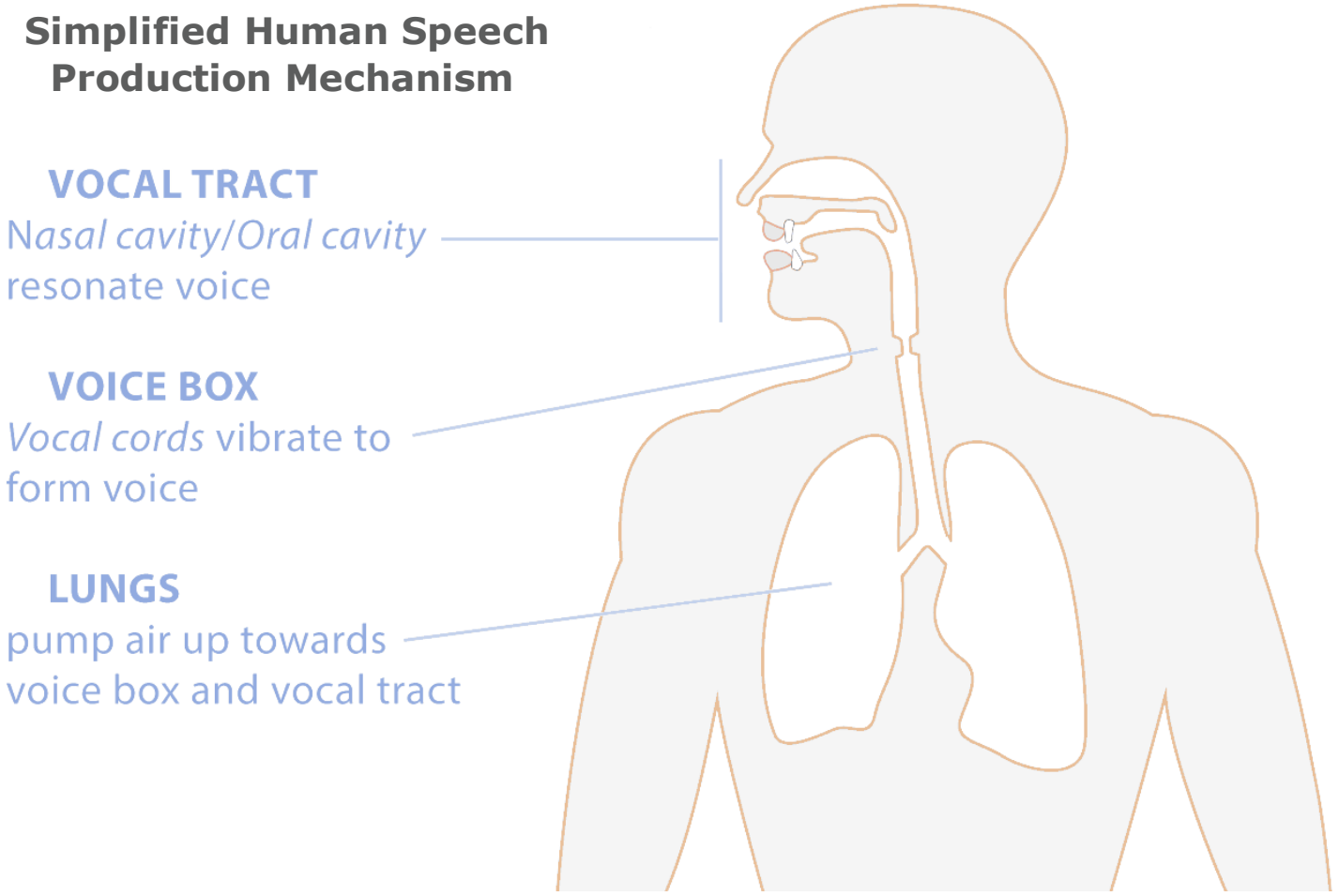
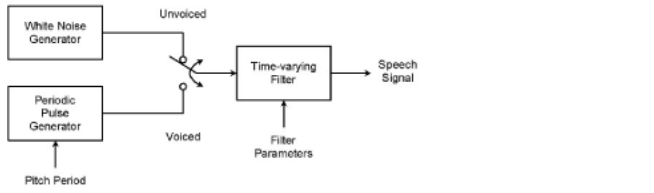
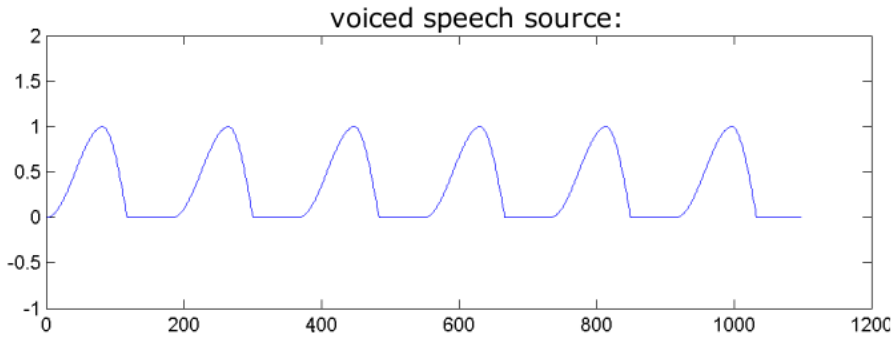
- - Speech Modelling and Infinite Impulse Response (IIR) Filters - -

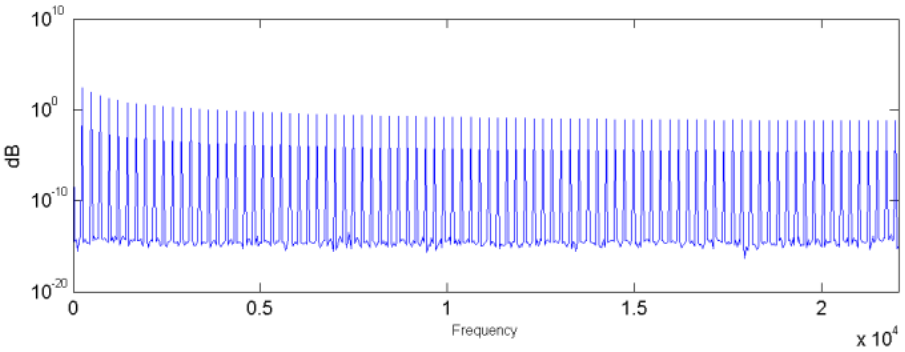


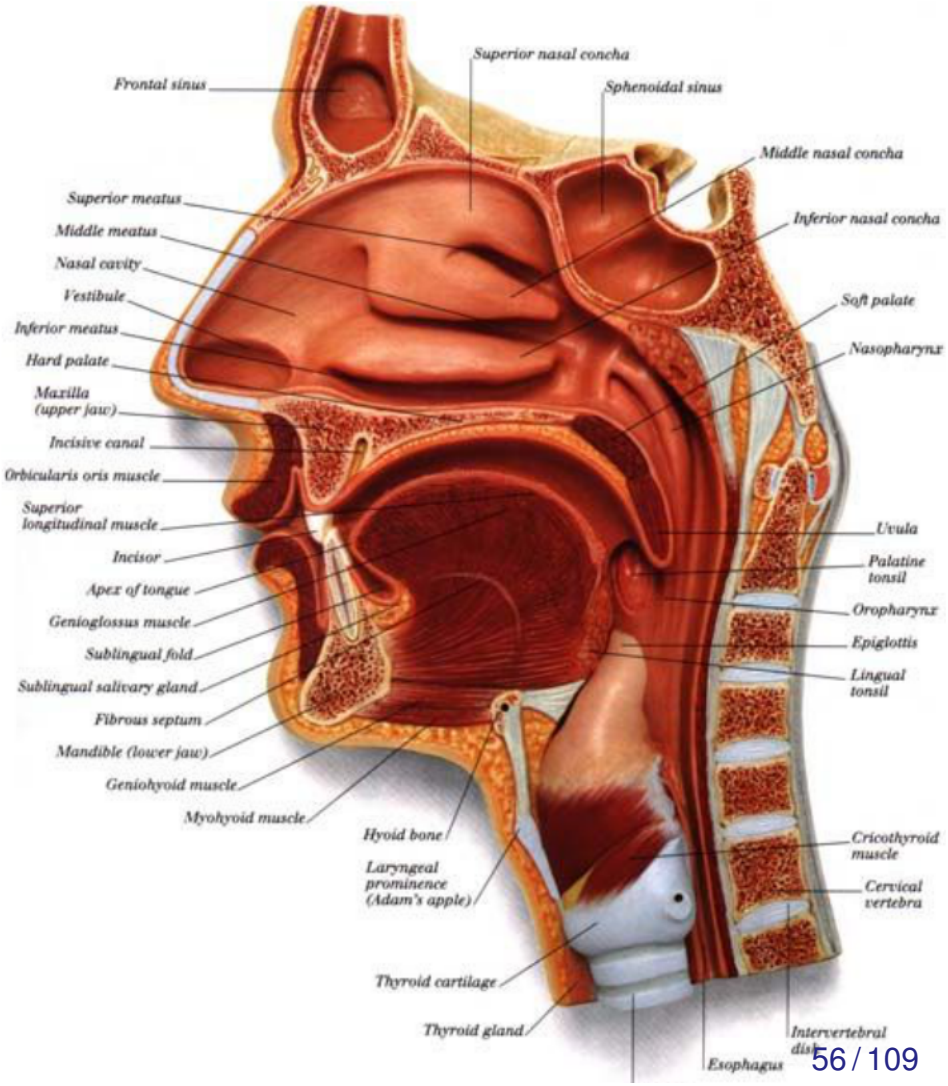
- main classification: speech is either voiced or unvoiced.

- - Speech Modelling and Infinite Impulse Response (IIR) Filters - -

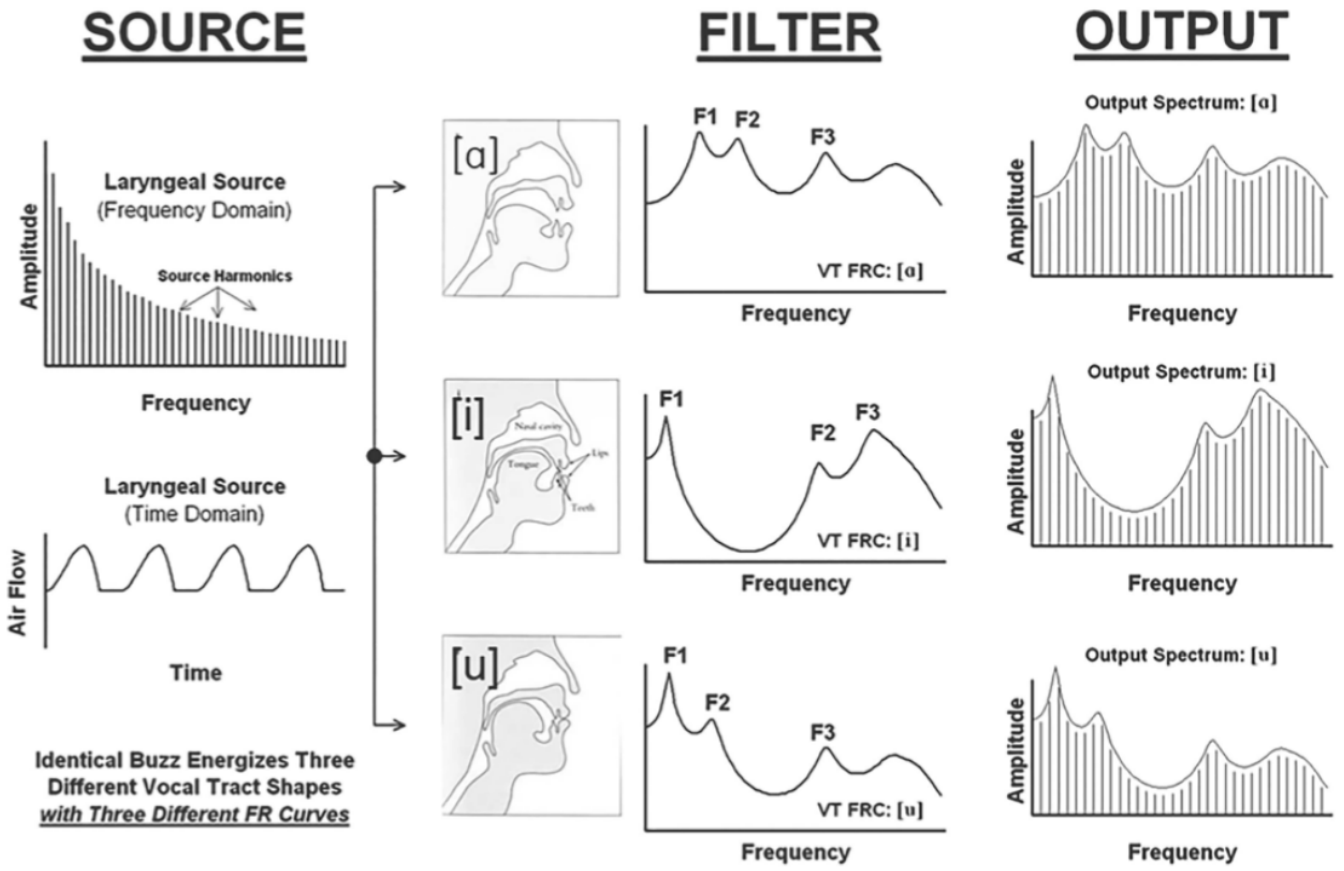








- - Speech Modelling and Infinite Impulse Response (IIR) Filters - -



Most important frequencies: pitch (F_0), i.e., the quasi-periodic frequency the vocal folds vibrate during voiced speech, and formants (F_1, F_2, F_3, ...), i.e., the vocal tract resonances.

- - Speech Modelling and Infinite Impulse Response (IIR) Filters - -

- Speech mechanism transfer function:
  - voiced speech, i.e., sounds derived from quasi-periodic vibrations of the vocal folds (basically vowels or consonants sounding as vowels):

$$H[z] = \frac{Y[z]}{X[z]} = \frac{1}{a_0 + a_1 z^{-1} + a_2 z^{-2} + ...}$$

  - unvoiced speech (all the other sounds):

$$H[z] = \frac{Y[z]}{X[z]} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + ...}{a_0 + a_1 z^{-1} + a_2 z^{-2} + ...}$$

- In the system transfer function, the roots of the numerator are known as **zeros** and the roots of the denominator as known as **poles**. In general, voiced speech corresponds to all-pole transfer functions, whereas unvoiced speech corresponds to functions with both poles and zeros. Vocal sounds amplify energy in specific frequencies (formants) and thus correspond to poles of $H[z]$, whereas nasal sounds drain energy from certain frequencies (anti-formants) and thus correspond to zeros of $H[z]$.

- - - Signal Processing Basics - - -

- We can characterize IIR filters and observe how they work just by converting the system transfer function $H[z] = \frac{Y[z]}{X[z]} = \frac{b_0+b_1z^{-1}+b_2z^{-2}+...}{a_0+a_1z^{-1}+a_2z^{-2}+...}$ from the Z-domain to the time-domain. In this case, $y[n]$ is a recursive filter, i.e., each of its samples is not only a function of $x[n]$ but also of past samples of itself. Contrary to this, FIR filters are not recursive.
- Example: write down the difference equation that corresponds to the system transfer function $H[z] = \frac{3+2z^{-1}-4z^{-2}}{2+0.4z^{-1}+1.3z^{-2}}$. Complementarily, depict the block diagram for the IIR filter.
- The poles and zeros of $H[z]$ can be represented graphically and their positions inside the z-plane reveal details on speech signals.
- Example: find the poles and the zeros of $H[z] = \frac{4-z^{-2}}{2-z^{-1}}$ and place them in the z-plane. Repeat it to $H[z] = \frac{1}{2-z^{-1}}$ and to $H[z] = 2 - z^{-1}$.
- **Today's Short Test (ST8)**: Find the difference equation matching the IIR filter $H[z] = \frac{1-2z^{-1}}{1-0.5z^{-1}}$. Then, plot its poles and zeros in the z-plane.