# DAT335 – Music Perception and Cognition
# Cogswell Polytechnical College
# Spring 2009

## Week 3 – Class Notes

### Frequency Selectivity, Masking, and Critical Bands (cont...)

## Profile Analysis

It has been proven experimentally that for stimuli without distinct envelope fluctuations, subjects are able to compare the outputs of different auditory filters to enhance the detection of a signal. In this case, the ability to detect an increment in the level of one component in a complex sound relative to the level of others was investigated. Results showed that subjects were able to detect the changes in relative level of the signal of only 1-2 dB. This small threshold cannot be obtained by monitoring the output of a single filter. It is believed that subjects perform this task by detecting a change in the shape, or profile, of the spectrum of the sound; thus the name “profile analysis”. Furthermore, subjects can compare the output of different auditory filters and can detect when the output of one changes relative to that of others, despite an overall level variation.

Profile analysis is said to be most effective under the following conditions:

1. The background or complex sound has a large spectral range.
2. There are many components within the spectral range.
3. The signal falls within the frequency range of the background.
4. The level of the signal component is similar to or slightly above the levels of the components in the background.
5. The background is composed of components with equal levels rather than levels which differ from component to component.

Because the first two aspects of profile analysis are quite similar to CRM. It could be argued that profile analysis is a special case of CMR, where the masker fluctuates randomly in level across stimuli but not within stimuli.

The fact that we can hear changes in the spectral shape shows that we can determine the timbre of a sound and that we can recognize vowels, regardless of the level of those sounds. Functionally, these actions are the same as profile analysis.

## Nonsimultaneous Masking

Type of masking that occurs when the signal is presented just before or after the masker. Two basic forms of nonsimultaneous masking can be distinguished: 1) backward masking, the probe (brief test signal) precedes the masker; and 2) forward masking, the probe follows the masker.

Forward masking has the following properties:

1. Forward masking is greater the nearer in time to the masker that the signal occurs.
2. The rate of recovery from forward masking is greater from higher masker levels. The masking decays to zero after 100-200 ms.
3. A given increment in masker level does not produce an equal increment in the amount of forward masking.
4. The amount of forward masking increases with increasing masker duration for durations up to at least 50 ms.
5. Forward masking is influenced by the relation between frequencies of the signal and the masker.

While the basis of forward masking is not clear, it is thought that four factors may contribute:

1. The response of the BM to the masker continues for some time after the end of the masker. This effect is known as “ringing”.
2. The masker produces short-term adaptation or fatigue in the auditory nerve or at higher centers in the auditory system, which reduces the response to a signal presented just after the end of the masker.
3. The neural activity evoked by the masker persist at some level in the auditory system higher than the auditory nerve, this persisting activity masks the signal.
4. The masker may evoke a form on inhibition in the central auditory system, this inhibition persists for some time after the end of the masker.

## Evidence for Suppression from Nonsimultaneous Masking

The response to a single tone of a given frequency can be suppressed by another tone with a different frequency, resulting in what is known as two-tone suppression. For complex signals, similar phenomenon can occur and are given the general name of lateral suppression or suppression. This behavior can be described as strong activity at a given CF can suppress weaker activity at adjacent CF's.

Nonsimultaneous masking tests are used to reveal the effects which can be directly attributed to suppression. Consider a subject that listens to a signal at a given frequency that is alternating with a masker at that same frequency, and a measurement is made of the pulsation threshold (level of an alternating signal and masker at which there is a change in perception from continuous to pulsating). Then, a second tone is added to the masker and the pulsation threshold is measured again. It will be noted that the addition of the second tone produced a reduction in the pulsation threshold, attributed to the suppression of the signal by the masker's second component. If the signal is suppressed, then there will be less activity in the frequency region around the signal's frequency. As a result, there will be a drop in the pulsation threshold. The masker's second tone is a more effective suppressor is its more intense and lies above the signal's frequency. Further mapping using the same technique, show that the regions found were similar to the (two-tone) suppression areas observed in single neurons of the auditory nerve.

Also, nonsimultaneous masking shows that frequency selectivity is greater than in simultaneous masking. PTC's are obtained by determining the level of a masker required to mask a signal as a function of masker frequency. The PTC's determined in forward masking show sharper curves (especially in the high frequency side) than those obtained from simultaneous masking tests. In nonsimultaneous masking, suppression does not affect the signal. Meaning that for maskers with frequencies above or below that of the signal, the effect of suppression is to sharpen the excitation pattern of the masker. Therefore, resulting in an increase of the masker level needed to mask the signal (the masker is less effective). The suppression levels is shown as an increase of the PTC slopes.

The PTC measured in nonsimultaneous masking is likely to be closely related to BM tuning curves and neural tuning curves.

## Frequency Selectivity in Impaired Hearing

Frequency selectivity can be impaired by damage to the cochlea. The tuning of the BM is dependent on its physiological state. However, damage to the tuning can be reversed if the damaging agent is removed quickly enough. Otherwise, the tuning damage can be permanent.

Studies in human frequency selectivity show that for people with cochlear damage, such as degenerative hearing loss, noise-inducted hearing loss, otosclerosis, and Ménière's disease, show a flat PTC profile than those who do not not have cochlear damage (normal hearing, and conductive hearing loss). This indicates that loss in frequency selectivity is associated with cochlear damage.

Loss in frequency selectivity has several perceptual consequences. First, there is greater susceptibility to masking by interfering sounds. This is because an impaired ear has broader auditory filters, thus they pass more noise through it and the detectability of th signal is greatly reduced. As a result, detecting sounds, like speech, tend to be more difficult in noisy situations.

Second, there is a difficulty in the perceptual analysis of complex sounds. When frequency selectivity is impaired, the ability to detect differences in the spectrum of the sound, and as a result the timbre, is reduced. For an impaired listener it becomes more difficult to distinguish different vowels or different musical instruments. In this case, hearing aids do not help overcoming these difficulties. While these aids can make sounds more audible, they do not offer any correction of frequency selectivity impairment.

# The Perception of Loudness

Loudness can be defined as the attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud. The human ear has a remarkable range of sound intensities. The ratio of intensities, the intensity of the loudest sound we can hear without damage compared to the faintest, corresponds to 1,000,000,000,000:1.

The measurement of loudness can be problematic since it is a subjective quantity and as such cannot be measured directly.

**Loudness Level and Equal-Loudness Contours**

Loudness level is a measure of loudness that indicates how intense a 1000 Hz tone must be in order to sound equally loud, instead of indicating how loud a tone is. This procedure is rather straight-forward. In order to determine the loudness of a given sound, the subject must adjust the level of a 1000 Hz tone until it has the the same loudness as the test sound. The 1000 Hz tone and the test tone are alternately played. The level of the 1000 Hz which gives equal loudness is the loudness level of the test sound, and it is measured in *phons*. By definition, the loudness level of a 1000 Hz tone is equal to its sound pressure level in dB SPL.

This test can be repeated for different frequencies of a sinusoidal test sound, so as to generate an equal-loudness contour, otherwise know as Fletcher-Munson curves. The shapes of these curves vary according to the method used and across studies. However, although a standard exists for these curves, there is no agreement to determine a set of “correct” values.

Equal-loudness contours have a similar shape to MAF curves, but they tend to become flatter at higher frequencies.

There are various implications for the reproduction caused by the shape of the equal-loudness curves. First, the relative loudness of the different frequency components in a sound changes as a function of the overall level. Meaning that the “tonal balance” is altered with a change of loudness. Human hearing is more sensitive to low frequencies at high intensities, as well as being less sensitive to low and high frequencies at low intensities.

The shape of equal-loudness curves have been used in the design of sound level meters (these devices attempt to give an approximate measure of loudness of a complex sound). However, they do not directly measure and sum the intensities of different frequencies; rather, they employ weighting networks that weights the intensity at each frequency according to the shape of the equal-loudness contour before summing intensities across frequency. These weighting networks are poor estimates of the shape of an equal-loudness contour.

Sound level meters have several inherent problems. First, they can only be reliably used for steady sounds of relatively long duration. Transient sounds do not provide enough of a response so as to give a subjective impression of its loudness. Second, the do not provide a satisfactory way of summing the loudness of components in widely separated frequency bands.

The loudness of a complex sound, with a given amount of energy, depends on whether its energy is contained within a narrow range of frequencies or is spread over a wide range of frequencies.

**The Scaling of Loudness**

In order to derive loudness scales, two methods are commonly used. The first, magnitude estimation, presents sounds with different levels, and the subject is asked to assign a number to each one according to its perceived loudness. The reference sound, called modulus or standard, is presented, and the subject judges the loudness of each test sound relative to that of the reference. The second method, magnitude production, requires that the subject adjusts the level of a test sound until it has a specified loudness, either in absolute terms or relative to that of a standard.

It has been suggested that the perceived loudness *L* was a power function of the physical intensity *I*:

$$L = kI^{0.3}$$

*k* is a constant dependent on the subject and the units used.

Also, it has been proposed that the unit of loudness is the *sone*. One sone is defined, abitrarily, as the loudness of a 1000 Hz tone at 40 dB SPL. However, at low levels the loudness change is more rapidly with sound level than implied with the previous equation.

Power law relationships between intensity and loudness have been confirmed in numerous experiments using different techniques; however, they are heavily criticized since these techniques are susceptible to bias effects. The results of these experiments are affected by factors such as range of stimuli presented, first stimulus presented, subject instructions, range of permissible responses, symmetry of the response range, and other factors related to experience, motivation, training, and attention. Consistent results are achieved only though the averaging of numerous subjects.

Overall, it is difficult to judge the loudness of a sound. Its perception is affected by the apparent distance of the sound source, the context in which it is heard, and the nature of the sound. Simply put, we are trying to make an estimate of the properties of the source itself. Obviously, measuring a sensation is a difficult and unnatural process.

**Models of Loudness**

The mechanisms underlying the perception of loudness are not fully understood. However, a common assumption is that loudness is somehow related to the total neural activity evoked to a sound. If this is true, the loudness of a sinusoidal tone will be determined not only by the activity in neurons with CF close to the frequency of the tone, but also by the spread of the activity to neurons with adjacent CF's. In other words, loudness depends on a the total sum of neural activity across different frequency channels. Some proposed models have incorporated this concept. Usually, they follow this basic structure:

Stimulus → Fixed filter for transfer of outer/middle ear → Transform spectrum to excitation pattern → transform excitation pattern to specific loudness → calculate are under specific loudness pattern

The overall loudness of a given sound, in sones, is assumed to be proportional to the total area under the specific loudness pattern. The area is approximately proportional to the total neural activity evoked by a sound.

**The Effects of Bandwidth on Loudness**

If a complex sound with a fixed intensity (or energy) has a bandwidth *W* that is lower than a certain bandwidth, or critical bandwidth of loudness $CB_L$, then the loudness of the sound is almost independent of *W.* The sound will be judged to be as loud as a pure tone or narrowband noise of equal intensity lying at the center frequency of the band. However, if *W* is increased beyond the $CB_L$, the loudness of the complex sound begins to increase.

For a given amount of energy, a complex sound is louder if its bandwidth is greater than one $ERB_N$, than if its bandwidth is less than one $ERB_N$.

The reason loudness remains constant for bandwidths less than one $ERB_N$, can be understood by considering how the specific loudness patterns change with bandwidth. For both excitation and loudness patterns, an increase in bandwidth up to $CB_L$ causes a flattening of the tips, as well as becoming broader. However, the total area remains almost constant. When the bandwidth is increased beyond the $CB_L$, the increase of the skirts is greater than the decrease around the tip, and as a result, the total area and the predicted loudness increases. This increase depends on the summation of specific loudness at different CF's, this the increase in loudness is described as “loudness summation”.

At low levels, the loudness of a complex sound is more or less independent of the bandwidth. At these levels, specific loudness changes rapidly with excitation level, and so does loudness. As a result, the total area under the specific loudness pattern remains almost constant as the bandwidth is increased. When a narrowband sound has a very low sensation level, and if the bandwidth is increased so as to maintain a constant energy level, the energy in each critical band becomes insufficient to make the sound audible. Near the threshold, loudness must decrease as the bandwidth is increased for a small value. As a result, if the intensity of a complex sound is increased slowly from a subthreshold value, the rate of growth of loudness is greater for a wideband sound than for a narrowband sound.

**Temporal Integration of Loudness**

It is generally agreed that at a given intensity, loudness increases with duration for durations up to 100-200 ms. For durations up to 80 ms, constant energy leads to constant loudness.