



Lecture 2-10: Complex Sounds in the Auditory Periphery

Overview

1. **The auditory periphery as a filterbank.** A filterbank is just a set of bandpass filters which cover some range of frequencies by varying their centre frequencies across the range. You have come across this idea before when we discussed the construction of spectrograms; you may also be familiar with **graphic equalisers** which are based upon a filterbank. Filterbanks are a good way of thinking about how the ear analyses **complex sounds**, that is, sounds with energy at more than one frequency, see figure 2-10.1. The mechanical properties of the basilar membrane divide out the different frequency components of the sounds to different regions. The nerve cells then signal information about the total amount of energy in the signal that falls into each region. It is important to appreciate that one region may receive energy from more than frequency component in the signal, and hence it responds to energy over a frequency range, just like a band-pass filter, see figure 2-10.2.
2. **Characteristics of the auditory filterbank.** In comparison to a spectrogram, the auditory filterbank has two special characteristics: firstly the spacing of the centre frequencies reflects the non-linear assignment of frequency to place found in the tonotopic mapping. That is to say that 1mm of the basilar membrane at the base extends over a wider range of frequencies than 1mm of the basilar membrane near the apex. In contrast, a spectrogram maps Hertz values linearly to mm up the page. Secondly the filter bandwidths based on the properties of the basilar membrane analysis vary as a function of frequency, with lower frequencies having narrower bandwidths than higher frequencies, see figure 2-10.4. In contrast, a spectrogram has the same filter bandwidth at all frequencies.
3. **Comparing a Linear Spectrogram with a Cochlear Spectrogram.** A **narrow-band** spectrogram has equally spaced filters of 45Hz bandwidth and is useful for displaying the individual harmonics of voiced sounds. A **wide-band** spectrogram has equally-spaced filters of 300Hz bandwidth and is useful for displaying excitation pulses and formant resonances. A **cochlear** spectrogram has approximately logarithmically-spaced filters with bandwidths that vary between about 35Hz (at 100Hz) to about 1100Hz (at 10kHz), see figure 2-10.5. This kind of spectrogram displays harmonics at low frequencies and striations and formants at high frequencies. This explains how the ear can be sensitive **simultaneously** to detailed low frequency spectral structure and to high frequency temporal variation.
4. **Loudness, Pitch and Timbre.** Loudness is related to the overall quantity of nerve activity. Pitch is related to both time and place cues in the low frequency region and to time cues alone in the high frequency region, see figure 2-10.3. Timbre can be related to many different aspects of a sound (including the way that its amplitude changes with time, such as the attack and decay of musical notes, known most generally as its **temporal envelope**). However, one very important special aspect of timbre, relevant to the perception of vowel quality, is mostly related to information about the shape of the spectral envelope coming from the auditory filterbank analysis (place analysis).

Reading

At least one of:

- Yost, Fundamentals of Hearing: An introduction (4th edition), Chapter 11, pp165-172; Chapter 13, pp 193-199). *Good description of masking.*
- Moore, Introduction to the Psychology of Hearing (5th edition), Chapter 3, Sections 1-5; Chapter 5. *Thorough but tough going.*

Learning Activities

You can help yourself understand and remember this week's teaching by doing the following activities.

1. Write an explanation of the conceptual model of the auditory periphery as bank of filters. Include experimental evidence for the model and a discussion of how the model explains the perception of the timbre of complex sounds independently from their pitch.
2. List ways in which the filters in the auditory system behave like the filters in a spectrograph, and ways in which they behave differently.
3. Use the web to research how hearing impairments affect cochlear functioning. Summarise the effects of those impairments on the auditory filterbank model.

If you are unsure about any of these, make sure you ask questions in the lab or in tutorial.

Reflections

You can improve your learning by reflecting on your understanding. Here are some suggestions for questions related to this week's teaching.

1. Why do auditory filters have a greater bandwidth at higher frequencies?
2. What parts of a complex periodic sound are separated out by the low-frequency auditory filters? What is different at high frequencies?
3. Sketch a typical spectrum of a vowel and the equivalent excitation pattern after auditory filtering.
4. Given that auditory filtering is typically broadened in sensori-neural hearing loss, what implications might this have for the perception of phonetic contrasts?

Examination questions

These are questions from past exam papers. You may like to write outline answers to these, or to discuss them in tutorial.

1. Describe the 'filterbank' model of the auditory processing of complex sounds. Explain how the filters arise in the hearing mechanism, and compare their properties with the filters used in the production of wide-band and narrow-band spectrograms. [2005/6]
2. Explain how the processing of sound in the peripheral auditory system can be modelled with a bank of bandpass filters. Outline how information about the pitch and timbre of a complex periodic sound is represented in the output of the filterbank. [2009]
3. Describe, using diagrams, how the frequency of a pure tone (sinewave) is analysed by the cochlea and represented in its neural output. If the pure tone were replaced by a complex tone of the same fundamental frequency, what aspects of the neural pattern would change and which would stay the same? [2009/10]

Figure 2-10.1 Excitation Pattern for a Vowel Sound

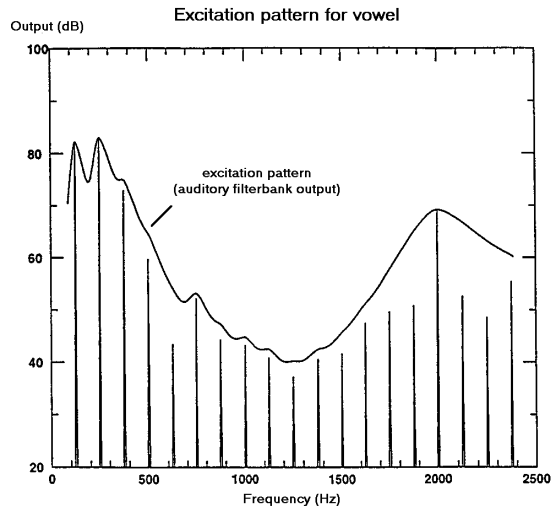


Figure 2-10.2 Auditory Periphery as a Bank of Filters

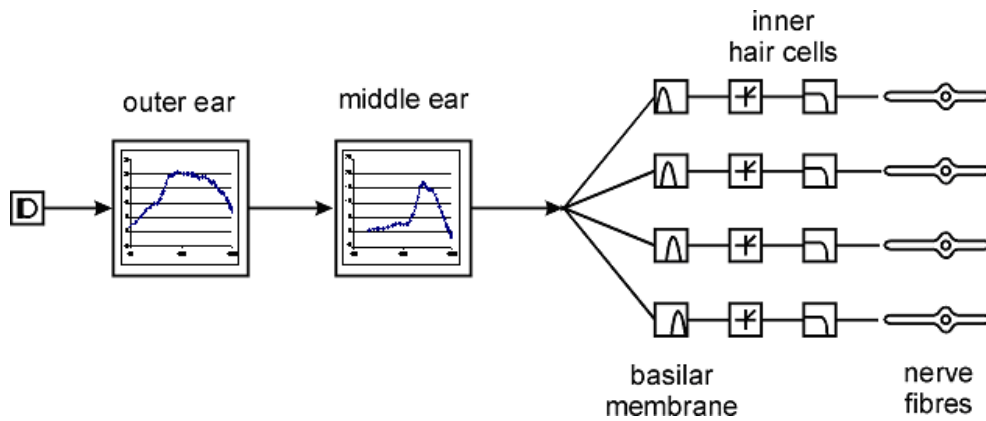


Figure 2-10.3 Auditory Filterbank Analysis of a Complex Tone

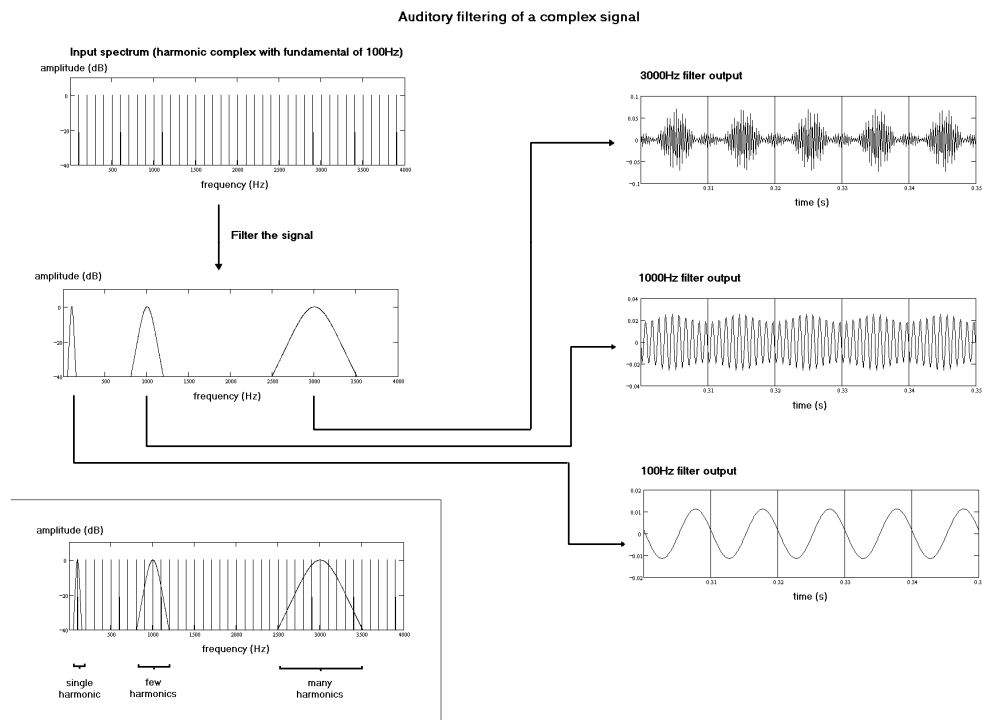


Figure 2-10.4 Auditory Filter Bandwidth

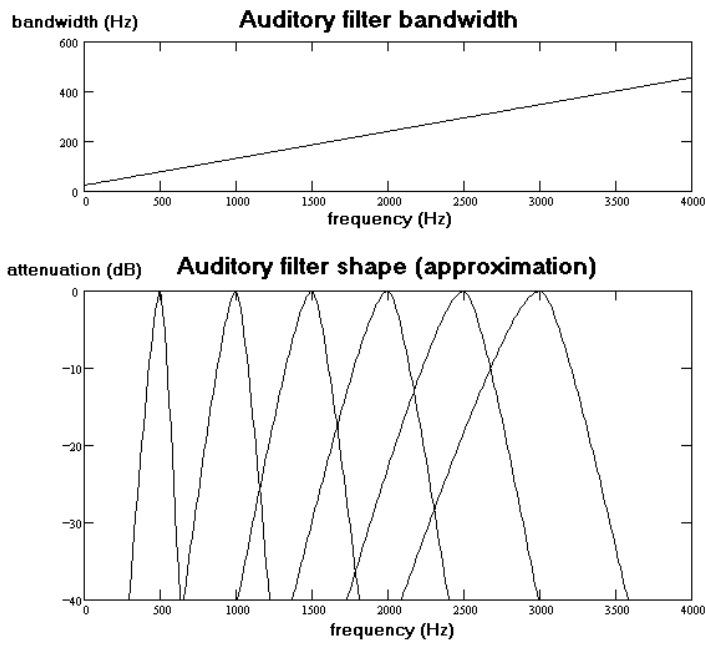
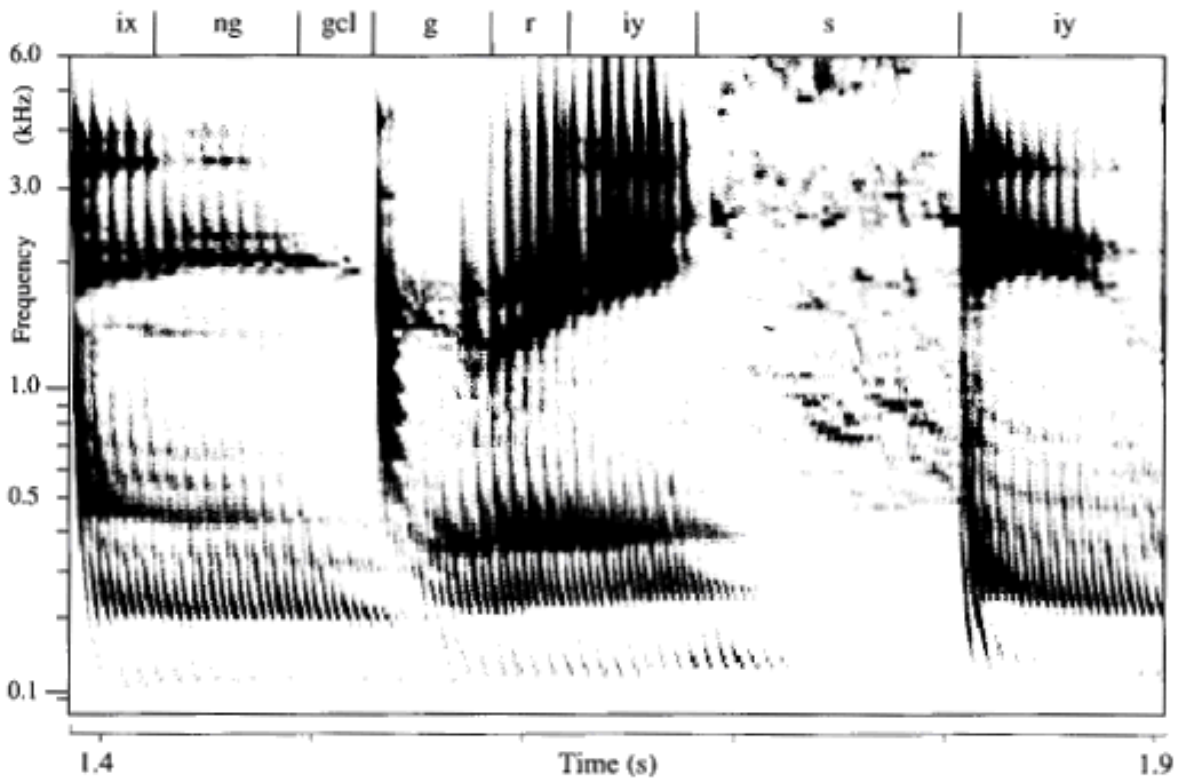


Figure 2-10.5 Cochlear Spectrogram



From: Giguere, C., & Woodland, P. C. (1993) Speech analysis using a nonlinear cochlear model with feedback regulation, in *Visual Representations of Speech Signals*, edited by M. Cooke, S. Beet, & M. Crawford (John Wiley & Sons, Chichester).

Lab 2-9 Cochlear Simulation

Introduction

The cochlea is a complex bio-mechanical structure which performs extraordinary feats of sound sensitivity and selectivity. To gain an understanding of the cochlear encoding of loudness, timbre and the pitch of pure and complex tones, we use a computer simulation. The simulation is simplified but exhibits many aspects of the behaviour of a real cochlea.

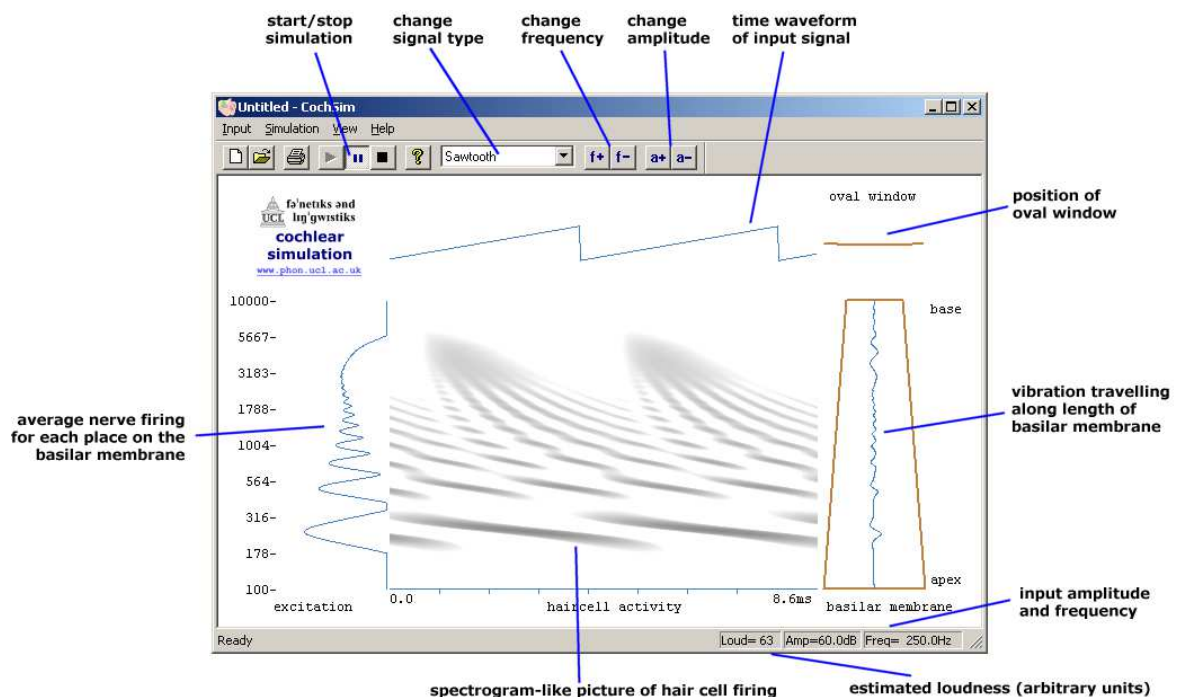
In this experiment we look at how this 'model cochlea' reacts to a range of simple stimuli and hence gain a better understanding of some aspects of cochlear function, including: place and temporal coding, the band-pass filter analogy, the determination of fundamental frequency and the phenomenon of beats.

Learning Objectives

By the end of the laboratory session you should:

- understand what is meant by place and temporal coding of the pitch of pure tones
- understand how changes in the amplitude or the frequency of a pure tone can be detected in changes in the excitation pattern
- understand how the fundamental frequency of a complex tone is encoded in the temporal pattern of the excitation
- understand how the band-pass filter analogy can be used to predict the excitation pattern for a complex sound
- understand what causes 'beats'.

Simulation Display



Observations

1. Put a sinewave through the simulation. Change its frequency with the f+/f- buttons and its amplitude using the a+/a- buttons, observing the effect on the simulation. Make sure you understand what each of the graphical components of the display represents before tackling these questions:

a) Where can you see the effect of 'place' coding of frequency in the simulation?

b) What property of the basilar membrane causes this mapping from frequency to place?

c) Where can you see the effect of 'temporal' coding of frequency in the simulation?

d) What property of the haircells in the organ of Corti causes this patterning?

e) Set the sinewave frequency to about 4000Hz. Why does the temporal patterning disappear?

f) Find the threshold level for a sinewave at 250Hz and at 4000Hz. (i.e. find the input dB level that gives a loudness just greater than zero). Why are these thresholds not the same?

Threshold at 250Hz = _____

Threshold at 4000Hz = _____

2. Process a sawtooth waveform with a fundamental frequency of 250Hz and an input amplitude of about 60dB.

a) Why are the harmonics not equally spread along the basilar membrane?

b) How could the fundamental frequency of the sawtooth be determined by the place coding mechanism?

c) How could the fundamental frequency of the sawtooth be determined by the temporal coding mechanism?

d) Process a sawtooth waveform missing the first harmonic. How does the absence of the first harmonic affect the place and temporal coding of its pitch? Given that subjects report that the removal of the first harmonic does not affect the pitch, what does this suggest about the relative importance of the two mechanisms?

3. Process an [i] vowel of about 60dB in level at fundamental frequencies of 150 and 250Hz. Study the excitation pattern to see which aspects change and which stay the same as the fundamental frequency changes. You can print these out if that makes them easier to study.

a) How is the vowel quality represented in the excitation pattern (independent of fundamental frequency)?

b) Check your answer to a) by processing an [ɑ] vowel. Estimate the formant frequencies of the two vowels.

4. Process two sinewaves through the simulation. The first sinewave is fixed at 3925Hz. Move the second sinewave towards it, looking at the excitation pattern that arises from their interaction. In particular, study the excitation pattern when the second sinewave is at 3775Hz (150Hz different) and 4000Hz (75Hz different).

a) What do you think would be the pitch of these complexes and why?

5. (If you have time) Using the WASP program, record an [ɜ] vowel at a sampling rate of 44100 samples/sec. Save a section of the centre of the vowel to a WAV file, then load this file into the simulation. Estimate the fundamental frequency and the formant frequencies from the simulation, then check against the values given to you by WASP.
