Articulatory Phonetics

Lecturer: Dr Anna Sfakianaki

HY578 Digital Speech Signal Processing
Spring Term 2016-17
CSD University of Crete
What is Phonetics?

Phonetics is a branch of Linguistics that systematically studies the sounds of human speech.

1. How speech sounds are produced
2. How speech sounds are transmitted
3. How speech sounds are received

It is an interdisciplinary subject, theoretical as much as experimental.
Why do speech engineers need phonetics?

- An engineer working on speech signal processing usually ignores the linguistic background of the speech he/she analyzes. (Olaszy, 2005)
  - How was the utterance planned in the speaker’s brain?
  - How was it produced by the speaker’s articulation organs?
  - What sort of contextual influences did it receive?
  - How will the listener decode the message?
Phonetics in Speech Engineering

Combined knowledge of articulatory gestures and acoustic properties of speech sounds

Categorization of sounds  speech Segmentation

Speech Database Annotation

Algorithms

Speech Recognition  Speech Synthesis
Phonetics in Speech Engineering

Speech Disorders
- diagnosis
- treatment

Pronunciation Teaching Tools
- L2
- Foreign languages

Speech Intelligibility Enhancement
- Hearing aids
- Other tools
Tuesday
Articulatory Phonetics
- Speech production
- Sound waves
- Places and manners of articulation
  - Consonants & Vowels
- Waveforms of consonants - VOT
- Suprasegmentals

Thursday
Acoustic Phonetics
- Formants
- Fundamental Frequency
- Acoustics of Vowels
  - Articulatory vs Acoustic charts
- Acoustics of Consonants
  - Formant Transitions

Friday
More Acoustic Phonetics...
- Interpreting spectrograms
- The guessing game...
- Individual Differences
Peter Ladefoged

- Professor UCLA (1962-1991)
- Travelled in Europe, Africa, India, China, Australia, etc.
- Interested in listening to and describing every sound used in spoken human language, which he estimated at 900 consonants and 200 vowels (*The Sounds of the World's Languages*).
- He was president of the International Phonetic Association (1986-1991) & the Linguistic Society of America.
- Had a brief career in Hollywood as the chief linguistic consultant on the 1964 film *My Fair Lady*.
- Exemplary teacher

Home Page: [http://www.linguistics.ucla.edu/people/ladefoge/](http://www.linguistics.ucla.edu/people/ladefoge/)
Most speech sounds result from movements of the tongue and the lips.

Speech movements are named **articulatory gestures**.

Making speech gestures audible involves
- pushing air out of the lungs
- producing a noise in the throat or mouth

Tongue and lip movements form the noise coming from the larynx.
Joy Nash trapped in an MRI machine…

https://www.youtube.com/watch?v=0-aEN2xHBCc
Speech Production

- The tongue and lips move rapidly from one position to another.
- The actions of the tongue are among the fastest and the most precise physical movements that people make.
The basic source of power for speech: the respiratory system

- Pushing air out of the lungs
- Lungs $\rightarrow$ trachea $\rightarrow$ larynx $\rightarrow$ vocal folds

Try to talk while breathing in instead of out. What do you observe?
Speech Production - Vocal folds

- In the larynx there are two small muscular folds, the vocal folds.
- If they are apart, the air has free passage into the pharynx and the mouth.
Abducted vocal folds:
- respiration
- Production of *voiceless* sounds

Adducted vocal folds:
- Production of *voiced* sounds (phonation)

Exercise: Voiceless vs voiced sound
[fffffffvvvvvvvvfffffffv v v v v v v v v]
- Put your fingertips against the larynx.
- Stop up your ears while contrasting.
Speech Production - Vocal folds

- Stroboscopy: female vocal folds vibrating at high and low pitches
  Video: [http://www.youtube.com/watch?v=UpOXecWC5Dw](http://www.youtube.com/watch?v=UpOXecWC5Dw)
Voicing

- Distinguishing sounds on the basis of voicing:
  - fat vs. vat
  - thigh vs. thy
  - sue vs. zoo
  - φάρος vs. βάρος
  - σώνει vs. ζώνη

More pairs?
Vocal Tract

- Oral Tract
  - mouth
  - pharynx

- Nasal Tract

- Articulators
  - tongue
  - lips

Flap at the back of mouth:
- air goes in and out through the nose
- production of [m] and [n]
The four main components of the speech production mechanism:

1. airstream process
2. phonation process
3. oro-nasal process
4. articulatory process
The way in which we hear a sound depends on its **acoustic structure**.

Why do we want to be able to describe the acoustics of speech?

- Understanding how do certain sounds become confused with one another
- Better description of vowels in terms of acoustics than articulatory gestures
- Understanding how computers synthesize and recognize speech
- Audio recording provides permanent data we can analyze and study
Speech sounds differ from one another in three ways:

1. **pitch/frequency**
2. **loudness**
3. **quality**

How is sound produced:

- Articulatory movements superimposed on outgoing flow of lung air → small variations in air pressure → **sound wave** → vibrations in listener’s eardrum.
Sound waves (Video)

http://www.youtube.com/watch?v=-rFnzHXX1vk
Sound waves (Video)

- Sound, vibration and acoustic characteristics

http://www.youtube.com/watch?v=dbeK1fg1Rew&feature=related
Word duration: 0.6s
Recurrence of major peaks in air pressure: 0.01s →
Vocal folds vibrate 100 times a sec. →
1 pulse every 1/100 sec

Variations within each period → Vocal tract shape
(vowel quality)
CONSONANTS
- smaller amplitude
- irregular vibrations in air pressure
- Vocal folds do not vibrate (voiceless C).

VOWELS
- large regular pulses of air pressure
- Vocal folds vibrate.
Places of articulatory gestures

- Articulators: parts of the vocal tract used to form sounds
- Articulators forming the lower surface of the vocal tract
  - are highly mobile
  - move towards articulators that form the upper surface

Exercise: Try saying the word “capital” and note the major movements of your tongue and lips.
Soft palate/velum: muscular flap that can be raised to press against the back wall of the pharynx and shut off the nasal tract, preventing air from going out through the nose (velic closure).
For English sounds:
Visit the website
http://www.uiowa.edu/~acadttech/phonetics/about.html
and select “Articulatory Anatomy”

For Greek sounds:
Visit the website
http://speakgreek.web.auth.gr/ and then select
“Εργαλείο-Σύνδεση” and
“Phonetic Library” (choice
of Greek or English language)
Άνω αρθρωτές (Articulators on upper surface of vocal tract)

Corresponding Greek & English terms

- χείλος (lip)
- οδόντες (teeth)
- φατνία (alveolar ridge)
- ουρανίσκος (hard palate)
- υπερώα (soft palate/velum)
- σταφυλή (uvula)
Кάτω αρθρωτές (Articulators on lower surface of vocal tract)

Corresponding Greek & English terms

- κάτω χείλος — bottom lip
- κάτω οδόντες — bottom teeth
- άκρο — tip
- προράχη — blade
- πρόσθιο τμήμα — front
- κέντρο — center
- ράχη — back/dorsum
- ρίζα — root
- επιγλωττίδα — epiglottis
Examples

- “peculiar”
  1. lips come together
  2. back and center of the tongue are raised (towards hard palate or velum?)
  3. tip of the tongue on alveolar ridge

- “true” vs. “tea”

- “sigh” vs. “shy”
Tongue depiction

- Mid-sagittal vs. 3D view

Takano & Honda (2007)
“It ran a lot”

Young & Stone (2002)
Basic places of consonant articulation

- In order to form consonants, the airstream through the vocal tract must be **obstructed** in some way.
- Consonants can be **classified** according to the place and manner of this **obstruction**.
<table>
<thead>
<tr>
<th>Articulator</th>
<th>Articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>lips</td>
<td>labial</td>
</tr>
<tr>
<td>tongue tip and blade</td>
<td>coronal</td>
</tr>
<tr>
<td>back of the tongue</td>
<td>dorsal</td>
</tr>
</tbody>
</table>

Example: “topic”
Places of consonant articulation

LABIAL ARTICULATION

- **bilabial**
  The two lips come together.
  - *pie, buy, my*

- **labiodental**
  The lower lip is raised and nearly touches the upper front teeth.
  - *fie, vie*

CORONAL ARTICULATION

- **dental**
  Tongue tip/blade protruding between upper and lower teeth *(interdental)* or close behind the upper front teeth
  - *thigh, thy*
Places of consonant articulation

CORONAL ARTICULATION (cont’d)

- **alveolar**
  tip/blade of the tongue at the alveolar ridge
  *
  *tie, die, nigh
  *sigh, zeal
  *lie

- **retroflex**
  tongue tip at the back of alveolar ridge
  *
  *rye, row, ray / *ire, hour, air

- **palato-alveolar** or **post-alveolar**
  tongue blade at the back of alveolar ridge
  *
  *shy, she, show

Tip: Articulate and hold the position while taking breath in
Places of consonant articulation

CORONAL / DORSAL ARTICULATION

- **palatal**
  front of the tongue at hard palate
  - *you*

DORSAL ARTICULATION

- **velar**
  back of the tongue at soft palate
  - *hack, hag, hang*

Example:

- *fee* → *theme* → *see* → *she*
- *labiodental* → *(inter)dental* → *alveolar* → *palato-alveolar*
The oro-nasal process

- In most speech, the soft palate is raised so that there is a **velic closure (oral sounds)**.

- During production of **nasal sounds**:
  - There is an obstruction in the mouth.
  - The velum is lowered so that air escapes through the nasal cavity.

Example:  
- velar → alveolar → bilabial
Manners of articulation

- At most places of articulation, there are **several ways** in which articulatory gestures can be accomplished.
  - Oral tract may close off
    - for an instant
    - for a longer period
  - The articulators may
    - narrow the space considerably
    - simply approach each other
Manners of articulation: *stop*

- Complete closure of articulators involved so that the airstream cannot escape through the mouth.

- Types of stops:
  - oral stop
  - nasal stop
Oral stop

- articulatory closure in the mouth
- the nasal tract is blocked off (raised soft palate)
- pressure in the mouth builds up
- airstream is released → burst → plosives

Example: pie, buy → tie, dye → kye, guy
bilabial → alveolar → velar
Nasal stop

- Articulatory closure in the mouth
- Lowered soft palate $\rightarrow$ air goes through nasal cavity

Usually:
- stop = oral stop
- nasal = nasal stop

Example:  
- my $\rightarrow$ bilabial  
- nigh $\rightarrow$ alveolar  
- hang $\rightarrow$ velar
Oral vs. Nasal stop

Oral

Nasal
Fricative

- close approximation of two articulators
- airstream is partially obstructed
- turbulent airflow is produced
  (hissing sound - *noise*)

Example: fie, vie → thigh, thy → sigh, zoo → shy
labiodental → dental → alveolar → palato-alveolar
Approximant

- approximation of two articulators
- vocal tract not narrowed to such an extent that turbulent airstream is produced

Example: yacht → we → raw
palatal → labial-velar → alveolar
**Lateral (approximant)**

- Obstruction of airstream at a point along the center of the oral tract
- Incomplete closure between one or both sides of the tongue and the roof of the mouth
- Air flows freely over the side of the tongue

**Example:** lie, laugh, hill alveolar
Additional consonantal gestures

- tongue-tip **trill** (roll)
  *rye, raw* (Scottish English)

- **tap** (flap)
  *poζ* (Greek /r/) or *pitty* (American English)

- **affricate** (stop + fricative)
  *church, judge*

- **glottal stop** [?]
  *flee east vs. fleeced*
Summary

Consonants are described in terms of five factors

1. state of vocal folds (voiced/voiceless)
2. place of articulation
3. central or lateral articulation
4. soft palate raised or lowered (oral/nasal)
5. manner of articulation

Exercise

* sing *

1. voiceless
2. alveolar
3. central
4. oral
5. fricative

* sing *

1. voiced
2. velar
3. central
4. nasal
5. stop
<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>Alveolo-palatal</th>
<th>palatal</th>
<th>palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>nasal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(central) approximant</td>
<td>(ʼ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Waveforms of Consonants
Waveforms of Consonants

my two boys know how to finish
Waveform of /t/ vs. /d/

- /tʰ/
  - spike indicating noise burst
  - after burst very small semi-random variations during the aspiration

- /d/
  - no spike, smaller noise burst
  - very little gap between burst and vowel start
Voice Onset Time (VOT) is the duration of the period of time between the release of a plosive and the beginning of vocal fold vibration. This period is usually measured in milliseconds (ms).
- **Positive VOT**: where there is a delay in the onset of vocal fold vibration after the plosive release
- **Zero VOT**: where the onset of vocal fold vibration coincides (approximately) with the plosive release
- **Negative VOT**: where the onset of vocal fold vibration precedes the plosive release
The articulation of vowel sounds

- Articulators do not come very close together → the passage of the airstream is relatively unobstructed.

- We describe vowel sounds in terms of
  - the position of the highest point of the tongue
  - the position of the lips.
Tongue position
UCLA tongue video

- X ray video of tongue and lip movement during production of vowels /i, e, a, o, u/.

Video: http://www.phonetics.ucla.edu/vowels/chapter11/tongue.html
Targets for vowel gestures

1. heed
2. hid
3. head
4. had
5. father
6. good
7. food
Front vowels

- The **highest** point of the tongue is in the **front** of the mouth.
- The mouth becomes progressively more **open**.
- The tongue remains in the front.
  1. *heed*: high front
  2. *hid*: mid-high front
  3. *head*: mid-low front
  4. *had*: low front
Back vowels

- The tongue is close to the back surface of the vocal tract.

5. father: low back
6. good: mid high back
7. food: high back
Lip rounding

- In good and food there is movement of the lips called lip rounding.

Unrounded vowels
heed, hid, head, had, father

Rounded vowels
good, food
Articulatory description of vowels

1. **height** of tongue body
2. **front-back** position of the tongue
3. degree of lip rounding

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[e]</th>
<th>[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>[i]</th>
<th>[u]</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td></td>
<td></td>
</tr>
<tr>
<td>back</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Articulatory description of vowels

- Very difficult to become aware of the position of the tongue in vowels.
- Get some impression of **tongue height** by observing **position of jaw** while saying the vowels in “*heed, hid, head, had*”.
- Compare **he vs. who** → Feel your tongue going from front to back and feel your lips become more rounded.
Relative positions of highest points of the tongue

- Specification of vowels in these terms not so satisfactory.

- Vowels classified as “high” do not have the same height (see 1 vs. 7).
- “Back” vowels vary in their degree of backness (see 5, 6, 7).
- Shape of the tongue and pharynx width are not taken into account.
Suprasegmentals

- Vowels & Consonants = Segments
- Segments → Syllables → Utterances
- Suprasegmentals:
  - Features superimposed on the syllables
  - They can affect single segments as well as whole syllables.
Stress can have a grammatical function

- an ‹insult - to in›sult (noun - verb)
- a ‹walkout - to ›walk ›but (noun - verb)
- a ‹hot dog - a ›hot ›dog (compound noun – adjective+noun)
- ‹diplomat ‹di›plomacy ◀diplomatic
- ‹photograph ‹pho›ography ◀photographic
- ‹monotone ‹mo›hotony ◀monolithic

Contrastive Stress

- I want a red pen, not a black one.
Stress in English is produced by:

1. increased activity in the respiratory muscles, producing **greater loudness**
2. exaggeration of consonant and vowel **properties** (vowel height, stop aspiration)
3. exaggeration of **pitch**
Pitch

- Pitch of the voice is what you alter to sing different notes in a song.
- The pitch of a sound is an auditory property that enables a listener to put it on a scale going from low to high.
- When a speech sound goes up in frequency, it also goes up in pitch.
Intonation

- The **pitch pattern** in a sentence is known as intonation.

- This is my father.  statement

- Is this your father?  question
Intonation

- That’s a cat. statement
- That’s a cat? question

- It is the **relative values** of pitch, length, or degree of stress of an item that are significant.
- The absolute values are never linguistically important!

Visit the websites:
- [http://soundsofspeech.uiowa.edu/english/english.html](http://soundsofspeech.uiowa.edu/english/english.html) (Interactive Phonetic Library for American English)
- [http://homes.chass.utoronto.ca/~danhall/phonetics/sammy.html](http://homes.chass.utoronto.ca/~danhall/phonetics/sammy.html) (Interactive Sagittal Section)
- [http://www.phonetics.ucla.edu/course/chapter1/linkschapter1.htm](http://www.phonetics.ucla.edu/course/chapter1/linkschapter1.htm) (Material from UCLA Phonetics Lab Data, Ladefoged “A course in phonetics”, 5th ed.)