory and visual acuities.

**Presbyopia** is the most common disorder related to accommodation and is defined as age-dependent subjective loss of the accommodation amplitude from a maximum of about 14 diopters at young age to below three diopters (Charman, 2008). The perceived subjective loss is a combination of objective loss related to the biomechanical changes in accommodation, as well as depth-of-field effects, which can mitigate the associated blur (e.g., by constricting the pupil). The typical onset of subjective presbyopia occurs around the age of 40, but the amplitude (i.e., the dioptric power range of accommodation) is continuously lost with age even when it is not subjectively noticed. These changes are correlated with other age-related changes to the lens, ciliary muscles, zonules, and other parts of the eye. However, there is no consensus at present as for which combination of degradations is the root cause of presbyopia. This impairment often manifests as hyperopia (far-sightedness) that can be corrected for with glasses, at least up to a certain degree of satisfaction.

**Prebycusis** is broadly defined as age-related hearing impairment that can emerge from peripheral, central, and/or cognitive causes (Gates and Mills, 2005; Humes et al., 2012). Co-occurrence of various symptoms of decline in all three types of causes is very common in old age, so several causes may coalesce in actual cases of prebycusis. The peripheral component famously manifests as a sloping high-frequency cochlear loss, which can be measured even in animals that are reared in quiet. The main peripheral effect (called **strial prebycusis**) is caused by age-related metabolic degeneration of the stria vascularis in the cochlea and, consequently, a drop in the endolymphatic voltage that is required for normal OHC amplification. However, the complete impact of prebycusis may be a cumulative effect of noise trauma, ototoxicity, genetic susceptibility, and other factors, so the effect of aging alone cannot always be easily isolated. Presbycusis leads to decline in speech intelligibility in noise and in quiet. Additional temporal processing deficits (e.g., poorer gap detection, time-compressed speech intelligibility, localization) are characteristic in prebycusis as well.

The main question we would like to answer is whether there is an analog in prebycusis to the drop in accommodation amplitude of presbyopia. Whereas ocular accommodation reacts to distance—the main parameter that determines the focal length of the lens—hearing is sensitive to coherence, so we can expect that auditory accommodation is designed to react to it as well. Two mechanisms are particularly pertinent for coherence processing: defocus and phase locking. Defocus is the main imaging property that has been discussed that differentiates acoustic objects according

to their relative degree of coherence. Phase locking to the carrier is required for coherent imaging and detection, which can lead to eventual demodulation of the complex envelope. If either the defocus or the ability to phase lock deteriorates as a result of aging, then the analogy to a drop in the accommodation amplitude may be valid. While some inherent decoherence (noise) in the system was hypothesized as a plausible mechanism of accommodation (§ 16.4.7), uncontrolled application of internal decoherence can have effects that interact with both defocus and phase locking and as such may adversely impact hearing.

### Hypothetical effects of medial olivocochlear reflex impairment

We speculated in § 16.4.2 that the medial olivocochlear (MOC) reflex (MOCR) has two complementary functions that work to weaken the coherent detection. One is a decrement in phase-locking strength, which entails a weakening of downstream TFS representation. It does not necessarily mean that the envelope representation is enhanced, though, although it seems like a reasonable processing strategy. The second MOCR effect is a reduction (or change) of the time-lens curvature that can decrease the defocus of the imaging system that differentiates incoherent and coherent objects. If the curvature  $s$  is independent of other dispersive shifts (i.e., of  $u$  and  $v$ ), then it is expected to improve the incoherent imaging without affecting the coherent imaging. This is because with a smaller defocus, the absolute value of the defocus term  $W_d = \frac{1}{u} + \frac{1}{v} + \frac{1}{s}$  in Eq. 13.53 becomes smaller, which makes the theoretical incoherent MTF closer to the coherent one, in terms of its cutoff frequency (see Figure 13.2). Aging-related accommodation impairment, therefore, can potentially degrade the effectiveness of the time-lens curvature and phase-locking MOCR inhibitory role, which may entail poor adaptation to different degrees of coherence. Such an impairment may be accompanied by a static change to the time lens or the other dispersive elements, which leads to a change in the operating point of the dynamic system.

We therefore predict that auditory accommodation should be associated with loss of defocus amplitude, which can manifest also in strengthening of envelope processing. Furthermore, loss of phase locking may reduce the defocus amplitude from the coherent side. These changes may be driven by decline in the MOC function, although other processes may contribute to it.

### Medial olivocochlear reflex and aging

Even though its exact role is disputed (see § 16.4.2), the interaction of the MOC system with the aged hearing function is robust, which constitutes the most superficial similarity required between presbycusis and presbyopia. For example, changes in the endolymphatic electric potential have a global effect following excitation of the crossed olivocochlear bundle reflex, which directly implicates the OHCs and the amount of intermodulation distortion in the cochlea (§ 16.4.2), at least in the vicinity of the impacted channels (Mountain, 1980). More recently, it has been shown that in aged mice with enhanced cholinergic receptors (of the type that exists in the efferent synapses with the OHCs) the auditory system hardly shows any of the normal age-related deterioration that is found in wild type mice (Boero et al., 2020). This was shown using normal ABR measures and OHC operation (DPOAE amplitudes), as well as only a minimal loss of afferent synaptopathy—all of which were comparable to young mouse controls. All mouse groups were raised in soundproof conditions, so noise-induced effects were ruled out. These results reinforce earlier findings that demonstrated that age-induced decline occurs in parallel in the DPOAE and the ABR of mice and humans (Kim et al., 2002; Jacobson et al., 2003). Also, it was found that a loss of MOC function (suppression of DPOAEs in middle-age groups) preceded the deterioration of the OHCs themselves (loss of DPOAE amplitude). This early MOC loss was more prominent at high frequencies ( $> 4$  kHz) in both normal-hearing humans and mice. In humans, the decline of the contralateral MOCR was evaluated from the weaker DPOAE and click-evoked OAE responses in aged listeners (41–60 years old) (Lisowska

et al., 2014). While the MOCR effect is small (typically 0.5–2 dB), the DPOAE baseline was lower on average, suppression effects were smaller, and also the proportion of DPOAE enhancement events due to contralateral broadband noise increased in the aged group<sup>176</sup>. All effects were stronger at high frequencies (2–3 kHz). Nevertheless, inasmuch as any of those effects translates to speech reception, there seems to be only a negligible MOCR involvement in the improvement of speech intelligibility measures in noise, at least in studies that employed OAEs as a proxy for the detection of MOCR inhibition (Gafoor and Uppunda, 2023a,b).

### Aging-related hearing and accommodation impairments

Aging effects are best observed in studies that separately controlled for both hearing status and aging. Hearing status and age interact in conventional hearing performance measures—speech intelligibility in noise and reverberation, and interrupted speech intelligibility (Gordon-Salant and Fitzgibbons, 1993, 1995), but not in temporal tasks, such as duration discrimination and gap detection where age alone is the dominant factor (Fitzgibbons and Gordon-Salant, 1994, 1996, 2001; Snell, 1997; Snell and Frisina, 2000).

In a different series of studies, the midbrain (FFR) and cortical (magnetoencephalography, MEG) responses to speech in quiet and speech in noise of young normal-hearing, older normal-hearing, and older hearing-impaired listeners were compared (Presacco et al., 2016, 2019). As in the psychoacoustic studies, speech intelligibility scores were correlated with the hearing status of all groups. It was found that the FFR strength for the /da/ syllable—both its transient and its steady-state portions—was higher in young listeners than in old listeners, independently of their hearing status and whether it was measured in quiet or in noise. Interestingly, the low-frequency envelope (1–8 Hz bandwidth) reconstruction fidelity based on the MEG data was higher for the older listener groups. This effect may be caused by an enhanced cortical response that is observed in old age listeners (e.g., an increase in the cortical N1 amplitude, Tremblay et al., 2003). Although we expected an improvement in incoherent processing at a lower subcortical processing level, these cortical changes may be an indication of a downstream effect of such a change that begins with adaptation in upstream auditory processing. In this overarching interpretation, these results could be in accord with accommodative MOCR impairment prediction, which has not been measured in the cited literature<sup>177</sup>.

The loss in defocus amplitude that leads to a reduction in phase-locking fidelity and coherent detection, as well as to changes in the incoherent imaging fidelity, can be thought of as resulting from extension of TMTF bandwidth. Whether this extension has any effect on the image sharpness depends in part on the modulation spectral content of the object and on its degree of coherence. Reduction in coherent detection contrast might be gathered from available studies that measured the differences between young and old normal hearing listeners in broadband and tonal TMTFs, summarized below and in Figure 17.2.

Tonal TMTFs were measured by He et al. (2008a) in young and old normal-hearing groups. The TMTFs were found to be similar in shape in the two groups, but less sensitive overall in the older listeners. The two subject groups had normal audiograms below 4 kHz, but the older group (60–80 years old) had mild losses at 6–8 kHz. Carriers tested were 500 and 4000 Hz—two spectral regions that differently probe the temporal sensitivity of the listeners. While both thresholds were elevated

<sup>176</sup>Enhancement rather than inhibition of the DPOAE is sometimes observed as a result of the MOC activation, although at a lower frequency than inhibition. It has been interpreted as a consequence of constructive interference of the DPOAE and reflection products in the ear (Abdala et al., 2009).

<sup>177</sup>A recent analysis of old normal-hearing and hearing-impaired listener data suggested that “*patients with impaired audibility cannot adjust the corresponding proportion of the envelope and TFS under noise conditions the way that individuals with NH [normal hearing] are able to*” (Hao et al., 2018). While our present analysis is mainly concerned with aging effects and not with cochlear hearing loss, these ideas are close to ours.

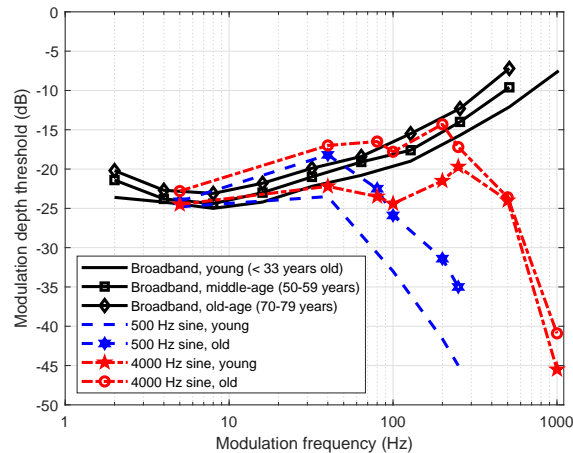


Figure 17.2: Comparison of young and old normal-hearing TMTFs based on broadband carrier data by [Takahashi and Bacon \(1992\)](#) and sinusoidal carrier data by [\(He et al., 2008a\)](#).

by 5–10 dB in the older group throughout the modulation frequency range, the cutoff frequencies, where the modulation are spectrally resolved in adjacent auditory filters, were the same in both groups. However, the threshold at 500 Hz was elevated also in the resolved region (blue curves in Figure 17.2), while at 4000 Hz the thresholds of both groups were the same once the signal was resolved (red curves in Figure 17.2). Therefore, coherent stimuli suffered a loss of contrast in older listeners, which was not necessarily associated with a loss of sharpness. These findings indicate a deterioration in the ability of old-age listeners to use temporal cues. Age-related decline in phase locking that leads to degradation of coherent detection might have impacted the performance, since the modulated stimulus was coherent. Similar studies did not always find consistent results. For example and additional references see [Carcagno and Plack \(2021\)](#), who described a small age effect on loss of contrast in coherent AM detection (with masking noise), as well as less conclusive overall results on the impact of age-related temporal processing degradation.

[Takahashi and Bacon \(1992\)](#) compared normal-hearing subjects of four age groups: young ( $\leq 33$  years), 50–59, 60–69, and 70–79 years old. The older groups had typical presbycusis sloping audiograms with mild losses at 6 and 8 kHz. The incoherent broadband TMTFs of the older subject groups were slightly less sensitive than the young group's TMTF, but the difference was insignificant and may have resulted from higher threshold due to their mild hearing losses (black curves in Figure 17.2). Similar conclusions were obtained in two other experiments carried out in this study of modulation masking and speech intelligibility in modulated and unmodulated noise: while the results were generally poorer for the older groups, an age effect per se could not be statistically established. Either way, unlike what can be predicted by a loss of defocus—not only was an improvement in the incoherent MTF bandwidth not observed in this measurement, but also the sensitivity itself decreased slightly. Nonetheless, as the TMTF broadband depends on across-channel integration, it may not be the right way to tap into channel specific processes, which may require narrowband stimuli that are partially coherent. Still, both aged-listener TMTF studies fit the hypothesis of loss of phase locking precision as a result of aging. This conclusions may be bolstered by a study by [Mamo et al. \(2016\)](#). An isolated effect of degraded phase locking that was simulated by artificially adding jitter to a /da/ syllable and comparing its evoked-speech ABR in young and old normal-hearing listeners (normal pure-tone audiograms below 4 kHz). Adding the jitter effectively equalized the ABR amplitudes (envelope FFR) of the young group to the old group, which was not affected by the additional jitter.



In order for the presbycusis-presbyopia analogy to work, the auditory accommodation hypothesis—that accommodation exists and it relies on adaptation of phase locking (§ 16.4.7) and maybe on adaptive mixing of coherent and incoherent streams (§ 16.4.8)—will have to be substantiated with more evidence. Especially, evidence will have to be established for the causal effects of MOC degradation to phase locking and loss of defocus amplitude.

### 17.7.2 Loss of depth-of-field control

In vision, accommodation impacts the depth of field of the image by changing the focal length. While pupil constriction would be expected to impact it too, aberrations and directional effects that are caused by the pupil edge tend to cancel out most of the depth-of-field changes, at least at large pupil diameters (e.g., [Marcos et al., 1999](#)). As the visual system reacts to the effective image sharpness, accommodation can take advantage of the depth of field and get away with perfect focusing, as long as sufficient sharpness is obtained ([Bernal-Molina et al., 2014](#)).

Since the depth of field determines the extent of the defocus effect, we considered the possibility that the auditory accommodation function is to directly manipulate the depth of field, by controlling the focal time of the lens. This form of release from masking (or **antimasking**, as it is normally referred to in this context) should be then measurable for a combination of probe and masker with or without their MOCR activated. However, results to this effect have been inconsistent in both normal-hearing and hearing-impaired listeners and cannot be accounted for using cochlear gain changes alone ([Jennings, 2021](#)). In animal studies, the antimasking effect is to reduce the firing rate of a tone in noise, which provides a release from adaptation, but not to the extent that the MOCR masking threshold recovers the tone threshold in quiet. A major difficulty in identifying antimasking conclusively, though, is that physiological effects of reduction in suprathreshold masking in animals do not necessarily map to psychoacoustical effects in humans, where the MOCR stimulation is inferred indirectly using OAEs. Another difficulty is that listeners whose MOC was sectioned do not necessarily have hearing performance that clearly shows degraded forward masking performance. For example, one subject with presumed severed efferents had relatively normal tone-on-tone masking recovery ([Zeng et al., 2000](#)). Four subjects in the same study had reduced simultaneous masking overshoot effect in their operated ears, in which the broadband masking threshold of a probe that appears immediately after onset tends to be higher ([Zwicker and Scharf, 1965](#)). In contrast, cochlear implant users, whose OHCs and hence MOCR are bypassed by directly stimulating the auditory nerve, exhibited normal overshoot effect of sinusoidal AM tones in noise, which suggests that the masking effect is caused by a central mechanism ([Marrufo-Pérez et al., 2019](#)).

Thus, unfortunately, at present we cannot rely on any specific data that may confirm or disconfirm our speculation that the auditory depth of field may be accommodated to by the MOCR. Like several other hypotheses brought up here, it is in itself linked to several other strong hypotheses, which will have to be confirmed or rejected first, regarding the time lensing in the auditory system and the validity of the MOCR in accommodation.

### 17.7.3 Tinnitus as an accommodative impairment

This section is an extension of the hypothesis put forth in § 16.4.7—that fast and selective adaptation of the degree of synchronization to different elements in the stimulus is a plausible mechanism to achieve accommodation. It was further suggested that some forms of tinnitus (mainly transient) may be an excessive manifestation of this kind of accommodation.

Tinnitus research is too vast to allow for all but a brief overview in the present work, especially given the uncertainties involved and the number of proposed theories, but see [Møller et al. \(2011\)](#),

Eggermont (2012), Eggermont et al. (2012), Baguley et al. (2013), Henry et al. (2014a), Roberts and Salvi (2019), and Henton and Tzounopoulos (2021) for recent reviews. After a very brief introduction to tinnitus, the hypothetical connection with accommodation will be explored to include additional processes, by way of analogy to vision.

Subjective tinnitus is characterized by the perception of phantom sounds, which are the result of abnormal neural activity that has no acoustic correlates inside or outside the body (Eggermont, 2012). Subjects with tinnitus have enhanced synchronized activity between various circuits in the thalamus and cortex, in response to an increase in spontaneous firing rate that is caused by a spontaneous activity along the auditory pathways (Eggermont, 2012, Chapter 9). There are many possible causes for tinnitus (e.g., noise- or music-induced hearing loss, cochlear insults, ototoxicity), but none that is either necessary or sufficient for it to develop. Therefore, many tinnitus cases have an unknown etiology (Kleinjung and De Ridder, 2011; Eggermont, 2012). Nevertheless, it is commonly thought that an initial degree of cochlear or neural impairment is the underlying cause, even if it goes undetected with normal clinical techniques (Norena, 2011). Although “ringing” is the most common descriptor of the perceived tinnitus, the full range of sounds that can be perceived by tinnitus patients is very broad and no universal perception exists (Stouffer and Tyler, 1990). Accordingly, the correspondence between the nature of the phantom sounds and the etiology is not universal either, but for hearing-impaired listeners the tinnitus pitch spectrally corresponds to the impaired range (Norena et al., 2002).

In general, tinnitus may have either a peripheral or a central origin, but it appears that it is essentially a central response to an impairment in the periphery, in the inner ear (some injury to the OHCs, IHCs, and/or the auditory nerve) (Jastreboff, 1990; Schaette and Kempster, 2006; Norena, 2011). According to this model, several central gain mechanisms that exist along the auditory system can compensate for a lack of activity in the impaired channels by a gain increase—itsself a result of homeostatic plasticity (e.g., Hutchison et al., 2023) that maintains a stable level of neural activity and efficient coding through various neuroplastic mechanisms. Tinnitus is then a side-effect of spontaneous activity amplification by the hyperactive system, which is perceived as noise. However, the gain hypothesis has been recently challenged as it was observed that gain is correlated with age and not with tinnitus. so that not all people with high gain experience tinnitus (Johannesen and Lopez-Poveda, 2021). In another line of models, the dorsal cochlear nucleus (DCN) has been repeatedly implicated as the specific source of tinnitus, although it most likely does not sustain it (Henton and Tzounopoulos, 2021). Tinnitus-related hyperactivity has been observed in the VCN as well (Norena, 2011; Henry et al., 2014a). The IC, medial geniculate body (MGB), and primary auditory cortex (A1) may all be implicated in a cascade or feedback process that can lead to tinnitus, which appears to emerge at lower processing levels and involve some degree of neural hyperactivity (Henton and Tzounopoulos, 2021).

Just as accommodation normally functions with a near-triad that (nearly) reflexively combines vergence and pupil constriction (see §16.2), so does synchronization accommodation may work in tandem with other mechanisms. Because these processes are interconnected, the failure of one can lead to the failure of the others in unobvious ways. This can be illustrated with an analogy to two ocular accommodation disorders that are caused by incorrect central control of the accommodative reflex muscles. The first disorder is **refractive accommodative esotropia**—a form of **strabismus** (crossed eye), which is itself the result of an uncorrected hyperopia (far-sightedness) (Lembo et al., 2019). This condition occurs when the visual system tries to compensate for the defocus by using focal accommodation, but then also involuntarily converges because of the accommodation reflex. This problem is prevalent in infants and can usually be corrected with glasses for hyperopia or, if it fails, with a surgery. The second disorder is the **spasm of accommodation**—the permanent contraction of the ciliary muscles, which blocks any possibility for accommodation and thereby leaves

the patient effectively myopic or hyperopic (Goldstein and Schneekloth, 1996). Once again, reflexive side-effects exist in the form of pupil constriction and/or vergence, all are found in different and inconsistent amounts among patients. Both disorders reveal nontrivially coupled symptoms, which require the joint understanding of the refractive problem and the accommodation system in order to be diagnosed and treated successfully.

The analogy between these two ocular disorders to tinnitus is not an obvious one, because the accommodation principles of the temporal and spatial imaging systems are fundamentally different. However, to the extent that tinnitus points to an impaired signal processing in the auditory system that is also part of a reflexive system, then the accommodative impairments of the eye may provide some inspiration for understanding the causal chain of auditory impairments. For example, tinnitus is often comorbid with cochlear impairments, and central gain was proposed as a compensatory mechanism for the loss of signal strength. But this may also lead to **hyperacusis**—the abnormal sensitivity to loudness in otherwise normal hearing—whereas tinnitus itself also involves excessive synchronization. Thus, synchronization and gain may be two processes that are tied by a reflex, which may also include any of the other mechanisms discussed in §16.4. The two may or may not be tied also to the MOCR.

In § 17.7.1, the effect of presbycusis on accommodation was hypothesized to be the loss of defocus amplitude, which is caused by a reduction in phase locking precision. In this subsection the possibility that tinnitus is uncontrolled phase-locking synchronization has been discussed. The two impairments need not be mutually exclusive. If the defocus amplitude is decreased, tinnitus may be a compensatory makeup shift in the available defocus range, which recalibrates the operating point from which synchronization can be accommodated (see Figure 9.5 for a cartoon illustration of this operating point). Indeed, tinnitus and presbycusis typically occur together. Histopathological analysis of the temporal bones of two groups of presbycusis patients with and without tinnitus revealed that the tinnitus group had a greater loss of OHC in the basal and middle turns of the cochlea and more atrophied stria vascularis in the basal turn (Terao et al., 2011). Thus, it may support the hypothesis that tinnitus may be caused by a central mechanism in order to compensate for a deficiency in the peripheral signal. If the OHC function is to produce gain, then central gain compensation may be indeed likely. Our interpretation of the OHC as the starting point of phase locking would imply that phase-locking compensation may be produced centrally, which can then appear as self-oscillations that produce tinnitus.

Finally, it is curious to note that several obscure disorders may point at systemic connections between the assumed auditory and the known visual accommodative reflexes. For example, tinnitus can interact with vergence (Yang et al., 2010) and saccadic movements (Lang et al., 2013), while some hyperacusis cases are coupled with hypersensitivity to light (Philips and Hunter, 1982; Woodhouse and Drummond, 1993). Additionally, there are documented cases of tonal tinnitus that are induced by gaze that is shifted away from the center, which occur following surgical deafferentation of the auditory periphery after the excision of lesions from the cerebellopontine angle (the area between the pons and the cerebellum; Cacace et al., 1994; Mennink et al., 2020). These obscure impairments may point to deeper, albeit weaker, connections between the various elements in the auditory and visual systems.

## 17.8 Conclusion

In this chapter we explored the hypothetical connections between known hearing impairments and several dispersive, phase-locking, and accommodative dysfunctions. Given the amount of unknowns about temporal imaging and the relatively limited empirical data that were found to test against the predictions, some of the hypothesized disorders are highly speculative and the overall picture

is somewhat patchy. Additionally, the relevant analyses have not been consistent throughout and it appears that some disorders may be more plausible than others. For example, a certain form of age-related accommodation impairment seems to be relatively compelling, given the known age effects on the MOCR and on the envelope and TFR processing balance changes. Excessive higher-order aberrations for broad auditory channels seem compelling as well, as they can cause various undesirable distortions, but at this point they are much more difficult to establish. Other dispersive changes in the cochlear and neural parts of the auditory system are not clearly implicated in great changes to performance, so they may be difficult to establish as well. Finally, dysfunction of the OHCs has already been implicated in a large range of hearing impairments. On top of these we stacked changes to phase locking and time lensing, which can have complex effects that likely interact with the global processing strategy of the system and are not straightforward to probe.

Even though the above analysis may appear premature given the current state of knowledge, it also seems worthwhile. The intention behind this speculative chapter has been to apply the ideas that were developed in the preceding chapters in hope to shed some light on hearing impairments that have resisted a compelling theoretical account, and hence—a satisfactory treatment. Ideally, this discussion can motivate further explorations of these topics that can contribute to the eventual resolution of at least some of the most persistent hearing disorders.