Lecture 2

Phonetics

Representing the Sounds of Speech

- *Pronunciation*: the way words are spoken in a language
- Differences in pronunciation are quite standard in many languages
  - e.g. some American English speakers merge `pen` and `pin` to sound the same
  - e.g. some American English speakers merge `caught` and `cot` to sound the same
  - Having many *dialects* (≈ variations) of a language often results in this
  - Pronunciation is part of knowledge when speaking a language (recall: communication chain)
- Many ways to study pronunciation in spoken languages:
  - *Articulatory phonetics*: the way in which speech sounds are produced
    - Key questions: Which parts of the mouth are used? What are the configurations?
  - *Acoustic phonetics*: characteristics of sounds produced by articulations
    - Key aspects: How do speech wave patterns look? Why do they take certain shapes for certain phenomena?
  - *Auditory phonetics*: how humans process sounds in speech
    - How do humans respond to particular auditory stimuli? What are their brain patterns?
- *Phoneticians* investigate speech sounds and representations
  - Investigating articulatory phonetics: *x-ray photography* used to be quite popular
    - However, x-ray radiation puts habitually-observed speakers at risk of radiation poisoning, etc.
    - Less radiation risk: *x-ray microbeam* and *electromagnetic articulograph*
    - These other methods utilize point-tracking devices to determine locations of receptors on lip/tongue/jaw
  - Investigating articulatory phonetics: using *palatography* to observe contact between tongue and mouth
    - Palatographical approach requires using specialized instruments to measure air flow/pressure during speech
  - Investigating articulatory phonetics: *ultrasound* imaging, similar to pregnancy ultrasound (expecting mothers)
    - Gives full image of tongue during articulation (trade-off: more intrusive method)
  - Investigating acoustic phonetics: *sound spectrography*
    - Pictures of sounds, help explore physical properties (waves, beats, etc.)
    - Typically requires use of a waveform editor, spectrograph, or other phonetics analysis software
  - Investigating auditory phonetics: computerized mechanisms such as scans
    - Advanced study requires materials like *magnetic resonance imaging* (MRI)
    - Some aspects of auditory analysis require use of *computerized tomography* (CT)
- Basic approach still used by phoneticians: *impressionistic phonetic transcription*
  - Key idea: writing down speech sounds to capture pronunciation
  - e.g. `to-mae-to` vs. `to-mah-to`

  - Spelling is conventionalized (relatively recent, regional differences present—e.g. USA vs. Britain vs. Canada)
  - Phonetic spelling is not conventionalized, however (but there are widely-used *phonetic systems*)
  - Consider these examples from English, Webster's *New Int'l. Dictionary*, and *American Heritage Dictionary*):

English	**Webster's**	**A.H.**
tomato	tə'mātō	təmā'tō
tomato	tə'måtō	təmä'tō

Figure 2.1: A demonstration of how phonetic spelling may not necessarily be standard.

- Working towards the best choice of a *phonetic alphabet*:
  - Keep in mind: there is no "right" choice (recall: descriptive linguistics)
  - Choices are influenced largely by considerations of typographical/historical context (recall: arbitrariness)
  - The reader and author MUST agree on sound qualities assigned to symbols in a phonetic alphabet
  - So we can create some rules of thumb for our phonetic transcription system:
    1. Each symbol should only represent one "sound" (*phone*), and only one symbol per sound (one-to-one)
    2. If two sounds can alone distinguish two words apart, then they must have different representations
    3. If two sounds are similar but the difference arises in context, then that should be represented somehow
  - For the third rule, we have to keep in mind the idea of *co-articulation*
    - That is, one sound can influence a neighboring sound, as words are made via a single flowing action of sound
    - So we may want a way to distinguish two very similar sounds if context matters that much
    - However, if we wish to NOT capture this variation, then we need the option of representing the similarity
  - English is NOT the ideal choice for a phonetic alphabet (breaks the rules a lot):
    1. Same sound may be spelled using different letters (e.g. vowels in `sea` vs. `see`, etc.)
    2. Same letters could correspond to different sounds (e.g. '`a`' in `all` vs. `apple`)
    3. Single sound could be written with more than one letter (e.g. '`ck`', as in `lock`)
    4. Single letter could represent more than one sound (e.g. '`x`' is frequently used as '`ks`')
    5. Some letters may not represent a sound at all (i.e. *silent* letters, e.g. '`h`' at end, '`kn`' = '`n`' at start, etc.)
  - We need a system that has this one-to-one correspondence between sounds and symbols
    - Eliminates inconsistencies and ambiguities, standardizes spelling approach
    - This approach can work even for other languages, since **sounds, unlike words, are fundamental**
    - This way, we can transcribe a large variety of text, literature, and speech
- A pretty good system in wide use today is *International Phonetic Alphabet* (IPA):
  - Has all the useful properties of a proper phonetic alphabet as described above
  - Can transcribe broadly as well as provide very fine phonetic information
  - We will build up our knowledge of IPA bit-by-bit, but its use will be ubiquitous henceforth
  - IPA is very much the standard in the linguistic literature (peer-reviewed journals, etc.)
- We need to classify the sound system in a more fine-grained manner:
  - Speech stream can be divided into two categories: *segments* and *suprasegmentals*
  - Segments: discrete units of speech stream
    - Further subdivisions: *consonants* and *vowels*
    - Can be easily transcribed using discrete symbols that somewhat match up with English
  - Suprasegmentals: "ride on top of segments", i.e. properties that apply to entire strings of consonants/vowels
    - e.g. stress, tone, pitch, intonation
    - Much more difficult to represent with an alphabetic transcription system
    - Require a more complex system to be utilized for faithful transcription
- We will focus on segments for now and revisit suprasegmentals later on

- Detailed articulatory view of segments:
  - Anatomical (but not physical) basis of sound production
  - Both consonants and vowels are produced by positioning the *vocal tract* in various configurations
  - Vocal tract configurations (VTCs) are an important aspect to study in articulatory phonetics
  - Consonants: produced with a constriction in the vocal tract that impedes airflow
  - Vowels: produced with at most a slight narrowing—allow free air flow in the oral cavity
- We can also do an acoustic analysis on the differences between consonants and vowels
  - i.e. study the waveforms of both kinds of segments and assign patterns (*data-based learning*)
  - We will revisit this type of analysis later on
- Consonants and vowels play different roles in *syllables*, though
  - Syllable: unit of speech (every utterance by a human MUST contain at least one syllable)
  - Syllables are not the fundamental/indivisible units, though (phonemes are!)
  - Words can be *monosyllabic* (e.g. `the`, `spring`) or *polysyllabic* (e.g. `polysyllabic`)
  - Interesting: `polysyllabic` is an example of an *autological* word, i.e. one whose meaning describes its etymology/self
  - Further subdivisions of a syllable: *onset* and *rhyme*
    - Rhyme: a vowel and consonants that come after it
    - Onset: consonants that occur before the rhyme within the syllable
    - All syllables must have a rhyme, but onsets may be optional (dependent on language)
    - Implication: every syllable must have a vowel or a functionally vowel-like component within it
    - Further subdivisions of a rhyme: *nucleus* and *coda*
      - ☐ Nucleus: vocalic part of the rhyme
      - ☐ Coda: the final consonants in the rhyme
      - ☐ Completes the tree-like structure of a syllable (nuclei + coda = leaves)
      - ☐ These are appropriately-named terms, as nucleus = "heart" of syllable, coda = end/finishing touches
    - Nuclei carry suprasegmental information (stress, volume, pitch, etc.)
    - Recall: vowels are better suited for suprasegmental function than consonants are
    - *Syllabic consonants*: rare cases where consonants function as the nucleus of the syllable
- Further subdivisions of vowels: *monophthongs* and *diphthongs*
  - Monophthongs ([mɑnəpθɑŋz]): simple vowels, with a single VTC
  - Diphthongs ([dɪpθɑŋz]): complex vowels, with a sequence of two VTCs
  - Note: some IPA symbols used above to slowly ease us into using them comfortably
  - Diphthongs are single vowels, since the sequence is the nucleus in the syllable
  - e.g. rhyme of `knives` (diphthong: [ɑɪ]) vs. `naive` (monophthongs: [ɑ] followed by [i])
  - More on these distinctions later on
- IPA phonetic symbols:
  - IPA symbols are written in square brackets ([]), as seen above
  - This distinguishes them from ordinary words, as there is some syntactic overload
  - IPA symbols are NOT letters of English—they represent sounds of language
  - We can break English sounds into a *phoneme inventory* of sounds and give the phonetic symbols
  - Important: we combine sounds in day-to-day *running/continuous speech*, so some pronunciations may end up being affected by surrounding words (context, phonological assimilation, etc.—more later!)
  - Open research question: How does the human mind process running speech into meaningful constituent parts?

- A complete IPA for English segments (keep in mind that pronunciation can vary, so the examples may not be universal):

Symbol	Examples	Symbol name

Figure 2.2: IPA symbols for consonants found in English. Keep in mind that a symbol's name differs from its meaning.