Theories of speech production

Pratik Bhandari, PhD

Nepal Center for Cognition and Society
`nams@nccs.pratikb.com`

2025-07-07

How do the vocal folds vibrate?

Figure: Image of a man sputtering Bronx cheer.

- Try the *Bronx cheer* or *raspberry*.
- VFs are similarly vibrated by the air exhaled out from the lungs.

Neurochronaxic theory of vocal fold vibration

- VFs vibrate in air like strings (mid 18th century).
- Frequency (f_o) of neural impulses generate VF vibration (Husson, 1950).
- This is obsolete.

Myoelastic-aerodynamic theory of phonation

- Early in the 19th century (Helmholtz and Müller)
- Advanced by van den Berg in the 1950s

Myoelastic

- Muscles control the elasticity and tension in the VFs
- Set VFs into vibration and control their (f_o)

Aerodynamic

- Air pressure from the lungs blows VFs apart.
- Air pressure drops below the VFs when the air escapes through the open glottis.
- VF then close again.

Source-filter theory of vowel production

Figure: Waveform at the glottis changes when it comes out from the lips.

- **Source**: vibration of the VFs
  - Sounds like a buzz, not like speech as we hear it.
- **Filter**: vocal tract between the larynx and the lips.
  - Different configurations due to articulator movement/placement.

Source-filter theory of vowel production

Figure: Vocal tract acts as a filter to change the waveform from the glottis.

- Assumes **independence** between the source of sound and the filter.
- Measure and quantify them separately.

Fundamental frequency (f_o) of VF vibration

- The number of times the vocal folds vibrate per second — they are blown apart and come together.
- Average f_o:
  - Men: $\sim$125 Hz
  - Women: $\sim$100 Hz
  - Children: $\sim$300 Hz or higher (highly variable as they are growing).
- In running speech, average f_o is rarely achieved.

Source function

- VF vibration produces not just one frequency but complex waveform — fundamental frequency and harmonics
  - The filter emphasizes one or the other frequency, which we will see later.
- Drop in amplitude: -12 dB/octave

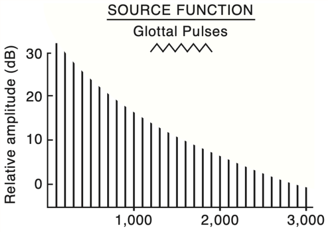

Figure: The spectrum of the glottal source.

Vocal tract: Resonance of a tube open at one end

- Length of an average male vocal tract = 17 cm
- Lowest resonant frequency of the tract (R_1)
  $= \frac{c}{\lambda} = \frac{344\text{ms}^{-1}}{(4*\text{length})} = \frac{34400\text{cm per sec}}{68\text{cm}} \approx 500Hz$

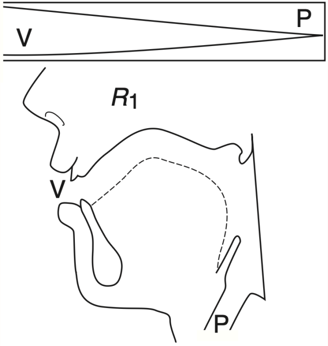

Figure: Vocal tract as a tube open one end; R, resonance; P, point of max pressure; V, point of max velocity.

Resonant frequencies of the vocal tract

- Even multiples: R_2, R_4, R_6, etc. are not effective.
  - They cancel out at the tube opening.
- Odd multiples: R_1, R_3, R_5, etc. are additive.
  - 500 Hz, 1500 Hz, 2500 Hz reinforce each-other.

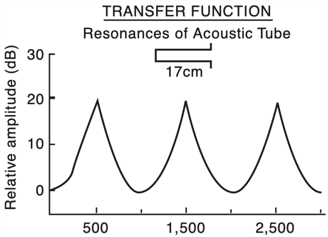

Figure: The frequency response of the vocal tract filter.

Output

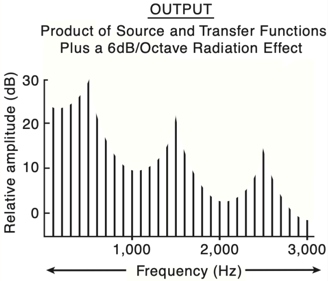

Figure: The spectrum of the speech after source passes through the filter.

Consult Raphael et al. (2011) for further elaboration.

Reference

Raphael, L. J., Borden, G. J., & Harris, K. S. (2011). *Speech science primer: Physiology, acoustics, and perception of speech* (6th ed.). Lippincott Williams & Wilkins.