Lecture 1-7: Source-Filter Model

Overview

1. **Properties of vowel sounds:** we can observe a number of properties of vowel sounds which tell us a great deal about how they must be generated: (i) they have pitch, so they are periodic signals, (ii) different vowels have different timbres, so they must have different harmonic amplitudes in their spectra, (iii) the same vowel can be spoken on different pitches, and different vowels can be spoken on the same pitch, so the pitch must be set independently from the vowel quality, (iv) the same vowel can be spoken on different voice qualities, so the voice quality must be set independently from the vowel quality, (v) vowel quality seems to depend mostly on tongue position: front-back and open-close, but (vi) vowel quality is also affected by the position of other articulators, the jaw, lips and velum.
2. **Source-filter model:** all of these characteristics of vowels can be explained by the source filter model of sound production in the vocal tract. This model of sound production assumes a **source** of sound and a **filter** that shapes that sound, organised so that the source and the filter are **independent**. This independence allows us to measure and quantify the source separately from the filter. For vowel sounds, the source of sound is the regular vibration of the vocal folds in the larynx and the filter is the whole vocal tract tube between the larynx and the lips. (Note: we can also apply the source-filter model to fricative sounds as we shall see later. For fricative sounds, the source of sound is the turbulence generated by passing air through a constriction, and the filter is the vocal tract tube anterior to the constriction.)
3. **Vowel Source:** vibration in the larynx is caused by blowing air between two tensed and approximated membranes: the vocal folds. The periodic buzz produced by the vibrating folds has a large number of harmonics up to 5000Hz or so, although the energy drops off with increasing frequency. Fundamental frequencies used in speaking are typically in the 100-200Hz range for men, and in the 150-300Hz range for women.
4. **Vowel Filter**: the frequency response of the vocal tract filter for vowels shows a small number of resonant peaks called **formants**. In a formant model of the vocal tract frequency response, each peak is considered to be a separate simple resonator; thus we tend to think of formants as individual resonances of the vocal tract (even though they are not really independent of one another) see figure 1-7.2. Studies of formant frequencies for different phonetic vowel qualities show a rough relation between the frequencies of the first two formants (F1, F2) and the position of the vowel on the vowel quadrilateral. This leads to the rule of thumb that F1 is associated with increasing open-ness of vowel articulation, while F2 is related to increasing front-ness of vowel articulation (see figure 1-7.1).

Reading

Choose at least one from:

- ☐ Rosen & Howell, Signals and Systems for Speech and Hearing (1st edition), The vocal tract as a linear system, pp92-99. *Accessible introduction to vocal tract response.*
- ☐ Ladefoged, Elements of Acoustic-Phonetics (2nd edition), Chapter 8: Resonances of the Vocal Tract. *Explanation of how resonances are related to articulation.*

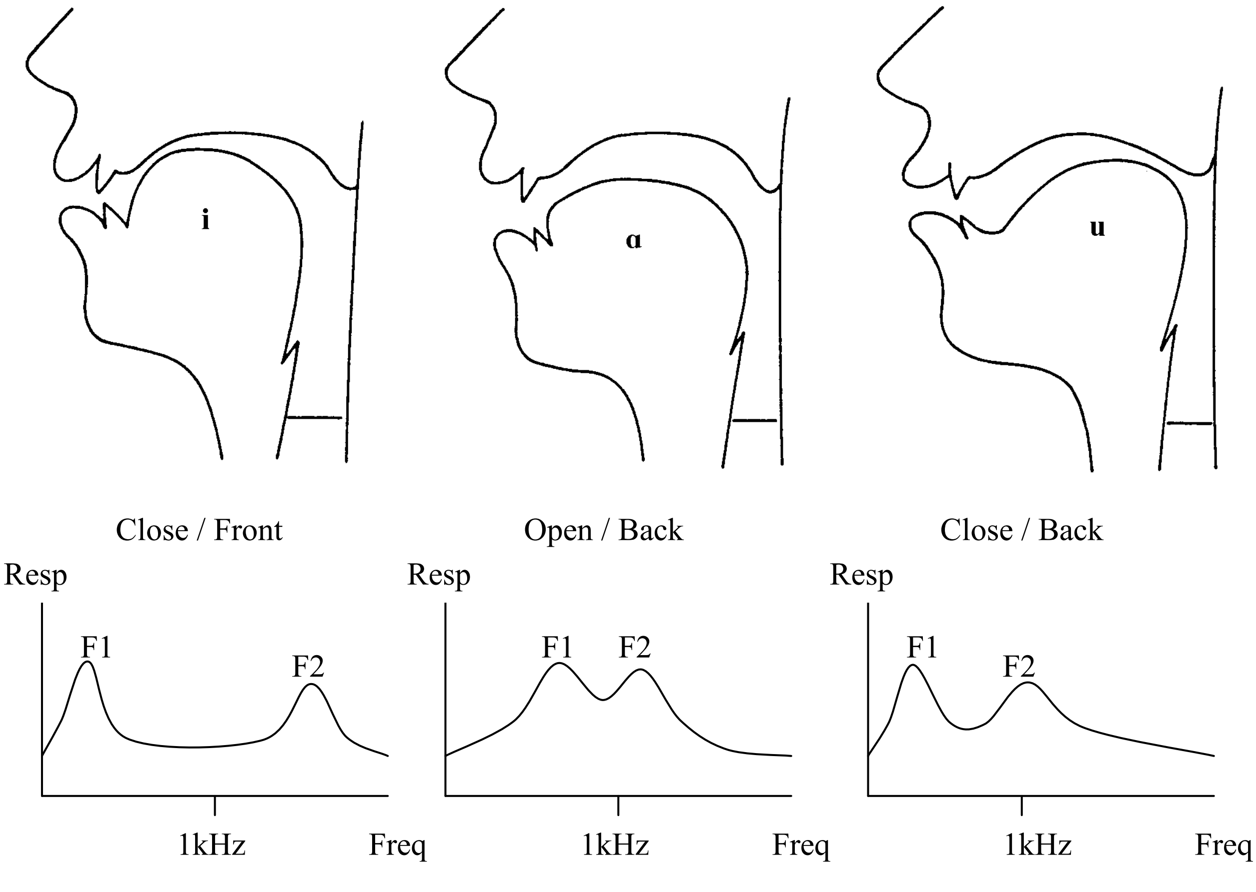
Learning Activities

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

1. Write an explanation of the source-filter model of speech production at a level that could be understood by a child at secondary school.
2. Research the history of the term 'formant'. Give arguments for and against the idea that formants are properties of vocal tract systems (as we teach) rather than properties of speech signals.
3. Draw a diagram of the frequency domain explanation of the generation of an /ɜː/ vowel using the source-filter model. Explain in words how your diagram would be different if the pitch of the vowel were raised.

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

Figure 1-7.1 Vocal tract shapes for different vowels lead to different frequency responses

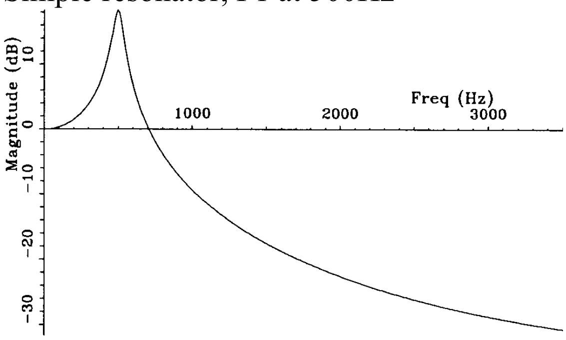


Definitions

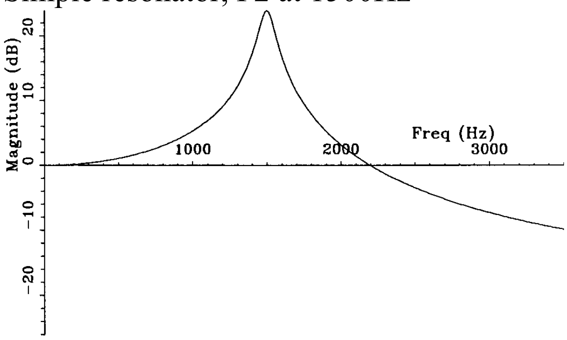
Source-Filter Model: a useful analytical model of how speech sounds are produced, which emphasises the independence of the source of sound in the vocal tract from the filter that shapes that sound.
Formant: a peak in the frequency response of an unobstructed vocal tract. One of the simple resonators that make up the complex resonant system of the vocal tract.

Figure 1-7.2 Vocal Tract Response is Like Combination of Simple Resonators

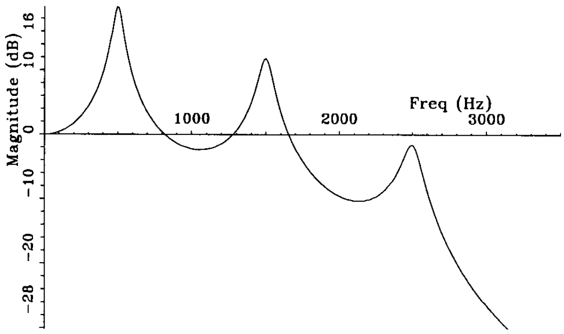
Simple resonator, F1 at 500Hz



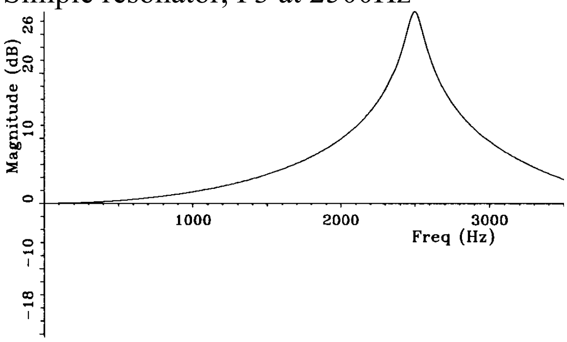
Simple resonator, F2 at 1500Hz



Complex resonator for vowel /3/:



Simple resonator, F3 at 2500Hz



The complex resonator can be measured ***as if*** *it were a chain of simple resonators.*

Reflections

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

1. What is meant by the source-filter model of speech production? Why is it important in the scientific study of speech?
2. Try and apply the source-filter model to other sound sources, such as musical instruments. When does the model break down?
3. What is the typical spectrum of vocal fold vibration? How does it change for changes in pitch?
4. What is a typical frequency response for a vowel?
5. Give a definition of a formant.
6. How does the frequency response for a vowel vary between close and open vowels? Between back and front vowels?
7. Use pictures to tell the story of the spectral development of a monophthongal vowel produced on a falling pitch.
8. What is the spectrum of a turbulent noise source in the vocal tract?
9. Why does [s] sound "brighter" than [ʃ] ?

Lab 1-7: Analysis and Generation of Vowels

Introduction

The source-filter model of vowel production states that the frequency content of a vowel may be explained by considering how the spectrum of the sound generated by the larynx is filtered by the vocal tract system. The independence of source and filter explains why vowels of the same timbre can be produced on different pitches, and why vowels of the same pitch can have different timbres.

The source filter model also helps to *quantify* vowels since we can separately measure the contributions of the source and the filter to the final vowel sound. The primary characteristic of the source is its fundamental frequency, while the primary characteristics of the filter can be reduced to the location in frequency of the vocal tract resonances or *formants*, see figure 1-7.2.

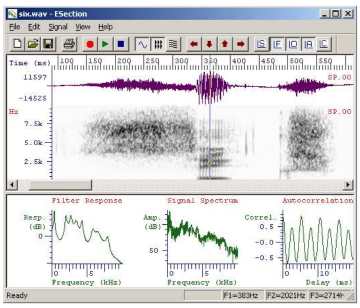
Scientific Objectives

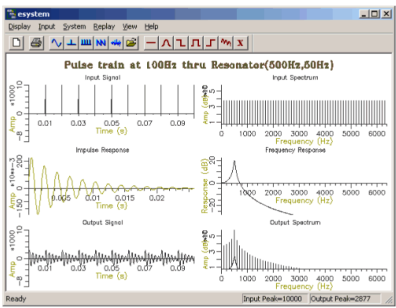
- To investigate the characteristics of vowel spectra that relate to pitch and timbre.
- To relate those characteristics to articulatory settings in the vocal tract.

Learning Objectives

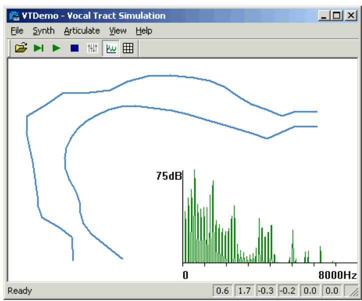
- To gain familiarity with the characteristics of vowel spectra and the terminology used to describe vowel spectra.
- To learn how to differentiate aspects of the spectrum concerned with the source from aspects concerned with the filter.
- To learn how to measure fundamental period, fundamental frequency and formant frequencies from waveform and spectrum displays of vowels.
- To understand the differences between the spectra of periodic and aperiodic sounds.

Apparatus


ESection



ESystem



VTDemo

The Lab PCs will be set up with three programs: ESection for analysing natural speech recordings, ESystem for signal generation, and VTDemo for vocal tract simulation. The analysis program allows you to study the waveform and the spectral cross section of a vowel so that you can determine its fundamental frequency and formant frequencies. The signal generation program allows you to filter a larynx-like signal through a vocal tract-like filter. The vocal tract synthesis program allows you to position the tongue jaw and lips of a simulated vocal tract to generate a replica of that sound. All these programs are also available on the web (http://www.phon.ucl.ac.uk/resource/sfs/)

Method

The analysis program displays the waveform at the top of screen, with two spectral displays at the bottom. The left spectral display is an estimate of the frequency response of the vocal tract filter, while the right spectral display is the signal spectrum. To select a region of the waveform and display a spectrum, first position left and right cursors on the waveform using the left and right mouse buttons. The frequencies of spectral peaks in the filter are shown in the status bar.

You can change the amount of waveform displayed by 'zooming' in to a region specified by the cursors. There are also commands for scrolling left and right along the waveform and to 'zoom out'. You can replay the signal and print any display. Be sure to enter your name as a title.

You met the ESystem program in the last laboratory session. You will be shown how to operate the vocal tract simulation.

Observations

1. Start the analysis program with the waveform 'ur'. This is a short section of a male [ɜː] vowel. Zoom in so that about 6 cycles are displayed across the screen.
   a. Use the timescale to determine the average period, and calculate the fundamental frequency.
   b. Position left and right cursors over six cycles to display its spectrum on the lower panel. Calculate the fundamental frequency by finding the frequency of the 10th harmonic and dividing by 10. It should be approximately the same as what you found in 1a.
   c. Position the cursors over a few cycles of the waveform and look at the estimated vocal tract filter frequency response to find values for the formant frequencies.
   d. Use the ESystem program to generate an artificial version of this sound using a sawtooth waveform (of the matching fundamental frequency) through a 3 resonance filter (of matching formant frequencies). Compare the spectra of the real and artificial versions. In what way do they sound different, and why?
   e. Switch to the vocal tract simulation and attempt to position the articulators so that a similar vowel is produced in terms of fundamental and formant frequencies. Are the position of the articulators what you expect?
2. Restart the analysis program with the waveform 'eh'. This is a pure vowel [e] spoken on a falling pitch by a male speaker.
   a. Sketch a (rough) graph of fundamental frequency against time for the vowel using one of the two methods you used in the last section.
   b. Estimate the formant frequencies.
   c. Switch to the simulation and build an animation of the vowel tract which produces a similar vowel sound, also falling in pitch. What aspects of the simulation are deficient?
3. Record and analyse some of your own vowel productions, for example [iː], [ɑː], [æ], or [uː] then attempt to build a simulation using the synthesizer. Note that you may want to change the “sex” of the synthesizer to get the best results.
   a. Record whether F1 and F2 is relatively high or relatively low for each case.
   b. Do the tongue positions for the simulation match your own? If not, why not?
   c. What would you predict for the movement of the formants in the production of an /ɔɪ/ diphthong? What do you observe? What is the cause of these differences?

Concluding Remarks

Formant frequencies are called F1, F2, F3, etc., for short. Some authorities call fundamental frequency 'F0', but we prefer the term 'Fx'. Why might that be?

Guidelines for Report

Your **introduction** should explain the basis of the source-filter model explanation of speech sound production. It should explain why vowel qualities can be characterised by the first few formant frequencies. Then describe how vowels can be synthesized artificially in two ways: an *acoustic* method using an excitation signal passing through a series of simple resonators, or an *articulatory* method using a model of larynx buzz passing through a model of the vocal tract pipe.

Your **results** section should include the measurements and diagrams of the sounds you analysed and synthesized in the lab including *at least one vowel of your own*. Be sure to explain what is shown in each diagram.

End your results section with a **discussion** of the strengths and weaknesses of the two approaches to vowel synthesis. Use the questions on the lab sheet as a guide for other things to discuss.