Representations of Speech

- Acoustic - Frequency & Intensity over Time
- Auditory - Pitch, Spectral Shape, Spectral/temporal Features
- Phonetic Features - Articulatory or Abstract
- Phonemes - Segmental, Language Specific
- Morphemes - Meaning Units
- Mora - Rhythmic Units, 1 or more Phonemes
- Syllables - Sequences of 1 or more Phonemes
- Words - Sequences of Phonemes/Syllables

Background

- In order to understand speech perception, need to know the basics of articulation.
- In order to understand speech perception, need to know the acoustics of speech.
- Understanding speech perception entails an understanding of hearing and the auditory system.
- Speech perception is both a biologically specialized process and a basic cognitive process (memory, attention, categorization).

Phonemes - American English Vowels

- /i/ - heed
- /ɪ/ - hid
- /e/ - hayed (diphthong)
- /ɛ/ - head
- /æ/ - had
- /aɪ/ - hide (diphthong)
- /oɪ/ - void (diphthong)
- /a/ - sod
- /ɔ/ - hawed
- /o/ - hoed
- /ʊ/ - hood
- /u/ - who’d
- /aʊ/ - how’d (diphthong)
- /ʌ/ - hud
- /ɚ/ - herd

Recording & Analysis Project

This project involves recording hVd words (and nonwords) that represent the vowel space of American English and measuring the formant frequencies for each token.

Record 3 tokens of each of the 15 vowels in hVd syllable frame in a fluent sentence. The formant frequencies are to be measured at the beginning, middle and end of the vowel.*

American English Consonants

stops & nasals

- /b/ - **b**at
- /p/ - **p**at
- /m/ - **m**at
- /d/ - **d**ot
- /t/ - **t**ot
- /n/ - **n**ot
- /g/ - **g**ap
- /k/ - **c**ap
- /ŋ/ - si**ng**

fricatives

- /f/ - **f**at
- /v/ - **v**at
- /θ/ - **th**in
- /ð/ - **th**is
- /s/ - **s**ip
- /z/ - **z**ip
- /ʃ/ - a**ss**ure (**sh**ip)
- /ʒ/ - a**z**ure
- /h/ - **h**ip

Consonants - Part 2

affricates

- /tʃ/ - **ch**ug
- /dʒ/ - **j**ug

approximates

- /w/ - **w**att
- /l/ - **l**ot
- /r/ - **r**ot
- /y/ - **y**acht

Synthesis Project

This project is designed to help you learn the two basic methods used to generate sets of syllables used in perceptual experiments. It consists of two components.

For one, use the speech synthesizer to make a synthetic VOT continuum that goes from /ba/ - /pa/ or /da/ - /ta/. For the second, I will provide a recording of a natural speech token (vowel) and you are to create a synthetic copy that sounds like the original talker.

Short Papers 1 and 2

There will be two short papers, one due at the class midpoint and the second due during finals week. For each paper, there will be two or three questions from which you pick one. Papers are to be about 5 - 6 typed, double spaced pages.

Example Topic:

Are phonetic distinctions cued by invariant acoustic properties? Cite data on both the acoustics and perception and be sure and define what you mean by invariant.

Mind and Brain

We believe that the brain provides the physical underpinnings for the mind. This can be seen in the description provided by Marr for the three levels of explanation:

1. Information Processing - Here, the goals of the system, its role in the environment, and the transformations involved between input and output are relevant. This is the most "abstract" level of description.

2. Algorithm - This involves a formal characterization of the transformations into an unambiguous language (e.g. math, computer program).

3. Hardware - The physical mechanism that implements the algorithm.

Our concern is levels 1 and 2. However, understanding what is known about level 3 constrains the nature of level 2. This, in turn, influences how we formulate level 1. Thus, neurophysiology is important to our understanding of language and perception. This will be apparent in our examination of articulation and basic hearing.

Issues in Speech Perception

Units and Representation - What are the objects of perception? What representations support word recognition?

Segmentation - Fluent speech is a continuous signal yet listeners perceive it as a sequence of discrete words.

Signal Variability - Listeners achieve constancy yet signal varies due to talker (ideolect, speaking rate, physiology), context (coarticulation, affect, pragmatics), environment (noise, other talkers,…).

Studying Perception

1. Articulation - Since the source of the signal is the talker, understanding articulation may provide insight into production and perception.
2. Correlation - Speech acoustics will show the acoustic qualities that are correlated with the talker's articulation.
3. Cues - Manipulating acoustic correlates in a perception study will show if the correlate "cues" a speech distinction.
4. Computation - Comparing different computational descriptions of cues will reveal how listeners process cues.

Instrumentation

For measuring the acoustic qualities of speech, we'll use Praat, available at:

http://www.fon.hum.uva.nl/praat/

Praat can be used to record and make temporal or spectral measures.

Sample spectrogram