The Speaking Mechanisms

FIGURE 4.1 The human vocal organs.

DP 4.1

Respiration – Part One
Managing Airflow

Talkers have an ongoing need to manage their airflow.

- **Primary Goal**: To produce a steady outflow of air to power speech.
- **Prerequisite**: To create an inflow of air that can be stored up and then expelled for speaking purposes.

**Airflow Facts:**

- #1: Air flows from a region of higher pressure to a region of lower pressure.
- #2: The greater the pressure difference between the regions, the greater the rate of airflow.
Talkers Inhaling & Exhaling

- Talkers vary lung volume to manage lung pressure relative to atmospheric pressure.

- To inhale lung volume must be expanded to create negative pressure.

- To exhale lung volume must be reduced steadily over time, to create and maintain positive pressure for long enough to express a thought.

- This is the business of the respiratory mechanism.
The Respiratory Mechanism:  
A System in Balance

- The lungs would be smaller if they could be.
- The thoracic skeleton would be larger if it could be.
- The lungs and the skeleton are coupled together, via the pleural linkage.
- They therefore come together at a relaxation point where all forces are in balance.
- The respiratory system will “spring back” to this balance point on its own, if it is given the chance.
Quiet Breathing

- Devote approx 40% of the cycle time to Inhalation & 60% to Exhalation.

- Complete a cycle in 2-3 seconds.

- Begin at REL (Resting Expiratory Level).

- Use inspiratory muscle forces to expand the lungs (and inhale).

- Relax and expire.

- Use only about 10% of the vital capacity (VC) of the lungs.
Speech Breathing

- Devote approx 10% of the cycle time to Inhalation & 90% to Exhalation.

- Extend cycle time for as long as needed to meet linguistic goals.
  - Can be 20 sec or more.

- Exchange a greater volume of air than for quiet breathing.

- How?
  - #1: Exert greater inspiratory muscle force than for quiet breathing.
  - #2: “Manage” the elastic forces of expiration in order to sustain and extend the expiratory phase.
Extending the Expiratory Phase

#1. Slow the rate of expiration down at the beginning, by opposing expiratory elastic forces with inspiratory muscle forces.

- Use what are nominally muscles of inspiration during the period of expiration!

#2. Later, extend expiration past the REL point, by engaging expiratory muscles.

NOTE: If the entire utterance is to sound equally loud throughout, and natural, all of these muscle activities must combine to produce a steady rate of expiration.
Spirometry

- A *Spirometric Tracing* is a chart showing lung volume as a function of time.
Spirometry

Quiet Breathing vs. Speech Breathing

Tidal breathing

Sustained tone

Speech
One method of making the speech air stream audible is to “chop it up” into a periodic series of bursts that are heard as a kind of buzz sound.

- **This is referred to as **Phonation**.
- **This is the principal activity of the larynx.**
Glottal Activity

Modal Register:

1.

\[
\begin{align*}
P & = \text{period of cycle} \\
HC & = \text{horizontal cursor} \\
VC & = \text{vertical cursor} \\
K & = \text{knee} \\
OP & = \text{open phase} \\
CP & = \text{closed phase}
\end{align*}
\]

Closed quotient \( (\text{CQ}) \) = \( \frac{CP}{P} \)

Closed to open ratio = \( \frac{CP}{OP} \)

Contact index \( (\text{CI}) \) = \( \frac{CP-OP}{CP} \)

F 6.16
Glottal Spectrum: Modal Register

- Fundamental frequency ($f_0$).
  - Men = ~ 100 Hz (shown below).
  - Women = ~200 Hz.
  - Children = ~300 Hz.

- All harmonics ($H_n$).
  - Intensity declines as harmonics get higher and higher.
  - Have sufficient power to be audible out to ~ 5000 Hz.
Voice Pitch “Contours”

- Drop the pitch at the end of a statement.
  - “You are over there.”

- Raise the pitch at the end of a question.
  - “Are you over there?”
Voiced, Voiceless, or Both

- **Voiced**: Phonation is required for the production of all vowels and many consonants.
  - “A very funny thing.”

- **Voiceless**: Voiceless consonants are produced in the mouth, on an air stream that passes freely through the glottis.
  - “A very funny thing.”

- **Both**: Voiced fricative consonants have both kinds of activation.
  - “A very funny thing.”
Theories of Phonation

The **Myoelastic Aerodynamic Theory**.

- The repeating cycles of phonation depend upon aerodynamic forces and upon the natural elasticity and tension of the system.

- Muscle forces are used to manage phonation, but not to move the vocal folds back and forth on each cycle.
Valves of the Vocal Tract

“The vocal tract can be conceptualized as a hollow tube that contains a series of valves… The valves are made up of the articulators, which come together in varying degrees and separate in various manners, thus constricting or obstructing the airstream in specific ways.” (Ferrand, p. 185)
“Newts” (/n/+/u/+t/+s/)

- In this lecture, we will consider some of the articulatory requirements associated with production of the word “newts” as in “Newts of several kinds are pictured below.”

[Images of newts]
The consonant /n/ is voiced, nasal, and alveolar. The combined realization of all of these features requires some care and precision on the part of the talker.
The Consonant /n/

- The soft palate must be lowered.
  - Relax *levator veli palatini*.
  - Possibly tense *palatoglossus*.

![Diagram of the human head showing muscles involved in articulation.](image-url)
The Consonant /n/ (con’t)

- The tongue must be pulled forward in the mouth.
  - **Genioglossus.**
The Consonant /n/ (con’t)

- The tongue tip must be elevated.
  - *Superior longitudinal.*

Fibers of the transverse muscle fan out from the region of the longitudinal sulcus.

F 8.09
• **Source-Filter Theory Revisited.**

  - All vowels begin with the same periodic sound source (The glottal waveform produced by phonation).
  - The vocal tract then filters this source due to resonance: Parts of the source spectrum are emphasized, other parts are de-emphasized.
Other Vowels

➢ To produce other vowels the tongue must be differently placed and differently shaped. The lips are also sometimes rounded.

![Diagram of vocal tract articulation for different vowels](image_url)

**FIGURE 4.9** Outlines of the vocal tract during the articulation of various vowels.

**DP 4.9**
The consonant /t/

- When producing the /t/ of “newts”, the tongue returns to a position similar to that for /n/, but in other respects /t/ is different.
  - /t/ is **voiceless**, so phonation must cease for a time during this gesture.
  - /t/ is **non-nasal**, so the valve to the nasal cavity must be closed off.
**The Consonant /s/**

- **Fricatives**: All fricative consonants begin with the same aperiodic sound source: The noise of frication, caused by turbulating air through a small opening.
  
  - *For /s/, this depends importantly on the integrity of the teeth.*

- The vocal tract then filters this source differently, depending on the place of articulation within the vocal tract.

F 8.15
Talkers typically have a lot to say and to say it they have to produce a lot of segments in a timely way.

Example: Think of any everyday conversation, say one about making plans for the afternoon.

Each Talker routinely produces 15 or more segments per second when having the conversation.
Coarticulation

- **Coarticulation** is an overlap in the production of neighboring speech segments. It speeds up the overall rate of segment production and helps keep the talker “in control.”

- Often, some of the articulators are needed to produce one segment, and some others are needed to produce the next segment.

  - *In such a case, the talker can begin work on the following segment even before the first one is through.*

  - /s/ as in “sue”, /s/ as in “see.”

  - /su/, /stu/, /stru/.
More on Coarticulation

- Coarticulation *blends things together acoustically.*

- **Result:** A single sound element, like a formant transition, may be related to two or more segments that overlapped when it was produced.
Suprasegmentals

- Each individual segment that is produced by a talker often has to fit into one or more larger patterns.

- These patterns are referred to as Suprasegmentals.

- **Example**: The overall duration of all the segments that make up a syllable may be constrained.

- “Taste” vs “Tasty” vs “Tastiest.”

- “Taste that please.” vs “How did that taste?”
Pitch Contours

F 8.32
The Vocal Tract

- Three cavities:
  - Nasal
  - Pharyngeal
  - Oral
The Nasal Cavity

- Sounds that pass through the nasal cavity will take on a very characteristic “nasally” quality.
  - Too much of this can be a problem.
  - Too little of this can be a problem.
  - Wrong timing can be a problem.

- The amount of nasalization of speech sounds is determined by the extent of opening at the velum.

F8.6
The Oral and Pharyngeal Cavities

- The oral and pharyngeal cavities together form a tube that extends from the glottis to the lips.

- Talkers control the overall length and cross sectional area of this tube by means of the articulators.

F 8.15
Oral and Pharyngeal Cavities (con’t)

- When *vowels* are produced the tube may be narrowed somewhat but it is always left open all along its length.

- When *consonants* are produced the tube gets greatly narrowed at some point along the way.
The Source-Filter Theory

- The source-filter theory says that a talker must do two things when producing speech sounds:

  - #1 **Energize the system by generating a complex sound (periodic, noisy, or transient). This is the SOURCE.**

  - #2 **Modify the spectrum of the complex sound by passing it through the various cavities of the vocal tract. This is the FILTER.**

- This lecture focuses on the source-filter theory as it pertains to vowels (and other voiced sounds).
Source-Filter Theory
Filtering in the Vocal Tract

- All vowel sounds start out the same at the glottal source, but then end up differently at the lips because they are filtered in different ways by the vocal tract.

- **Example:** The filters for /i/ and /u/.

F8.21
FORMANTS are resonances of the vocal tract that produce peaks in the output spectrum for a vowel.

- Formants can therefore be identified in the vowel spectrum.

- The first formant (labeled F1) is the lowest frequency peak, the second formant (F2) is the next higher peak, and so on…
Spectrograms

Spectrograms are acoustically displays that do two things important for speech:

• #1 They highlight the formants, making it easy to see what their frequencies are.

• #2 They “track” the formants over time, make it easy to see when the formant frequencies change.

F8.23
The Neutral Vowel Schwa

- The most neutral of all vowels is schwa.
  
  - *Produced with a relaxed vocal tract (and lips).*

- Results in a sort of tube-like vocal tract shape.
  
  - *The tube is closed on one end (at the glottis) and it has a constant diameter.*

- For an adult male talker, the spectrum has formant peaks at approx 500 Hz, 1500 Hz, and 2500 Hz.

- For an adult female talker the peaks will be 10 – 15% higher in frequency (due to smaller vocal tract size).
Monophthongal Vowels
(including schwa)

- Produced with a relatively fixed vocal tract shape.
- The associated formants (resonances) “shape” the vowel spectrum.
- The first two formants (F1 and F2) are the most important for vowel perception.
- The frequencies of F1 and F2 are closely related to: (1) tongue height and (2) tongue advancement.
The Vowel Quadrilateral

- Tongue height is **INVERSELY RELATED** to F1 frequency.
  - *The higher the vowel the lower the F1 frequency.*

- Tongue advancement is **DIRECTLY RELATED** to F2 frequency.
  - *The more advanced (fronted) the vowel, the higher the F2 frequency.*

---

![Vowel Quadrilateral Diagram]

F2

F1

---

Vowel Acoustics – Part One
The Vowel Quadrilateral

- Tongue height is **INVERSELY RELATED** to F1 frequency.
  - The higher the vowel the lower the F1 frequency.

- Tongue advancement is **DIRECTLY RELATED** to F2 frequency.
  - The more advanced (fronted) the vowel, the higher the F2 frequency.
Separate plots for men, women, and children reveal:

- Substantial average differences among talkers regarding both F1 and F2.
- Strong similarities across talkers regarding the relative positioning of vowels in the F1/F2 space.
- The relationships to vowel height and advancement noted above.
Spectrograms of monophthongal vowels highlight the formant frequencies.

F8.23 (a)
Spectrograms of Diphthongal Vowels

- Diphthongal vowels are dynamic (time varying) and those dynamics are clearly visible in a spectrogram.
More Diphthongs

- /i/, /a/, /u/
- /ai/, /au/, /iu/

Some Monophthongs & Diphthongs Recorded in the Lab

- Questions to be asked about diphthongs:
  - *When the formants are “stable” where are they?*
  - *When the formants are transitioning where do they go and how long do they take to get there?*
Consonants:

This, That, & the Other

- How consonants differ:
  - Manner *(how they are produced)*
  - Place *(where they are produced)*
  - Voicing *(whether phonated or not)*

- Previous lecture on consonants focused mostly on manner.

- This lecture will look mostly at place and voicing and some assorted other things.

- **General Question**: What can we find in the sound that seems to be related to the phonetic features?
Place of Articulation for Fricatives

- All fricatives start out with the same turbulent (“noisy”) source.
  - An approximately “white” noise.
  - Components extending up as high as 8,000 Hz.

- When the noise is produced at different places in the vocal tract, it gets filtered differently.
  - RESULT: Different portions of the noise spectrum are emphasized at different places of articulation.
Spectrograms of Fricatives

F 8.29 (b)
Place of Articulation for Stops: The Release Burst

- All stop consonants start out with the same transient source (a stop release).
  - Creates a brief “burst” of energy with components extending up as high as 8,000 Hz.

- When the burst is produced at different places in the vocal tract, it gets filtered differently.
  - RESULT: Different portions of the burst spectrum are emphasized at different places of articulation.
Spectrograms of Stops

F 8.27 (a)
Place of Articulation for Stops and Nasals:  
*The Slope of the F2 Transition*

- All stop consonants and nasal consonants involve some phonation.
- All stop consonants and nasals are produced in combination with a vowel and the vowel is also phonated.

**THEREFORE:** There will be formant transitions.

- The *slope of the F2 transition* is different for different places of articulation.
Spectrograms of Syllables that Start with a Stop Consonant

F 8.28a
Voiced vs. Voiceless for Stop Consonants

- For a voiced stop consonant, phonation begins close to the moment of stop release.

- For a voiceless stop consonant, phonation is delayed until some time well after the moment of stop release.

- The variable that describes this relationship is voice onset time (VOT).

- NOTE: VOT can be negative, zero, or positive.
/k/ vs. /g/

F 8.28b
Before and After the Vowel

- Note above, that the acoustics of a consonant are often very different depending on whether it occurs:
  - *Prevocically* (before the vowel).
  - *Postvoically* (after the vowel).

- A consonant may also occur *intervocically* (in between two vowels).
  - Example: “Achoo”
F1 Rising (Prevocalic Consonants)

- Recall that there is an inverse relationship between F1 and vowel height.
- All consonants are even “higher” than the highest vowel, in the sense that the vocal tract is even more constricted.
- F1 therefore starts lower for a consonant than it does for any vowel.
- **RESULT:** The will be a rising F1 transition as you move out of a consonant and into a vowel.

  *See examples above.*