SCIENCE OF TALKING – DISORDERS OF SPEECH PRODUCTION

JYRKI TUOMAINEN

SPEECH, HEARING AND PHONETIC SCIENCES

UNIVERSITY COLLEGE LONDON

January 14, 2010

GENERAL OUTLINE

1. A MODEL OF LANGUAGE PROCESSING
   - HOW IS LANGUAGE PROCESSING DISTRIBUTED IN THE BRAIN?
   - WHAT HAPPENS WHEN SOMETHING GOES WRONG IN THE BRAIN?
2. A MODEL OF SPEECH PRODUCTION
   - WHAT HAPPENS WHEN WE NAME A PICTURE?
   - WHICH AREAS OF THE BRAIN ARE RESPONSIBLE FOR SPEECH PRODUCTION?
3. INTELLIGIBILITY OF SPEECH
   - WHAT IS THE RELATIONSHIP BETWEEN SPEECH PRODUCTION AND SPEECH PERCEPTION?
   - ACOUSTIC AND PERCEPTUAL ASPECTS OF INTELLIGIBILITY

DISORDERS OF LANGUAGE AND SPEECH MOTOR CONTROL

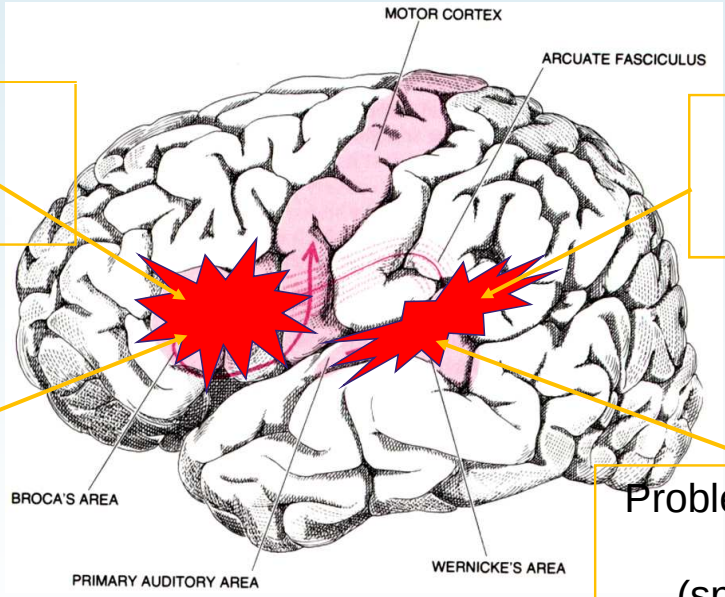
DEFINITIONS:

- **Aphasias** are acquired neurological *language disorders* that affect production and understanding of spoken, written, and signed language.
- **Dysarthrias** are speech disorders that result from neurologic impairments associated with *weakness, slowness, or incoordination of the musculature to produce speech*.
- **Apraxia of speech** in adults is a result of a focal brain damage that impairs especially the *processes of planning or programming speech movements* in the face of essentially normal strength, speed and coordination of the speech musculature.

WERNICKE-LICHTHEIM MODEL

Articulatory speech "memories"

Auditory speech "memories"

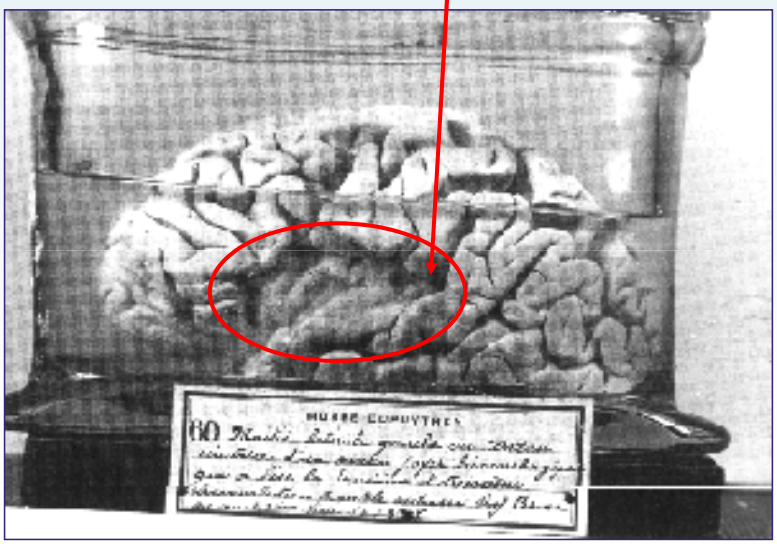


Problems with speech production → Broca's aphasia)

Problems with speech comprehension (speech production also affected!) → Wernicke's aphasia

WERNICKE-LICHTHEIM MODEL

Broca's 1st patient Leborgne's ("Tan") brain in formaldehyde solution. The damage is located in the posterior part of the inferior frontal gyrus (Broca's are) and extends to the posterior part of the temporal lobe. The damage is massive but Broca insisted that the posterior part of the *left inferior frontal gyrus* was the cause of the speech production problem.



A MODEL OF LANGUAGE PROCESSING

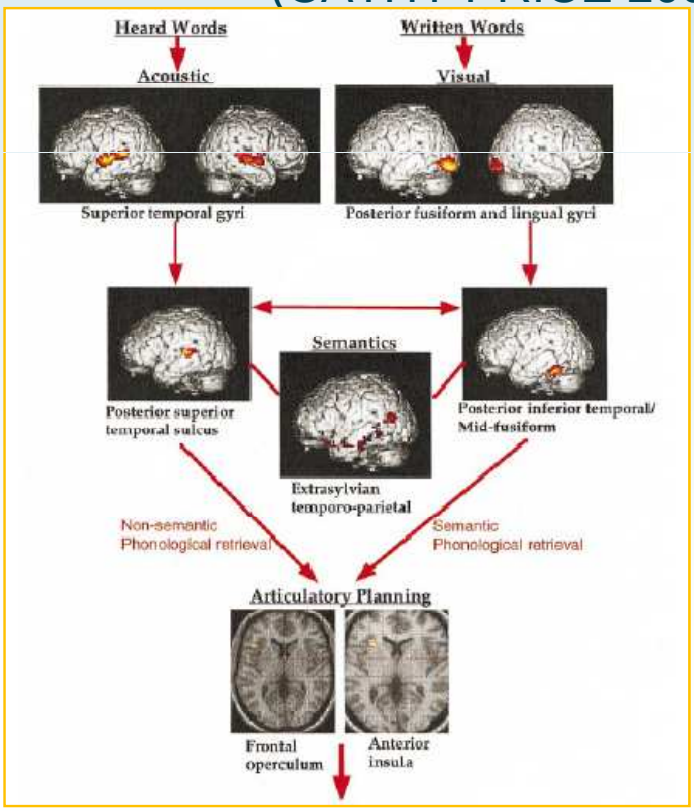
(CATHY PRICE 2000)

Auditory processing of speech occurs *bilaterally* in the temporal cortex

Visual processing of written words occurs *bilaterally* in the occipito-temporal cortex.

Phonological and semantic processing occurs in the *left* temporo-parieto-occipital areas (Wernicke's area).

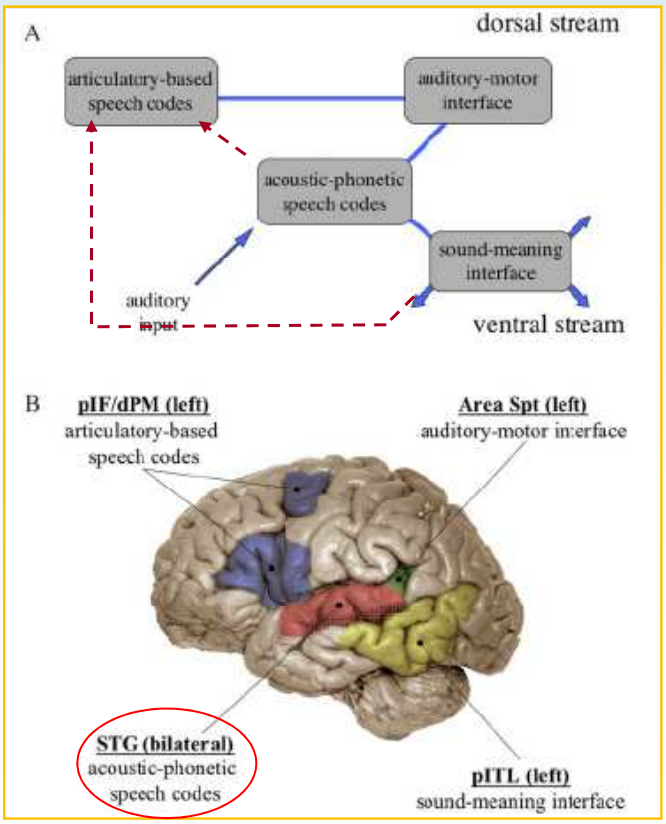
A common semantic system for understanding heard and written language.



A MODEL OF AUDITORY SPEECH PROCESSING (HICKOK & POEPPEL 2004)

- Early auditory-phonetic processing occurs *bilaterally*.
- The type of the *task* affects which neural network is activated even with same stimuli.
- Ventral stream for phonetic, phonological and semantic processing.
- Dorsal stream links sensory and motor areas in speech perception and production.

→ Auditory areas are activated during speech production and motor areas are activated during speech perception!



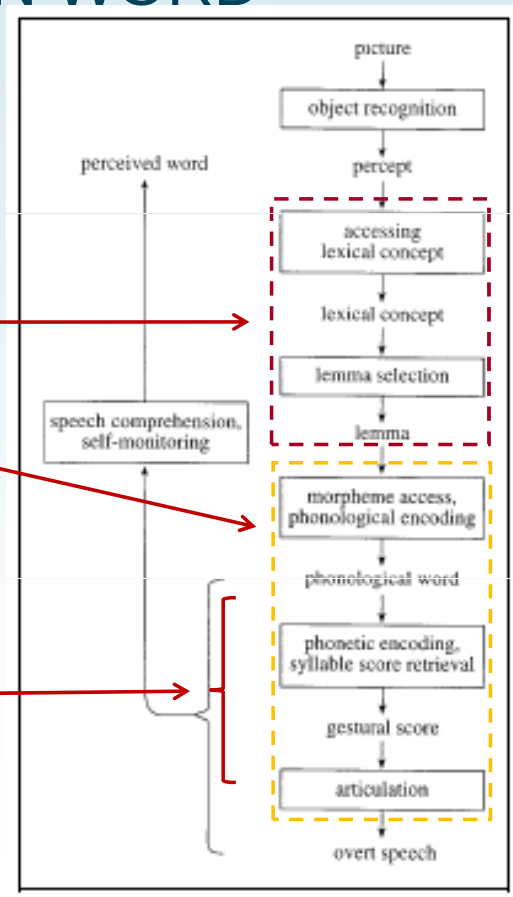
A MODEL OF SPOKEN WORD PRODUCTION

Current standard model:

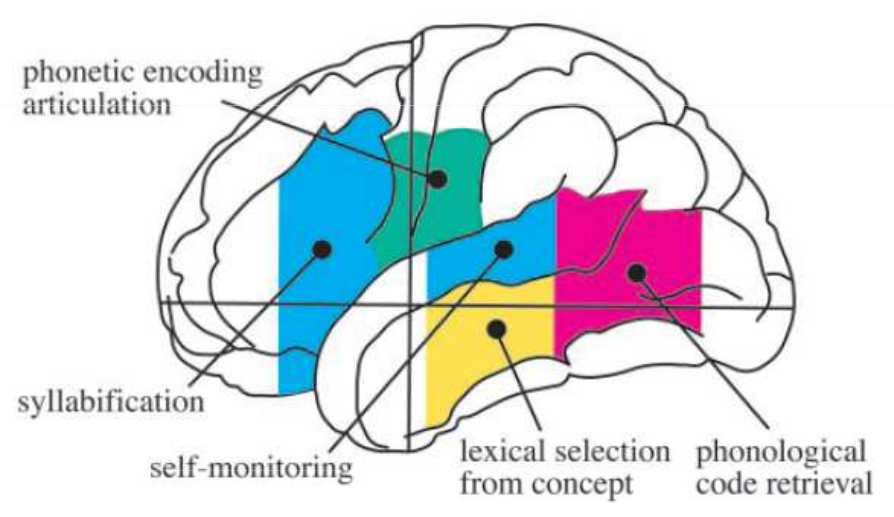
A serial two system model:

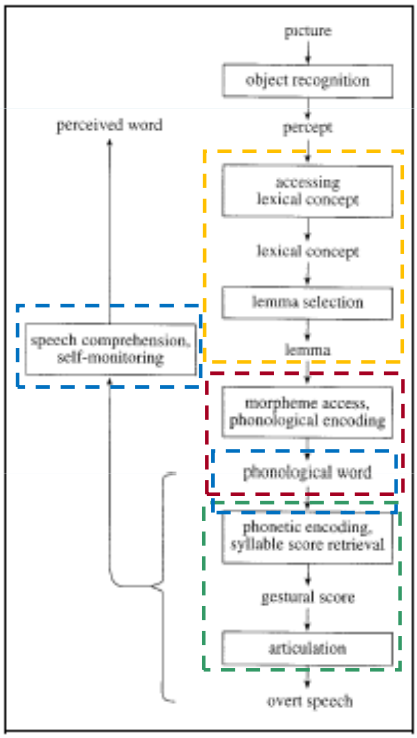
- LEXICAL SELECTION (two stages)
- FORM ENCODING (three stages)

Disorders in speech motor control (apraxia of speech and dysarthrias) result from a problem at phonetic and execution levels!



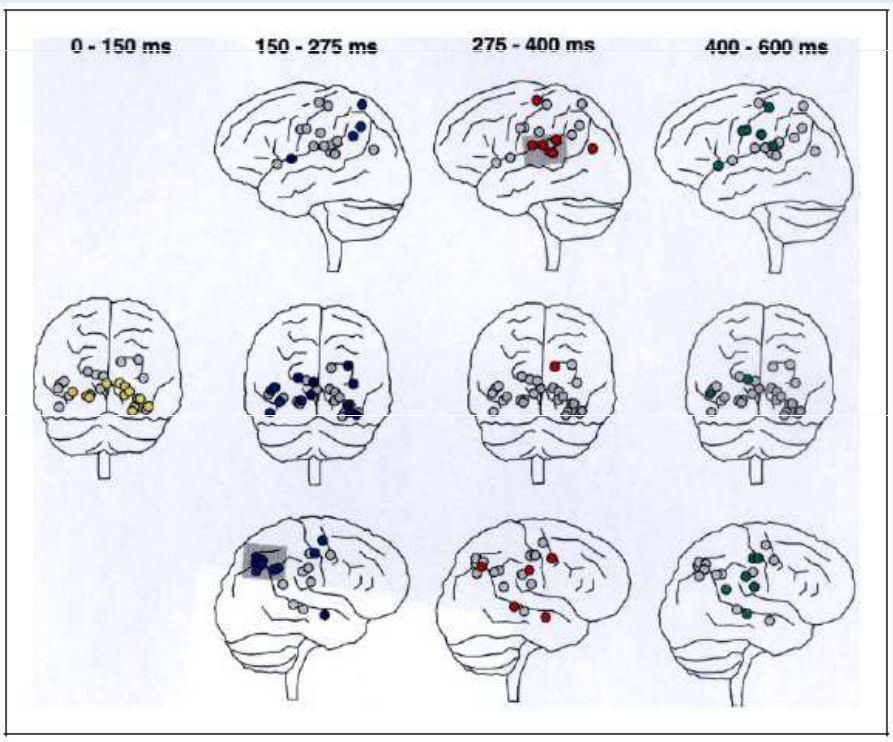
SPOKEN WORD PRODUCTION AND THE BRAIN

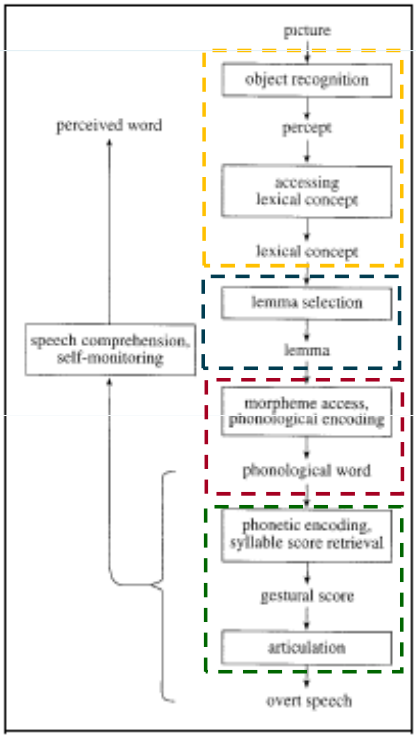




What about speed? How fast do subjects name pictures? What is your guess?

TIME COURSE OF SPOKEN *WORD* PRODUCTION





MODELING SPEECH PRODUCTION

PROBLEMS (1)

- **Serial order** in a processing system with multiple subcomponents and a serial output system (Lashley, 1951) (e.g., *Phonological sequencing*)
- Assume a system that is capable of producing two responses, R1 and R2, that can be evoked by two stimuli, S1 and S2 (S-R associations).
- If S1 and S2 and simultaneously activated, the system has to possess a mechanism to either *select* R1 or R2, or to *assign a sequence* in which to the execute of R1 and R2.

- *Serial order cannot be implemented using associative chaining*. More likely, parallel activation of subcomponents can be handled by competitive queuing. (Think about how a bartender serves customers: The loudest customer gets the drink first!)

MODELING SPEECH PRODUCTION

PROBLEMS (2)

- WHAT IS THE BASIC UNIT OF SPEECH PRODUCTION?

Speed of information transfer:

<u>Phoneme</u>

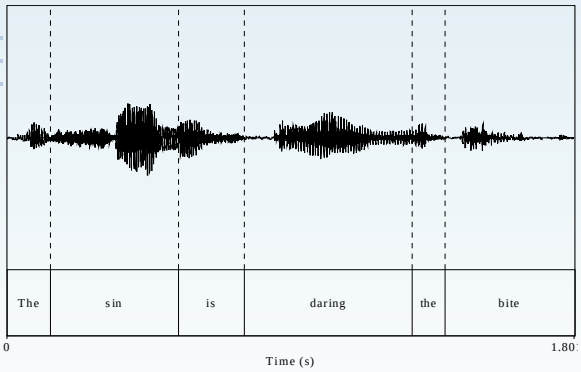
(~10 sounds/sec)

<u>Syllable</u>

(~3.8 syll/sec)

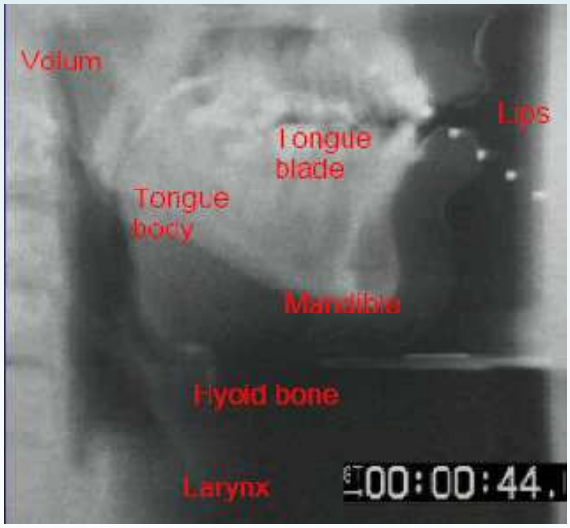
<u>Word</u>

(~3.3 words/sec)



- Syllable seems to be the unit that specifies the motor programs (large muscles of the tongue and the mandible form the basic articulatory unit (”carrier wave”) on top of which other articulatory movements are attached to).
- Syllable effect refers to experimental finding which shows that it takes somewhat longer to initiate articulation of a two-syllabic word compared a one-syllabic word (of equal length). This suggests all the syllables in a word (~ ”phonological phrase”) need to be planned before it is possible to start articulation.

SPEECH MOTOR CONTROL



- Approximate number of muscle pairs that move the
  - Tongue: 9
  - Velum: 3
  - Lips: 12
  - Mandible: 7
  - Hyoid bone: 10
  - Larynx: 8
  - Pharynx: 4
- NB: The respiratory system

http://www.phonetics.ucla.edu/course/chapter1.1/chapter1.1.htm

- Possibly *the* most complex motor behavior
  - A large number of degrees of freedom
  - A very complicated control problem

CLASSIFICATION OF DYSARTHRIAS

DE FACTO STANDARD:

– Darley, Aronson, Brown (1969, 1975); Perceptual and neurological characterization of speech motor disorders

TABLE I. Clinicoanatomic relationships hypothesized by Darley *et al.* (1969*a*, *b*). Shown for each perceptual type of dysarthria is the primary lesion site

Dysarthria type	Primary lesion site
Flaccid	Lower motor neuron (one or more cranial nerves)
Spastic	Upper motor neuron
Spastic–flaccid	Both upper and lower motor neurons
Ataxic	Cerebellum or its outflow pathways
Hypokinetic	Basal ganglia, especially substantia nigra
Hyperkinetic	Basal ganglia, especially putamen or caudate

CLASSIFICATION OF DYSARTHRIAS

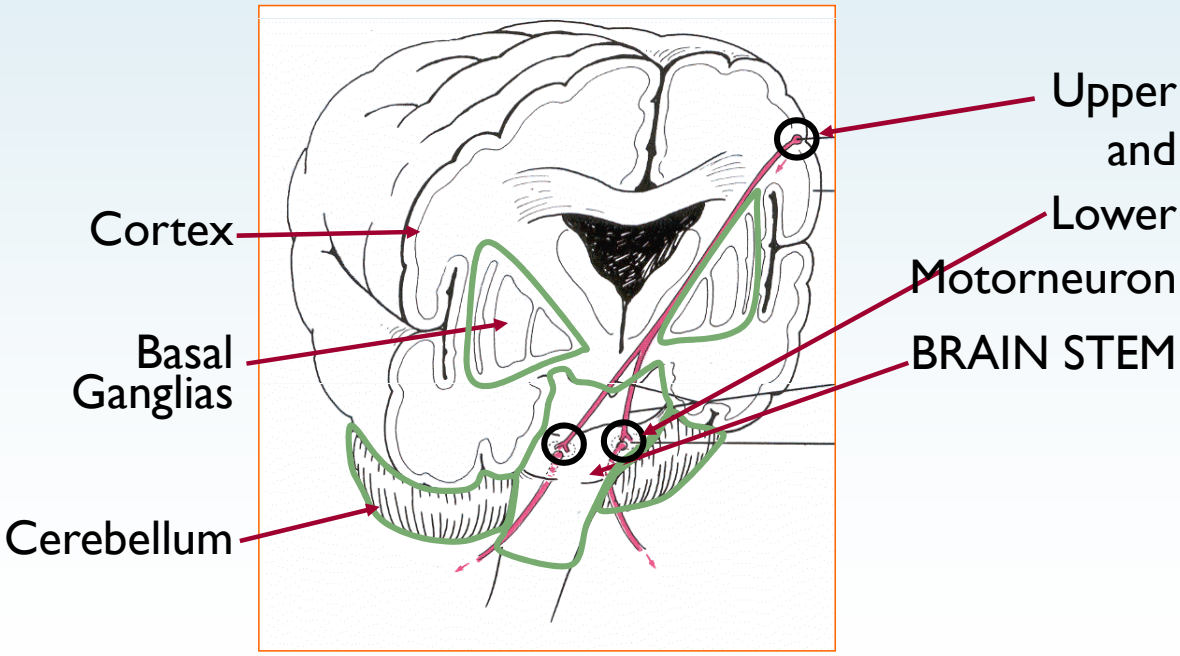
Darley, Aronson, Brown (1969, 1975); Perceptual and neurological characterization of speech motor disorders

TABLE II. Major clusters of deviant perceptual dimensions for dysarthria, as reported by Darley *et al.* (1999*a*, *b*)

Type of dysarthria	Clusters of deviant dimensions
Ataxic dysarthria	Articulatory inaccuracy, prosodic excess, phonatory–prosodic insufficiency
Spastic dysarthria	Prosodic excess, prosodic insufficiency, articulatory–resonatory incompetence
Flaccid dysarthria	Phonatory incompetence, resonatory incompetence, phonatory–prosodic insufficiency
Spastic-flaccid dysarthria	Prosodic excess, prosodic insufficiency, articulatory–resonatory incompetence, phonatory stenosis, phonatory incompetence, resonatory incompetence
Hypokinetic dysarthria	Prosodic insufficiency, phonatory incompetence
Hyperkinetic dysarthria (Chorea)	Articulatory inaccuracy, prosodic excess, prosodic insufficiency, articulatory–resonatory incompetence, phonatory stenosis
Hyperkinetic dysarthria (Dystonia)	Articulatory inaccuracy, prosodic excess, prosodic insufficiency, phonatory stenosis

A BIT OF NEUROLOGY

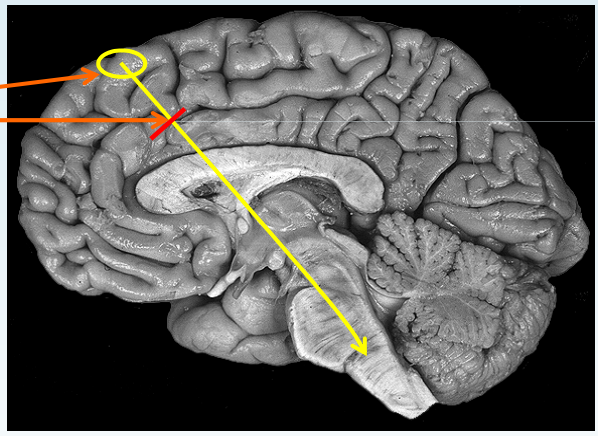
MACROANATOMY OF THE BRAIN



EXAMPLE: SPASTIC DYSARTHRIA

TYPICAL LOCATION OF THE LESION:

- Upper motor neuron
  - Bilateral lesion produces chronic symptoms
  - Unilateral lesion: Transient symptoms (most head muscles are bilaterally innervated!)

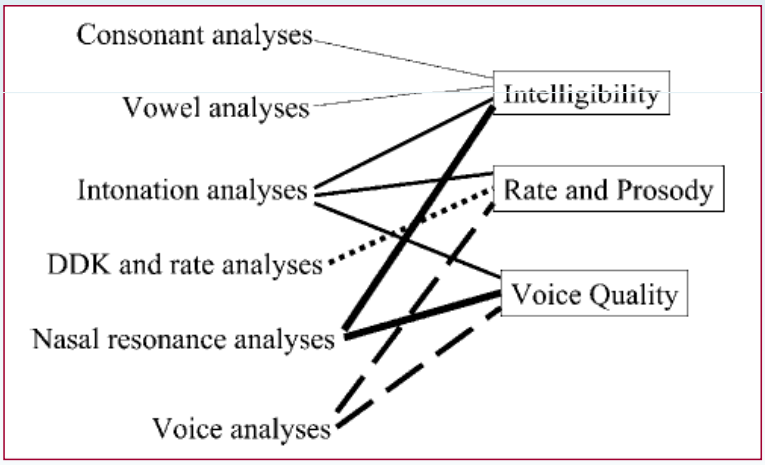


- TYPICAL NEUROLOGICAL SYMPTOMS:
  - Spasticity of muscles, reduced speech of movements, (hyper)sensitive reflexes; in severe cases primitive reflexes present (involuntary biting, sucking etc.)
- SPEECH:
  - Slow speech rate, imprecise consonants
  - Flat prosody (especially intonation)
  - Voice quality usually rough and stringent
  - Hypernasality

PERCEPTUAL EVALUATION OF INTELLIGIBILITY OF SPEECH

SETTING THE SCENE:

Speech subsystems related to intelligibility involve consonantal and vowel contrasts, prosody, nasality (but strangely, not voice; cf. Parkinson's disease)



Here's something more on voice:
http://www.youtube.com/watch?v=52UAEQfMTtU

AUDITORY-PERCEPTUAL EVALUATION OF INTELLIGIBILITY IS THE CLINICAL STANDARD!

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

1. GENERAL ACOUSTIC PARAMETERS EMPLOYED IN THE ANALYSIS OF DYSARTHRIC SPEECH
2. WHICH PARAMETERS ARE USEFUL IN DETERMINING ACOUSTIC CORRELATES OF SPEECH INTELLIGIBILITY?
3. WHICH PARAMETERS ARE USEFUL FOR THE CURRENT DATA?

KEY READINGS:

- Kent, R., Weismer, G., Kent, J., Vorperian, H., & Duffy, J. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders*, 32, 141-186.
- Kent, R. & Kim, Y.-J. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, 17(6), 427-445.

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

VALIDITY AND RELIABILITY OF ACOUSTIC MEASURES

- Inter- and intra-judge reliability measures should be provided!
- *Annotation accuracy:* For calculating a correlation coefficient, ask another judge to analyze 10 to 20% of the data (inter-judge). Same judge re-analyzes ~ 10% of the data to obtain intra-judge estimate. (Coefficients should be >.75).
- *Formant frequencies* (F1, F2, F3) can, with high-quality recordings, be tracked within +/- 60Hz from LPC spectra. Don't use the spectrogram for this!
- *Duration measures* from the waveform suggests an accuracy of about 10 to 25 ms.
- *Perturbation measures* (jitter and shimmer) are vulnerable to poor quality recordings increasing variability in the results! Pay attention to the quality of the recording!
- **Reliability measures are poorer for dysarthric speech!**

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

ACOUSTIC MEASURES AND SPEECH MOTOR CONTROL

VOCAL TRACT FUNCTION FOR CONSONANT ARTICULATION: SPECTRAL CORRELATES

- Spectrum of stop or fricative noise;
- Spectral moments of stop burst or fricative noise;
- Formant transitions of CV and VC transitions

VOCAL TRACT FUNCTION FOR CONSONANT ARTICULATION: ACOUSTIC-ARTICULATORY RELATIONSHIP

- *Sonorants* (glides, liquids and nasals):
  - Formant/antiformant patterns, steady-state vs. transitional segments
- *Non-sonorants* (stops, affricates and fricatives):
  - Frication event (a burst or transient noise for stops, a brief noise for affricates, and a longer interval for fricatives)

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

ACOUSTIC MEASURES AND SPEECH MOTOR CONTROL

CONSONANT ARTICULATION IN DYSARTHRIA:

- Reduction of first spectral moment (center of gravity) correlates with perceptual estimates of consonant precision.
- Averaged slope of the F2 transition in CV sequences correlates with perceptual measures (overall intelligibility index).
- Energy measures during voiceless stop gap: Increased energy in the silent interval is an index of either turbulent noise because of incomplete closure of vocal tract, or voicing due to poor laryngeal control.

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

ACOUSTIC MEASURES AND SPEECH MOTOR CONTROL

SOUND SEGMENT TIMING:
ACOUSTIC MEASURES
- Segment durations

SOUND SEGMENT TIMING:
ACOUSTIC-ARTICULATORY RELATIONSHIP
- Fairly straightforwardly related to certain steady state or transitional physiologic events in laryngeal and vocal tract functions

SOUND SEGMENT TIMING IN DYSARTHRIA
- Due to slowed speaking rate, segment durations are longer in general.
- Normalize the durations → more lengthening for consonants and vowels, but not for VOT!

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

CURRENT DATA:

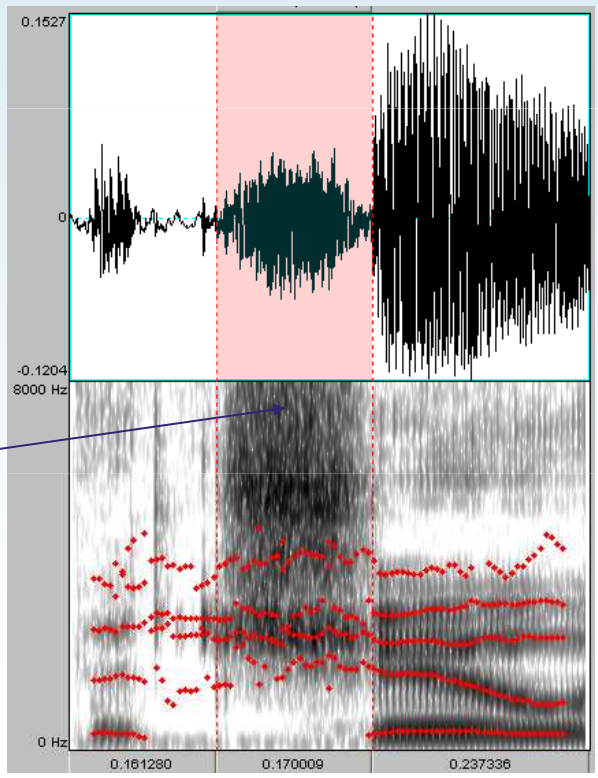
- *Syllable initial fricative*: (back vowel context: “sue/shoe”)
- *Syllable initial fricative* (front vowel context: “sin/shin”)
- Main acoustic cues: *Duration of the frication*, *spectral moments* (e.g., centre of gravity)

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

CURRENT DATA:

- *Syllable initial fricative*: (back vowel context: “sue/shoe”)
- *Syllable* initial (front vowel context: “sin/shin”)

- Main acoustic cues: *Duration of the frication*, *spectral moments* (e.g., center of gravity



ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

INSTRUCTIONS:

1. Download from \\bell\btemplab a folder “sounds” and place it on your desktop
2. Open Praat
3. “Read from file”: open file “sin.wav” in folder JP (the normal control)
4. Analyse the spectral moment and the duration of the /s/ sound.
5. Enter data on Excel

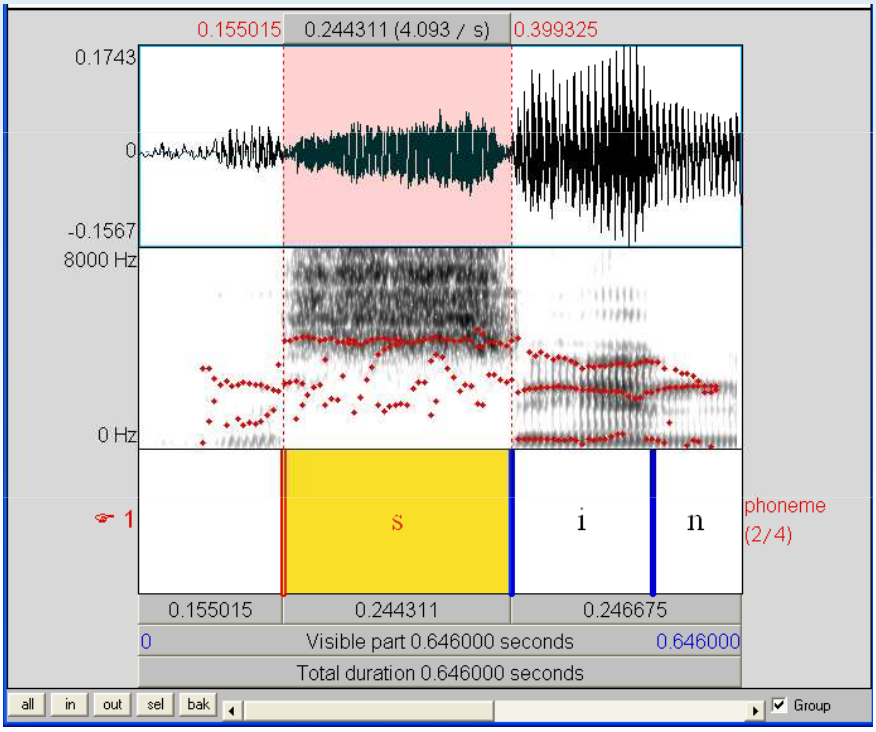
REMEMBER: You always need a control sample to obtain a reference point. The control sample should consist of either normal non-neurologic speakers or another neurologic population without speech disorders.

ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

INSTRUCTIONS:

Start by annotating the segments.

Read file (e.g. sin.wav); create textgrid; use 1 tier (name it e.g. ”phoneme”)



ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

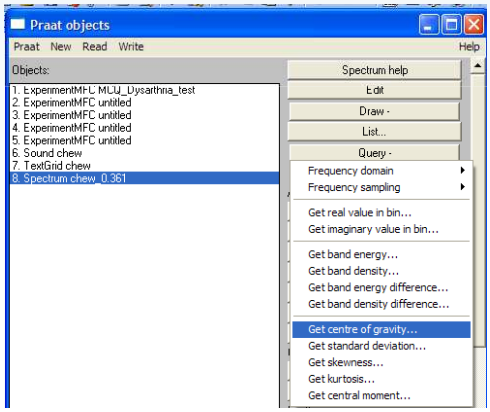
INSTRUCTIONS:

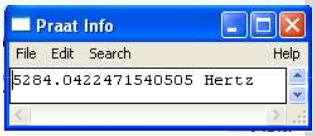
1. Analyse now the same “fricative words” from one of the dysarthric speakers (MH) to get used to the analysis procedure.
2. Analyze *Center of gravity*, *duration of the fricative noise*.
3. Enter the values on the Excel sheet on the designated columns.

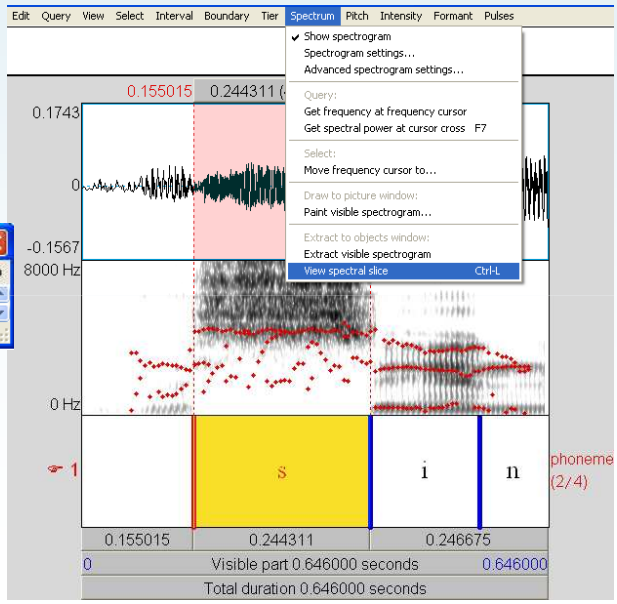
ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

How to measure Center of gravity (1st spectral moment)

- Select the fricative noise part of the signal
- Click on Spectrum → View spectral slice
- Go to Object window
- Select Spectrum object
- Click on Query,
- Get CoG (use power: 2)
- Read the value from Info window







ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

Enter data on Excel sheet, a sample file (data.xls).

	A	B	C	D	E	F	G	H	I
1		Acoustic variables							
2	Item	JP COG Hz	JP Noise duration		MH COG Hz	MH Noise duration		JF COG Hz	JF Noise duration
3									
4	*shoe*								
5	*sue*								
6	*shin*								
7	*sin*								
8									
9	***Total mean***	#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!
10	***Mean /s/***	#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!
11	***Mean /sh/***	#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!		#DIV/0!	#DIV/0!
12									
13	*COG*								
14	*t-test JP/MH*	#DIV/0!							
15									
16	***Duration***								
17	*t-test JP/MH*	#DIV/0!							
18									

- NOTE: Use coefficient of variation to get a hold of variability (CV=S.D./M)
- Increased variability is one measure of system instability!

THINGS TO PONDER ON...

1. A MODEL OF LANGUAGE PROCESSING
   - Language processing is distributed all over the place in the brain, but some aspects mainly in the left hemisphere.
   - When something goes wrong in the brain the symptoms depend on the site of the lesion.
2. A MODEL OF SPEECH PRODUCTION
   - Naming a picture consists of multiple processing stages and can be accomplished in appr. 600 milliseconds.
   - Brain areas activated during naming are located in the temporal and frontal brain areas, mostly but not exclusively in the left hemisphere.
3. INTELLIGIBILITY OF SPEECH
   - Speech production and speech perception are intimately tied together. Brain activation pattern depend on task requirements.
   - We are still far away from knowing the acoustic and perceptual correlates of speech intelligibility.