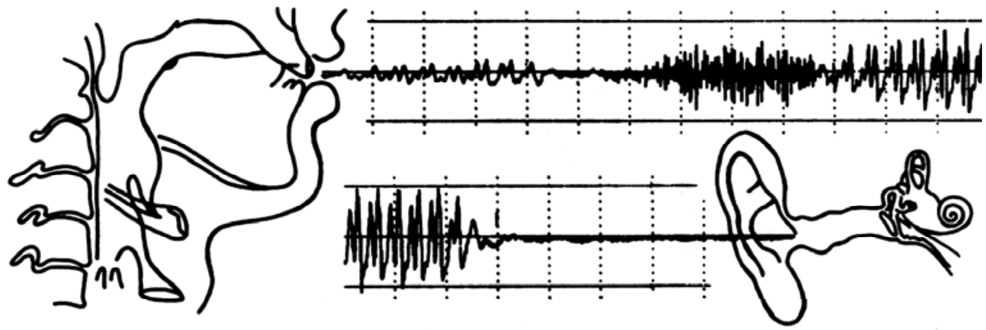
Harvard-MIT Division of Health Sciences and Technology
HST.722J: Brain Mechanisms for Hearing and Speech
Course Instructor: Joseph S. Perkell

Motor Control of Speech: Control Variables and Mechanisms



HST 722, Brain Mechanisms for Hearing and Speech

Joseph S. Perkell
MIT

Outline

- Introduction
- Measuring speech production
- What are the "controlled variables" for segmental (phonemic) speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

Outline

- Introduction
  - Utterance planning
  - General physiological/neurophysiological features
  - The controlled systems
  - Example of movements of vocal-tract articulators
- Measuring speech production
- What are the “controlled variables” for segmental speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

Utterance Planning

- Objective: generate an intelligible message while providing for □ "economy of effort" – stages: □
  - Form the message (e.g. Feel hungry; smell pizza; together with a friend).
  - Select and sequence lexical items (words). "Do you want a pizza?"
  - Assign a syntactically-governed prosodic structure.
  - Determine "postural" parameters of overall rate, loudness and degree of **reduction** (and settings that convey emotional state, etc.)
    - Extreme reduction: "Dja wanna pizza?"
  - Determine temporal patterns: Sound segment durations depend on:
    - Phoneme length
    - Overall rate
    - Intrinsic characteristics of sounds
    - Position and number of syllables in word
- Result: an ordered sequence of goals for the production mechanism

Serial Ordering

- Evidence reflecting serial ordering in utterance planning: speech errors
  - Examples from Shattuck-Hufnagel (1979)
    - Substitution — Any*m*ay — (Anyway)
    - Exchange — e*m*e*n*y — (enemy)
    - Shift — bad high*v*way dri_ing — (highway driving)
    - Addition — the p*l*ublicity would be — (publicity)
    - Omission — sonata _umber ten — (number)
    - ? — dignal sigital processing
  - See Averbeck et al. on neurophysiologial evidence concerning serial ordering

General Physiological/Neurophysiological Features

- Muscles are under voluntary control
- Structures contain feedback receptors that supply sensory information to the CNS:
  - – Surfaces: touch/pressure
  - – Muscles:
    - length and length changes: spindles
    - Tension: tendon organs
  - – Joints (TMJ): joint angle
- Reflex mechanisms:
  - – Stretch
  - – Laryngeal (coughing)
  - – Startle
- Motor programs (low-level, “hard wired” neural pattern generators)
  - – Breathing
  - – Swallowing
  - – Chewing
  - – Sucking
- Low-level circuitry could be employed in speech motor control. The picture is complex, and a comprehensive account hasn’t emerged.

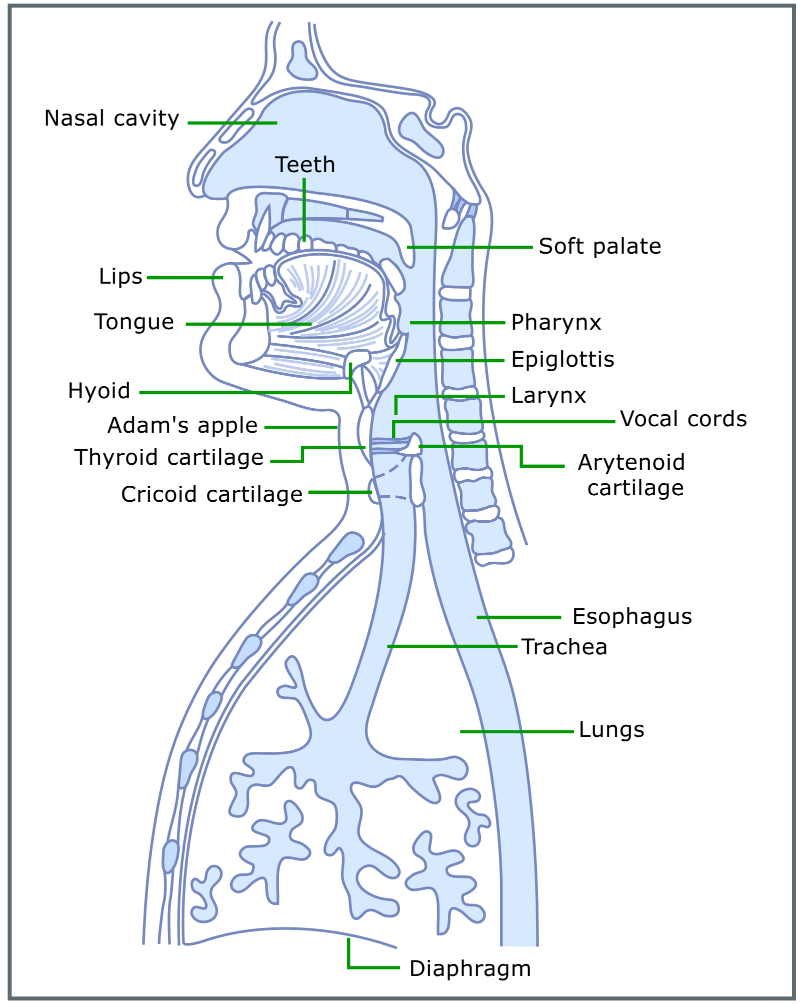


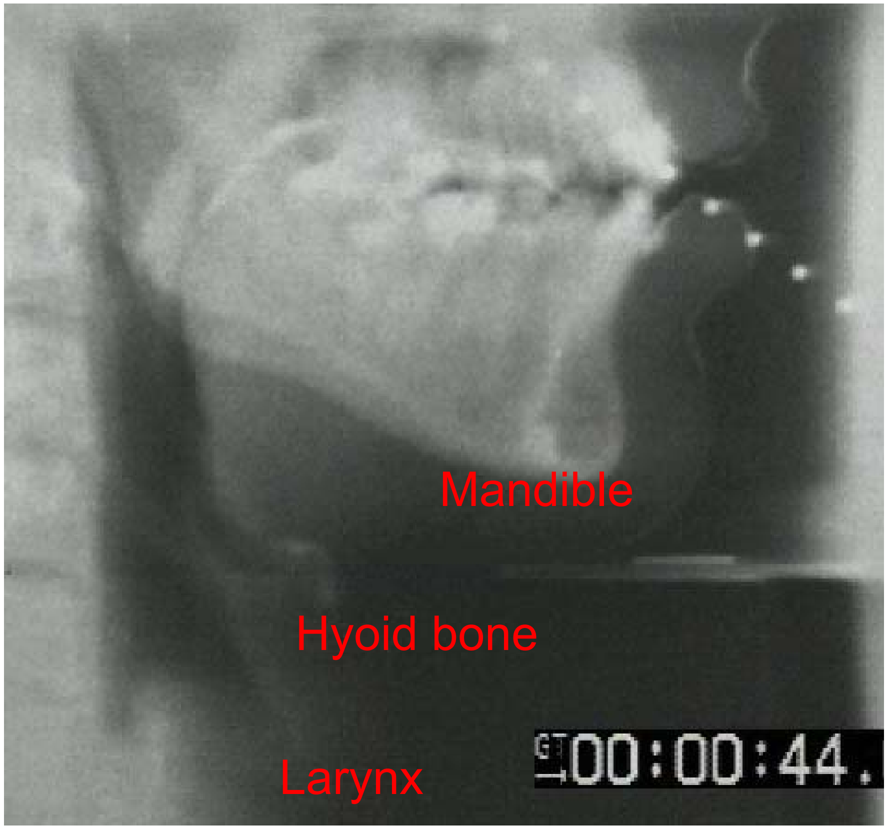
Figure by MIT OCW.

The controlled systems

- The respiratory system
  - most massive (slowly-moving structures)
  - Provides energy for sound production
    - Fluctuations to help signal emphasis
    - Relatively constant level of subglottal pressure
  - Different patterns of respiration: breathing, reading aloud, spontaneous, counting
  - Different muscles are active at different phases of the respiratory cycle – a complex, low-level motor program
- Larynx
  - Smallest structures, most rapidly contracting muscles
  - Voicing, turned on and off segment-by-segment
  - F0, breathiness – suprasegmental regulation
- Vocal tract
  - Intermediate-sized, slowly moving structures: tongue, lips, velum, mandible
  - Many muscles do not insert on hard structures
  - Can produce sounds at rates up to 15/sec
  - To do so, the movements are coarticulated

Figures removed due to copyright reasons.
Please see:
Conrad, B., and P. Schonle. "Speech and respiration." *Arch Psychiatr Nervenkr* 226, no. 4 (1979): 251-68.

Focus of lecture is on movements of vocal-tract articulators



- Consider the movements of each of these structures
- Approximate number of muscle pairs that move the
  - Tongue: 9
  - Velum: 3
  - Lips: 12
  - Mandible: 7
  - Hyoid bone: 10
  - Larynx: 8
  - Pharynx: 4
- Not including the respiratory system

- Observations:
  - A large number of degrees of freedom
  - A very complicated control problem

Outline

- Introduction
- Measuring speech production
  - Acoustics
  - Articulatory movement
  - Area functions
- What are the "controlled variables" for segmental speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

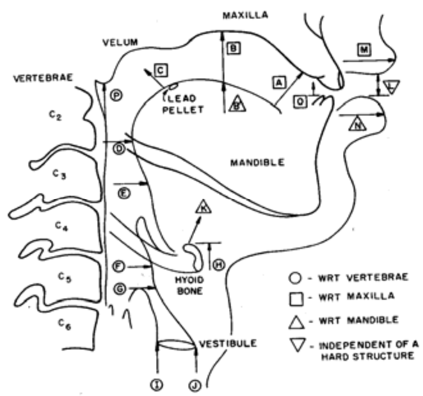
Measuring Speech Production

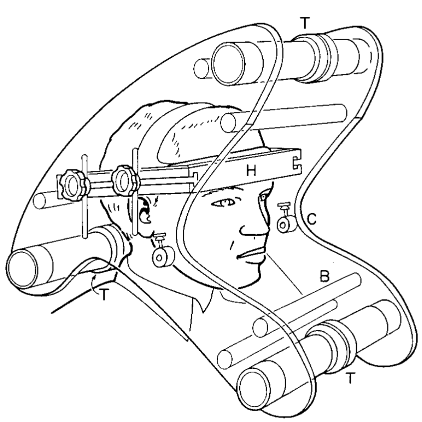
- Acoustics – important for perception
  - Spectral, temporal and amplitude measures
    - Vowels, liquids and glides:
      - Time varying patterns of formant frequencies
    - Consonants:
      - Noise bursts
      - Silent intervals
      - Aspiration and frication noises
      - Rapid formant transitions

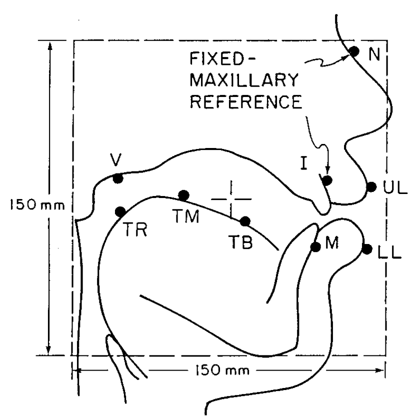
Figure removed due to copyright reasons.
Please see:
Steven, K. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1998, p. 248. ISBN: 026219404X.

"The yacht was a heavy one"

From: Perkell, Joseph S. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Research Monograph No. 53. Cambridge, MA: MIT Press. 1969(c). Used with permission.



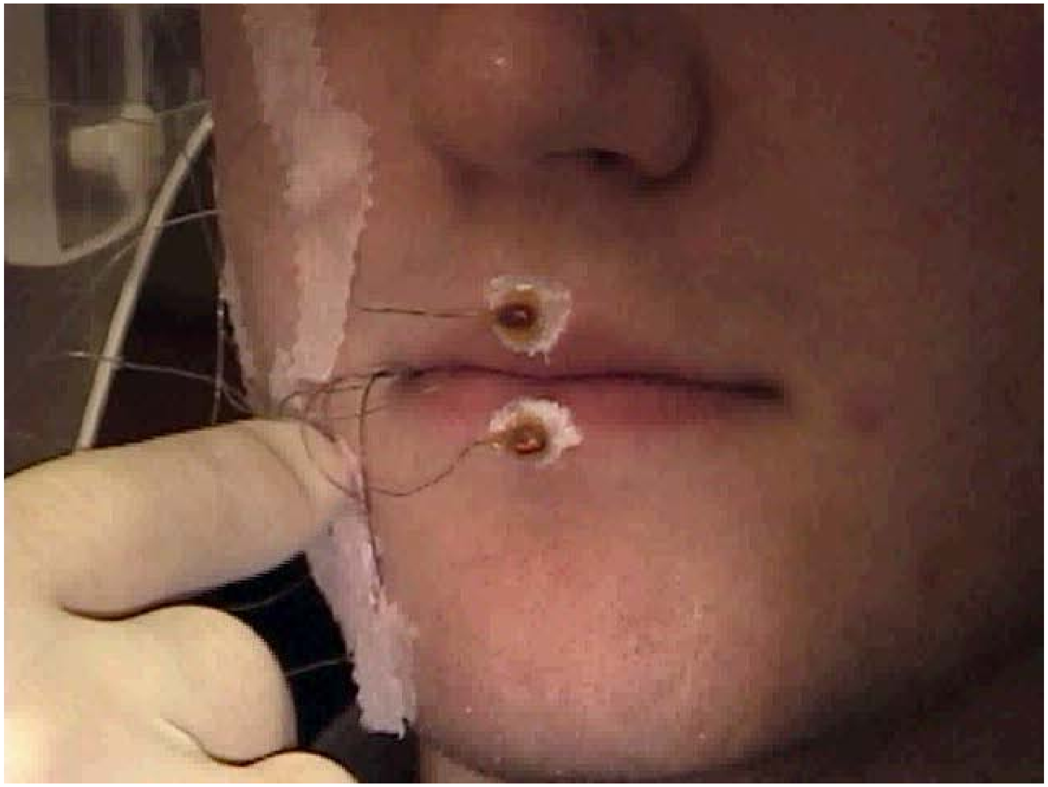




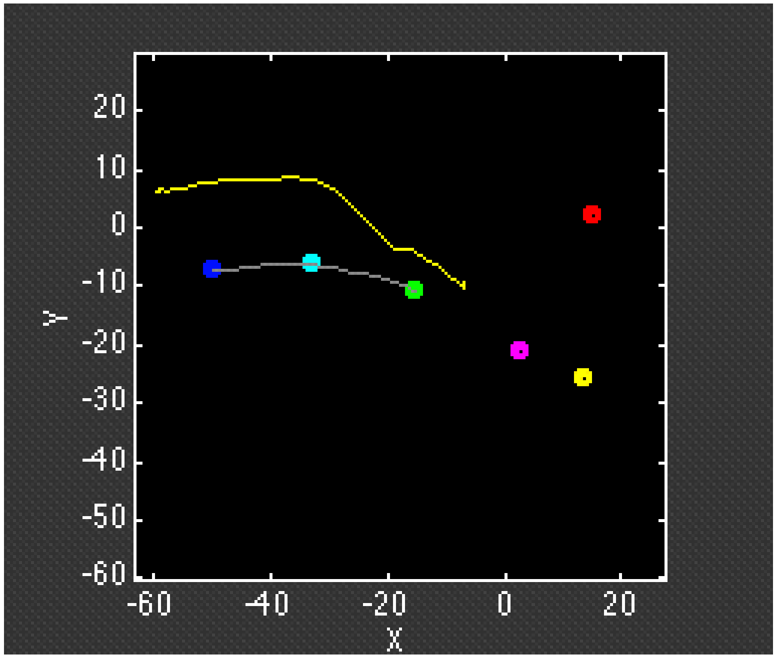
Reprinted with permission from:
Perkell, J., M. Cohen, M. Svirsky, M. Matthies, I. Garabieta, and M. Jackson. "Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *J Acoust Soc Am* 92 (1992): 3078-3096. Copyright 1992, Acoustical Society America.

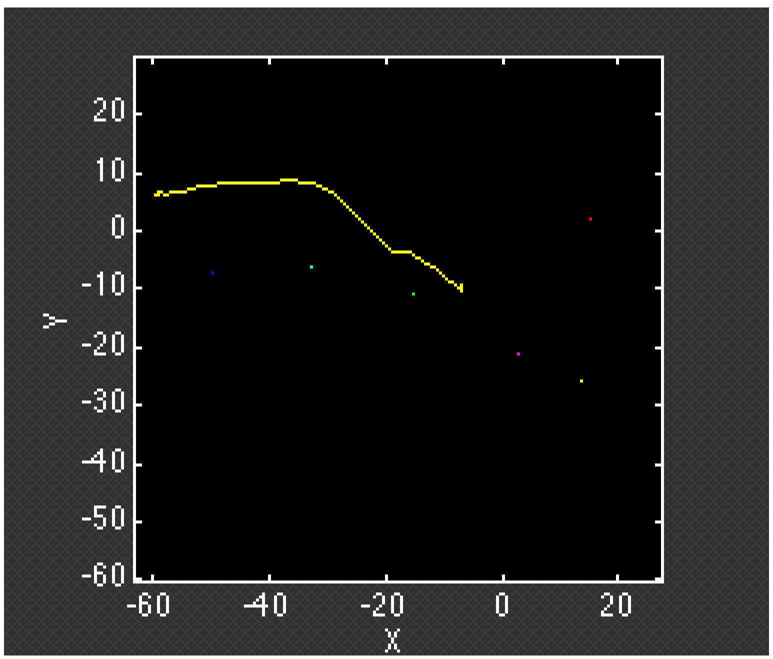
- Movements
  - From x-ray tracings
  - With an Electro-Magnetic Midsagittal Articulometer (EMMA) System
    - Points on the tongue, lips, jaw, (velum)
- Other parameters: air pressures and flows, muscle activity …

EMMA Data Collection

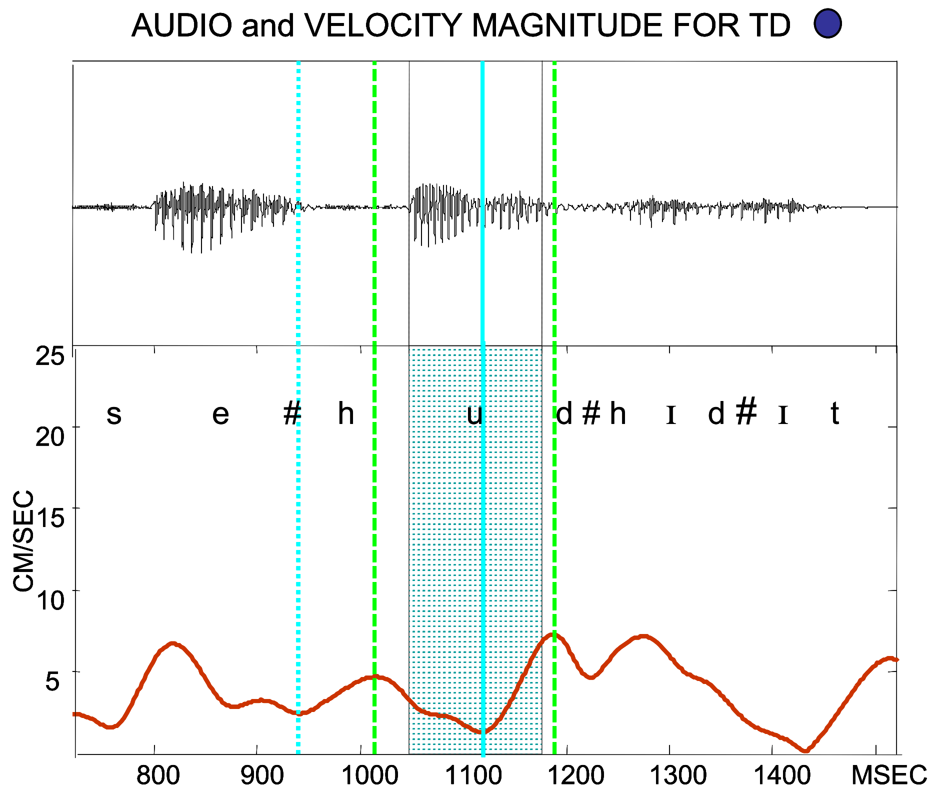


- Transducer coils are placed on subject’s articulators
- Subject reads text from an LCD screen
- Movement and audio signals are digitized and displayed in real time
- Signals are processed and data are extracted and analyzed





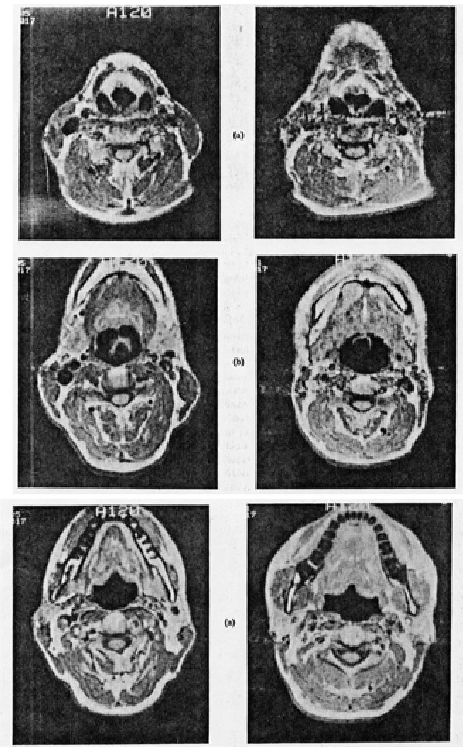
Analysis of EMMA data



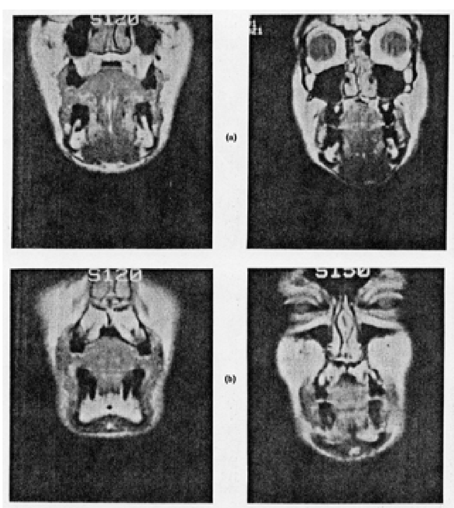
- Algorithmic data extraction at time of minimum in absolute velocity during the vowel:
  - Vowel formants
  - Articulatory positions (x, y)

3-Dimensional Area Function Data

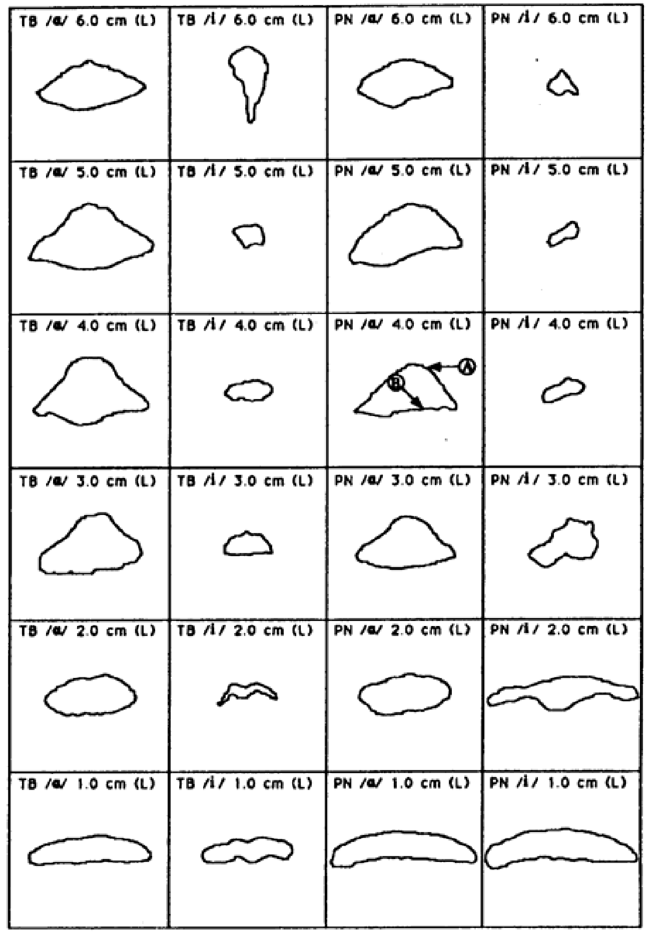
/i/ - 2 speakers



Transverse (horizontal)

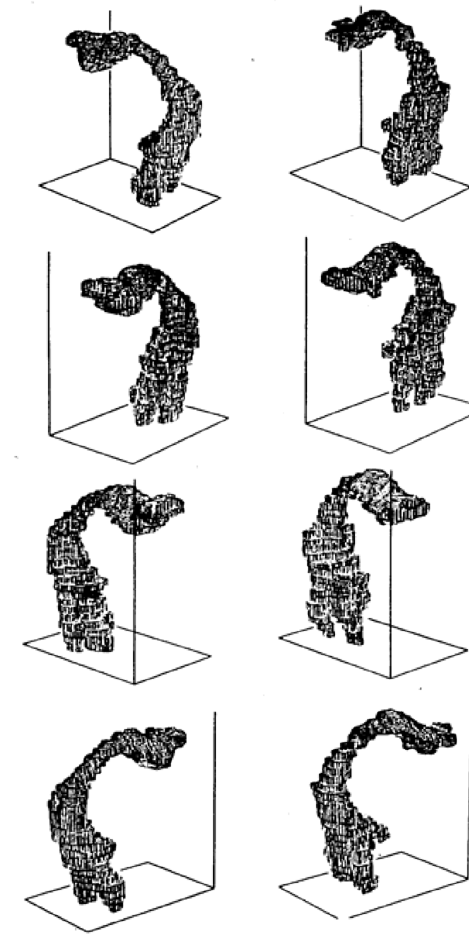


Coronal (vertical)



/ɑ/, /i/ - 2 speakers

/u/ - 2 speakers



- MR images of sustained vowels (Baer et al., JASA 90: 799-828)
  - Area functions are more complicated than they look in 2 dimensions
  - There are lateral asymmetries, but 2-D midsagittal (midline) movement data provide useful information

Reprinted with permission from

Baer, T., J. C. Gore, L. C. Gracco, and P. W. Nye. "Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels." *J Acoust Soc Am* 90 (1991): 799.

Outline

- Introduction
- Measuring speech production
- What are the “controlled variables” for segmental (phonemic) speech movements?
  - Possible controlled variables
  - Modeling to make the problem approachable: DIVA
  - A schematic view of speech movements
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

"What are the controlled variables?"

- The question has theoretical and practical implications
  - What are the *fundamental motor programming units* and most appropriate *elements for phonological/phonetic theory*?
  - What domains should be the main focus of research for diagnosis and treatment of speech disorders?

Figure removed due to copyright reasons.
Please see:
Steven, K. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1998, p. 248. ISBN: 026219404X.

"The yacht was a heavy one"

- Objective of Speaker:
  - To produce sounds strings with **acoustic patterns** that result in **intelligible** patterns of **auditory sensations** in the listener
- Acoustic/auditory cues depend on type of sound segment :
  - Vowels and glides: Time varying patterns of formant frequencies
  - Consonants: Noise bursts, Silent intervals, Aspiration and frication noises, Rapid formant transitions

Possible Motor Control Variables

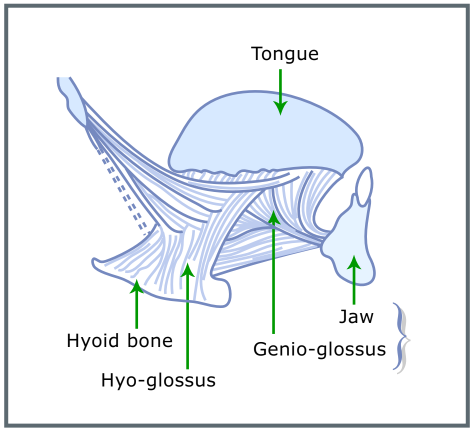
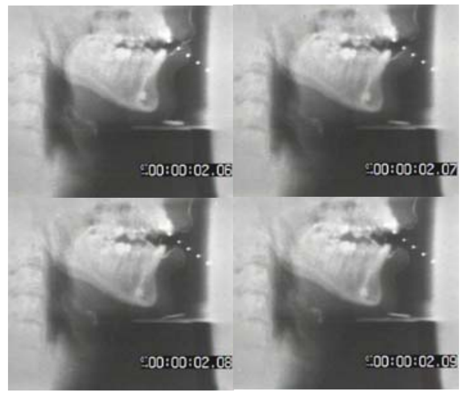


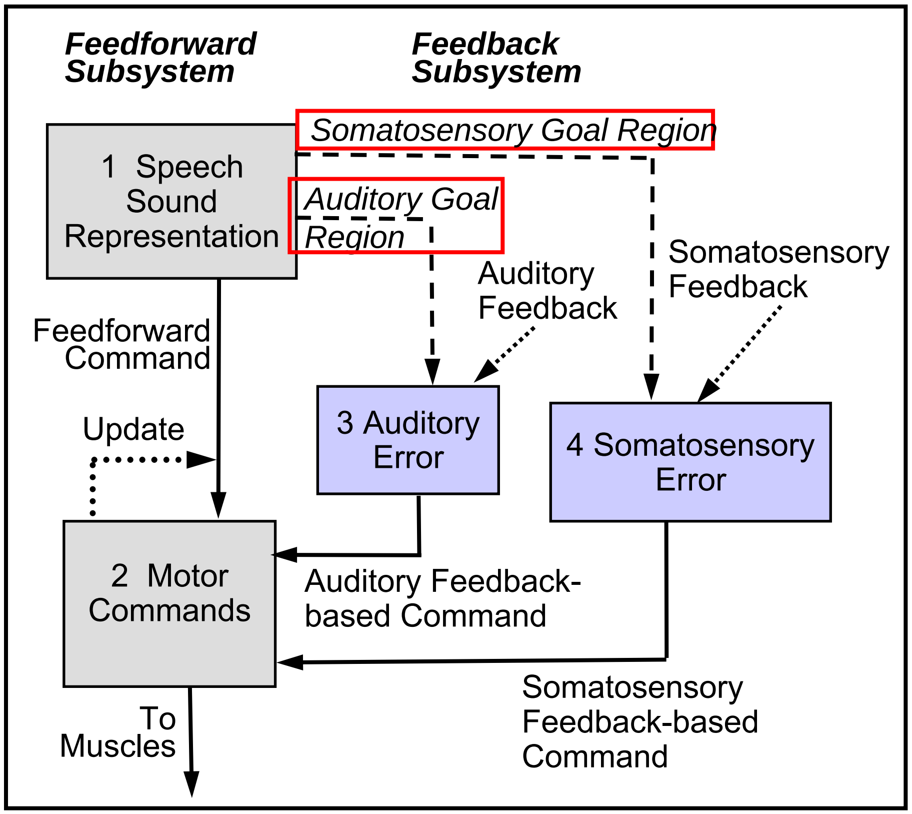
Figure by MIT OCW.



- Auditory characteristics of speech sounds are determined by:
    1. Levels of muscle tension
    2. Changing muscle lengths and movements of structures
    3. The vocal-tract shape (area function)
    4. Aerodynamic events and aeromechanical interactions
    5. The acoustic properties of the radiated sound
- Hypothetically, motor control variables could consist of feedback about any combination of the above parameters

Modeling to make the problem approachable: DIVA

"Directions Into Velocities of Articulators" (Guenther and Colleagues – Next lecture)



- A neuro-computational model of relations among cortical activity, motor output, sensory consequences
- **Phonemic Goals**: Projections (mappings) from premotor to sensory cortex that encode *expected sensory consequences* of produced speech sounds
  - Correspond to *regions* in multidimensional auditory-temporal and somatosensory-temporal spaces
- Roles of feedforward and feedback subsystems will be discussed later.

A Schematic View of Speech Movements

- Planned and actual acoustic trajectories illustrate:
  - Auditory/acoustic goal *regions*
  - Economy of effort (Lindblom)
  - Coarticulation
  - Motor equivalence
  - Biomechanical saturation (quantal) effects
- When controlling an articulatory speech synthesizer, DIVA, accounts for the first four and
  - Aspects of acquisition
  - Responses to perturbations

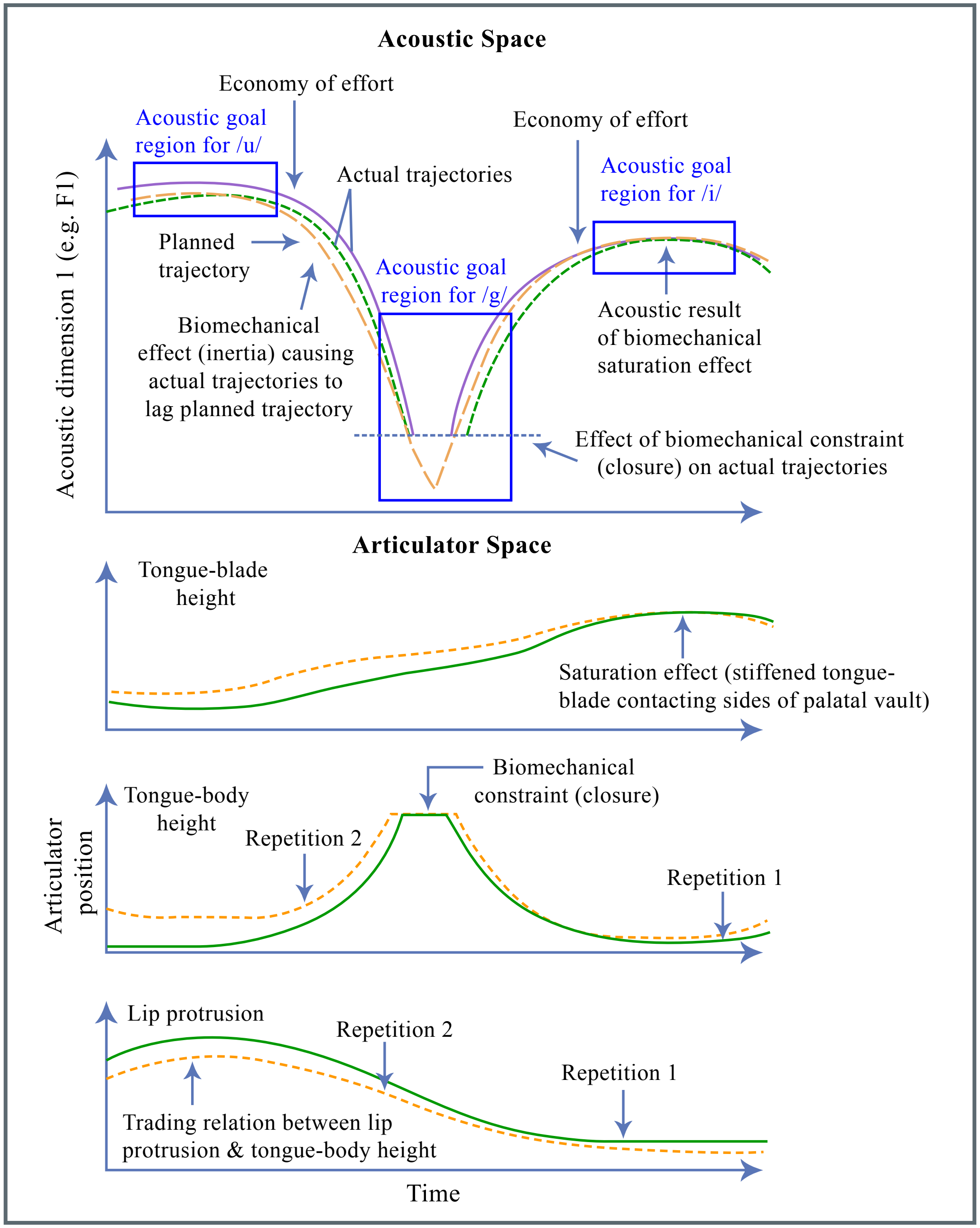


Figure by MIT OCW.

Outline

- Introduction
- Measuring speech production
- What are the “controlled variables” for segmental speech movements?
- Segmental motor programming goals
  - Anatomical and acoustic constraints: Quantal effects
  - Individual differences – anatomy
  - Motor equivalence: A strategy to stabilize acoustic goals
  - Clarity vs. economy of effort
  - Relations between production and perception
    - Vowels
    - Sibilants
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

Anatomical and Acoustic Constraints on Articulatory Goals

- Properties of speakers' production and perception mechanisms help to define goals for speech sounds that are used in speech motor planning
- Some of these properties are characterized by *quantal effects* (Stevens), which can also be called "saturation effects"

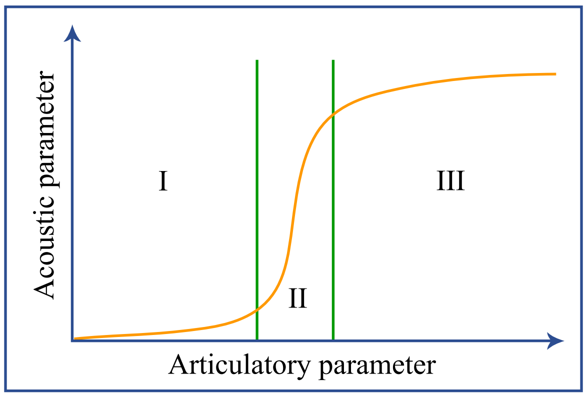


Figure by MIT OCW.

- Schematic example: A continuous change in an articulatory parameter produces two regions of acoustic stability, separated by a rapid transition
- Hypothesis: some goals are auditory and can be characterized in terms of acoustic parameters:formant frequencies, relative sound level, etc.

- Languages "prefer" such stable regions
- The use of those regions by individual speakers helps to produce □ relatively robust acoustic cues with imprecise motor commands □

Goals for the vowel /i/ - A. An acoustic saturation (quantal) effect for constriction location (Stevens, 1989)

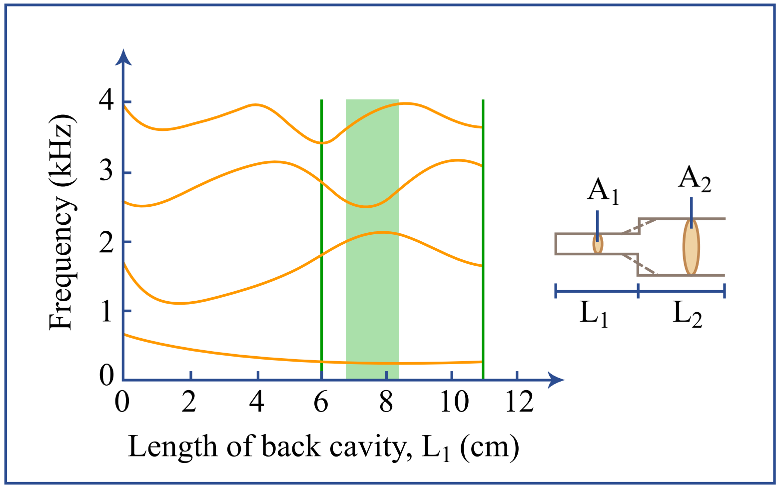
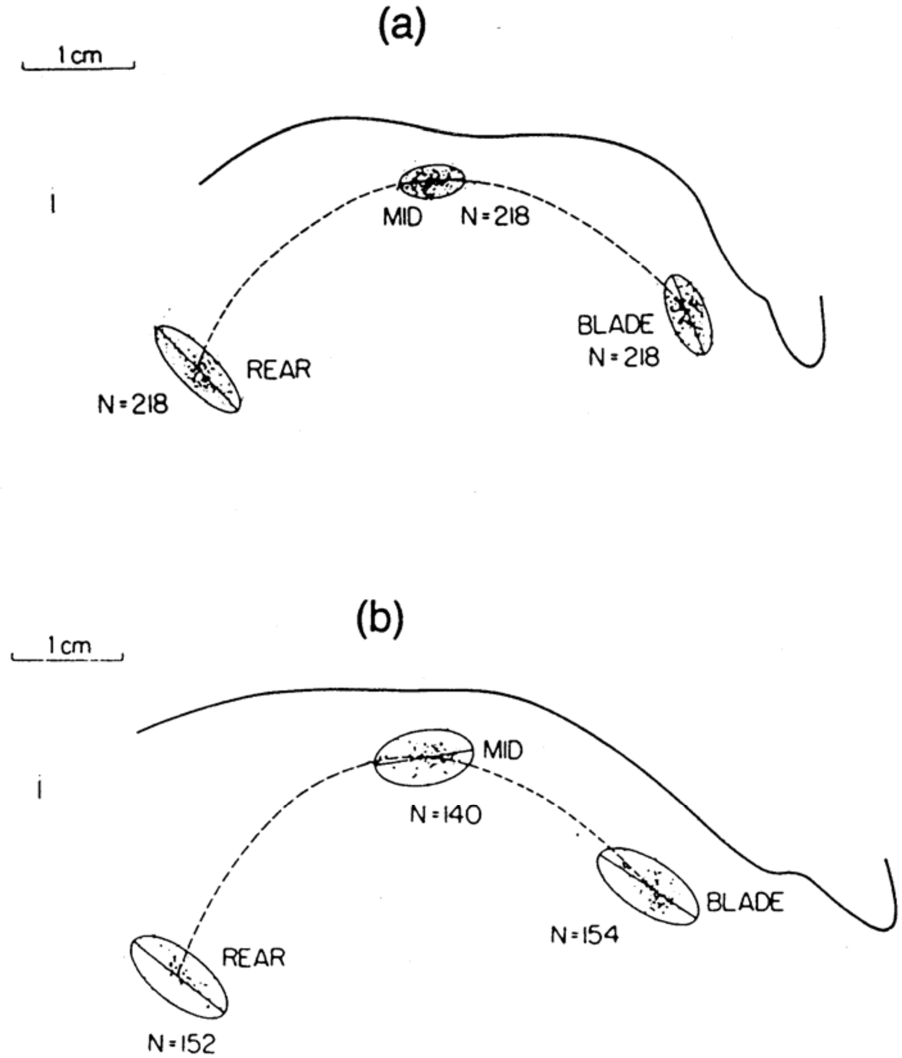


Figure by MIT OCW.

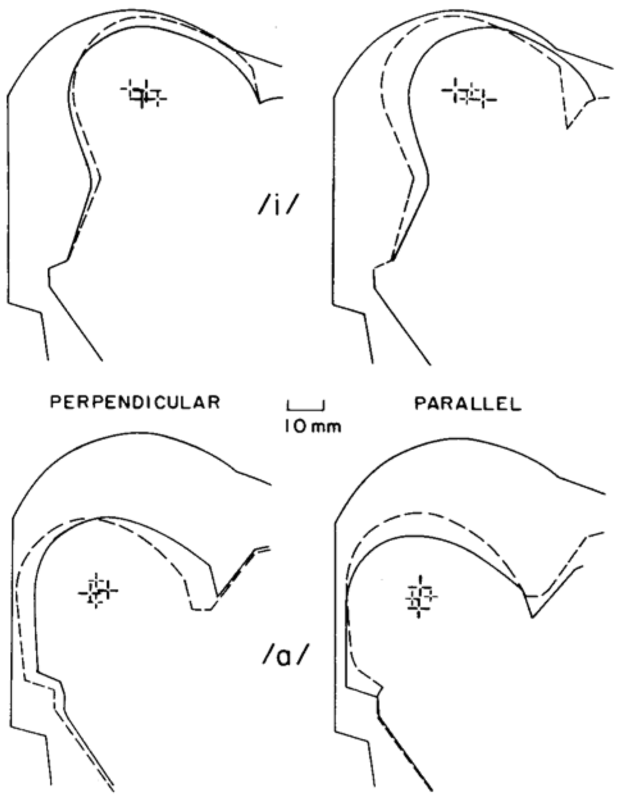


There is a range (in green) of back cavity lengths over which F1-F3 are relatively stable.

Many repetitions of /i/ in two subjects show a corresponding variation of constriction location.

However, as reflected in the articulatory data, the formants of /i/ are **sensitive to variation in constriction degree**.

Quantal and non-quantal articulatory-to-acoustic relations for /i/ and /ɑ/



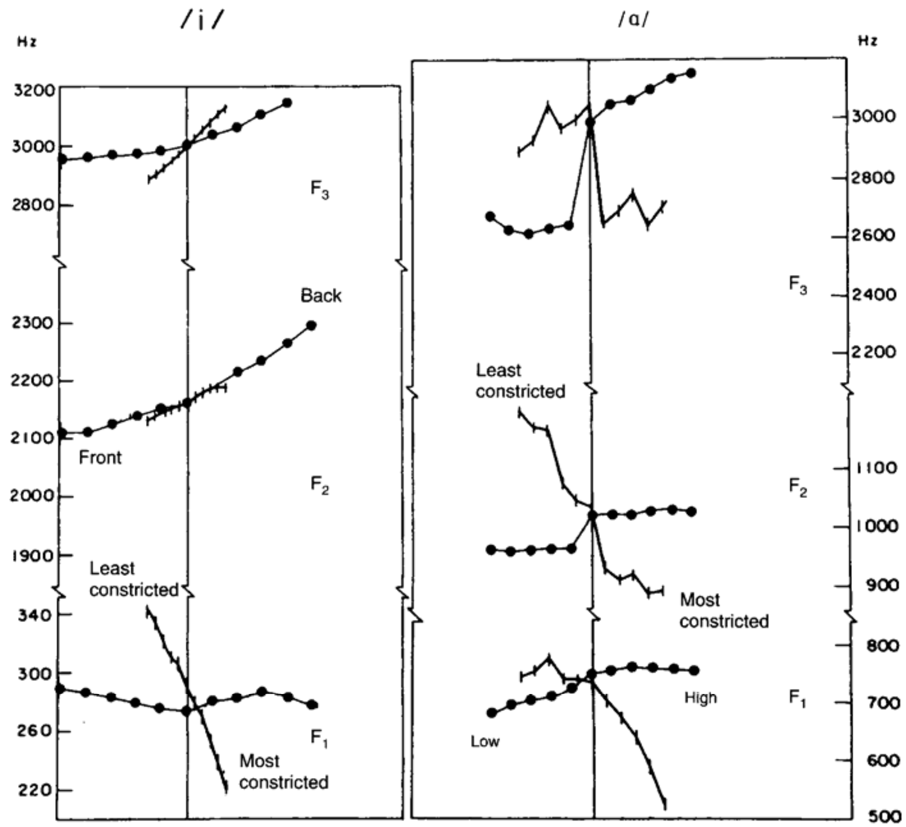
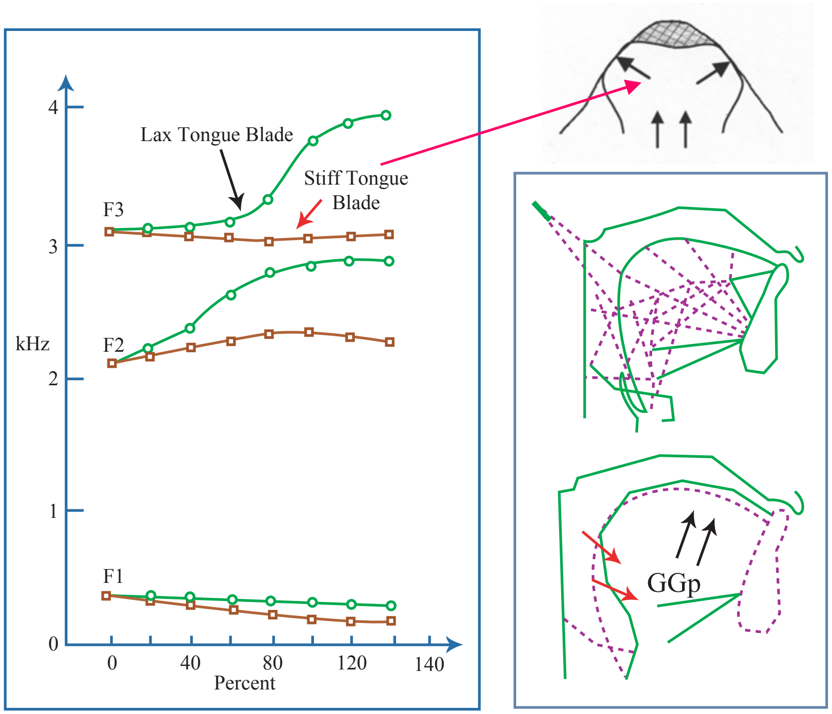


Figure 6. Formant values that result from manipulating the location and degree of constriction of an articulatory synthesizer for the vowel /i/ on the left, and /ɑ/ on the right. The filled circles represent formant values at 1 mm increments of front-to-back tongue-body movement for /i/ and low-to-high tongue-body movement for a /ɑ/, and the tick marks show formant values that represent formant values at 1 mm increments of tongue-body movement from the least constricted to the most constricted configuration for each vowel.

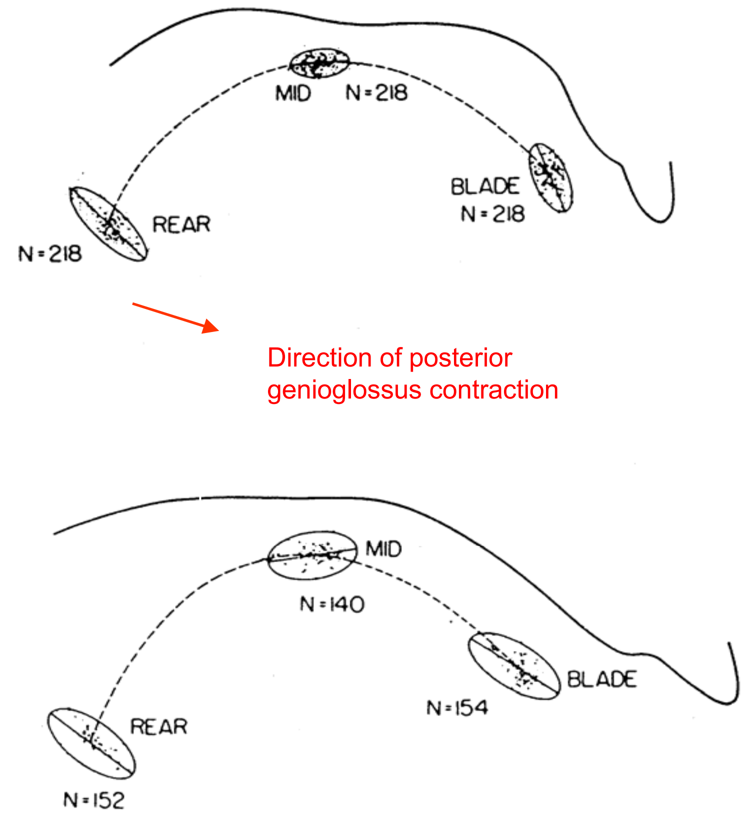
Reprinted with permission from:

Perkell, J. S., and W. L. Nelson. "Variability in production of the vowels /i/ and /a/." *J Acoust Soc Am* 77 (1985): 1889-1895.

A biomechanical saturation effect for constriction degree for /i/



Figures by MIT OCW.



- Constriction degree and resulting formants can be *stabilized*
  - Stiffening the tongue blade (with intrinsic muscles)
  - Pressing the stiffened tongue blade against the sides of the hard palate through contraction of the posterior genioglossus (GGp) muscles

- Constriction area (shaded) varies little, even with variation in GGp contraction (from a 3D tongue model by Fujimura & Kakita)

Tongue Contour Differences Among Four Speakers

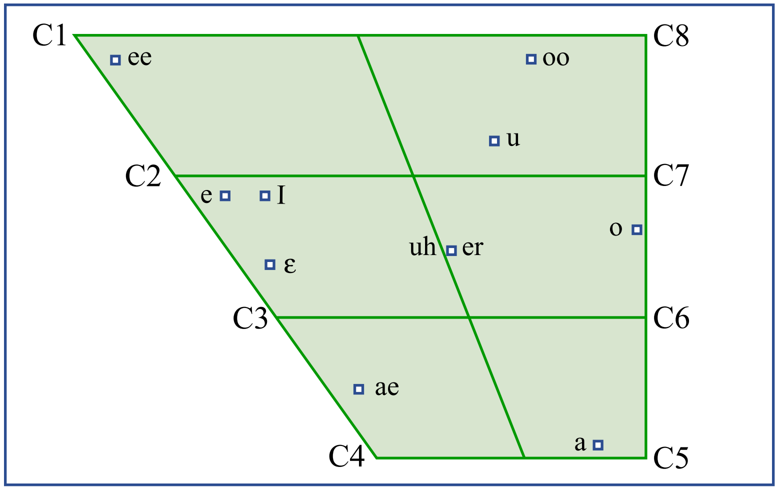


Figure by MIT OCW.

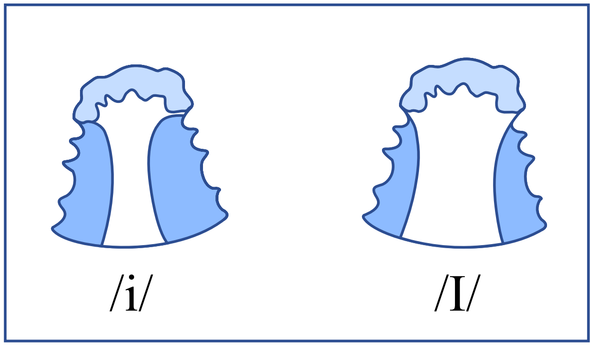
Figures removed due to copyright reasons.

- Note the tongue contour differences among the four speakers.
- An “auditory-motor theory of speech production” (Ladefoged, et al., 1972)

Effects of different palate shapes on vowel articulations

- Palatal shapes differ among individuals
- Palatal depth can influence
  - Spatial differences in vowel targets

Figures removed due to copyright reasons.
Please see:
Perkell, J. S. "On the nature of distinctive features: Implications of a preliminary vowel production study." *Frontiers of Speech Communication Research*. Edited by B. Lindblom and S. Öhman. London, UK: Academic Press, 1979, p. 372 (right), 373 (left).



Figures by MIT OCW.

Production of /u/

- Contractions of the styloglossus and posterior genioglossus
- Note: place of constriction & variation in constriction location

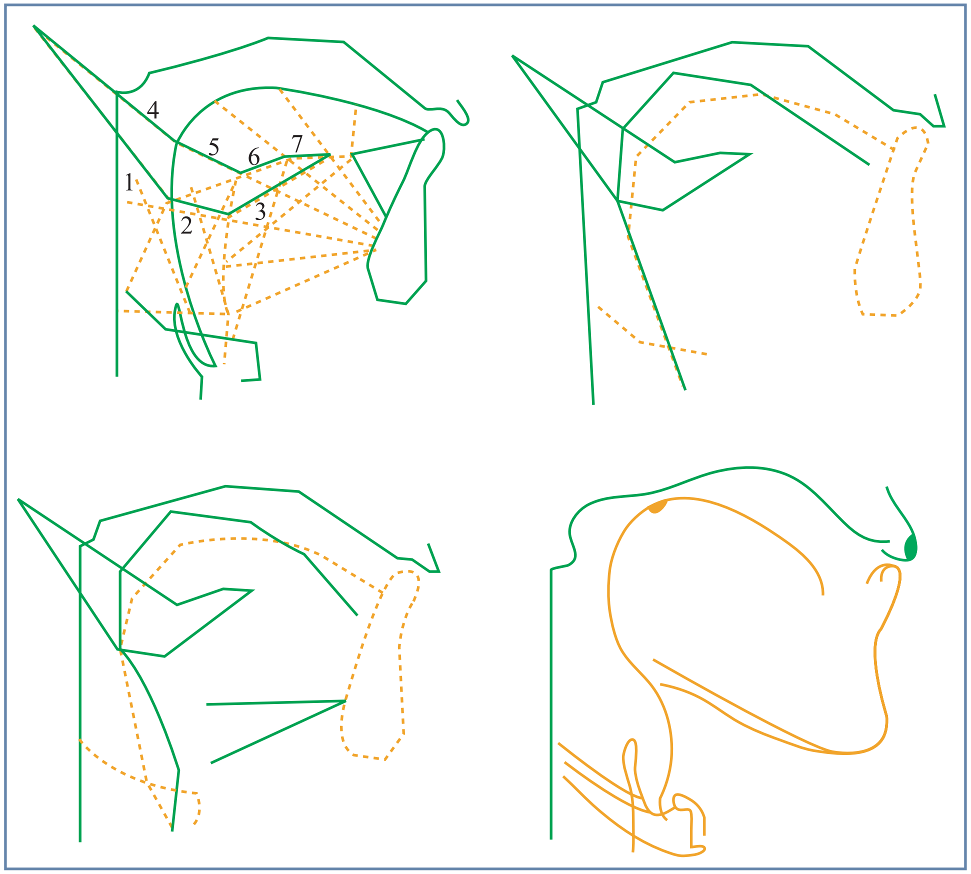
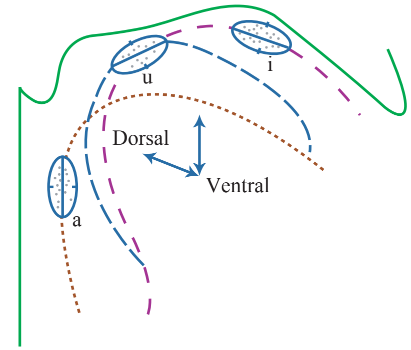
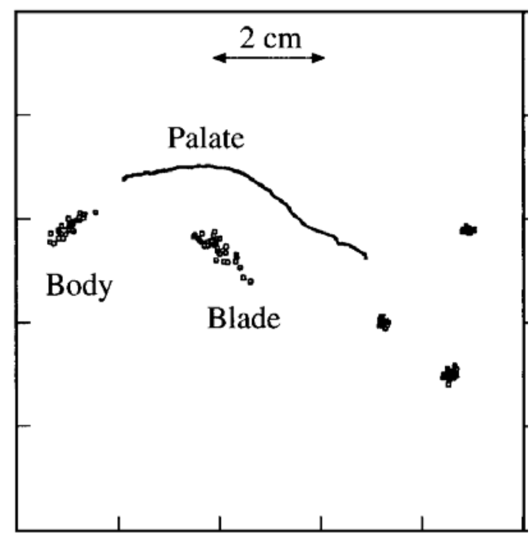


Figure by MIT OCW.

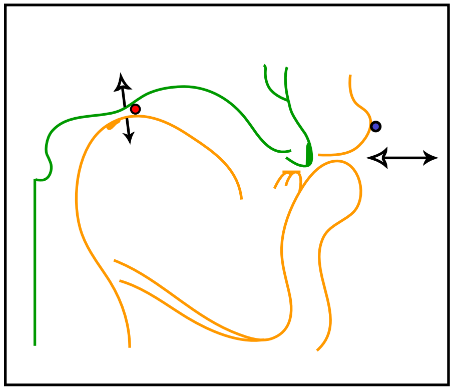


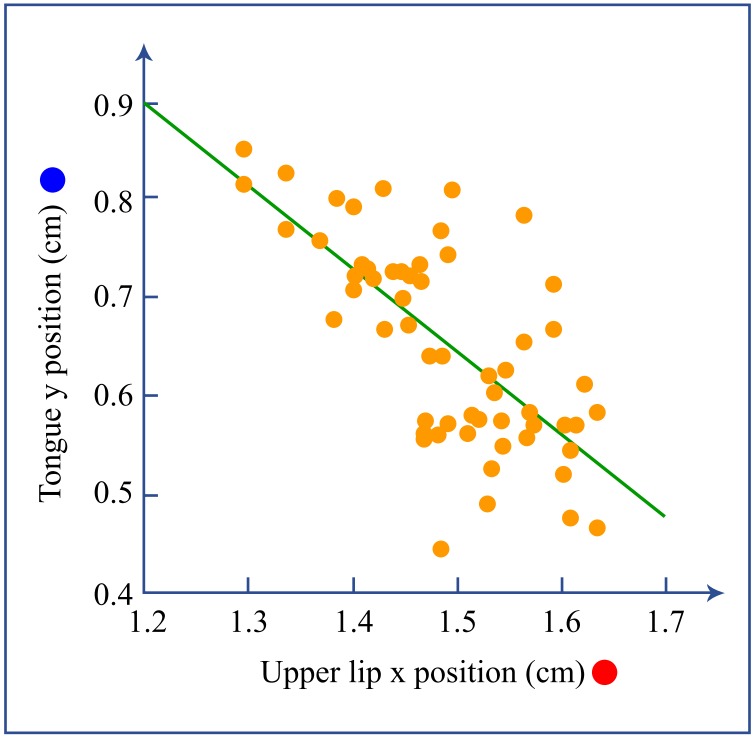
A schematic illustration of articulatory data for multiple repetitions of the vowels /a/ (tongue surface illustrated by······), /i/ (– –) and /u/ (— —), showing elliptical distribution of positioning of points on the tongue surface, with the long axes of the ellipses oriented parallel to the vocal-tract midline.

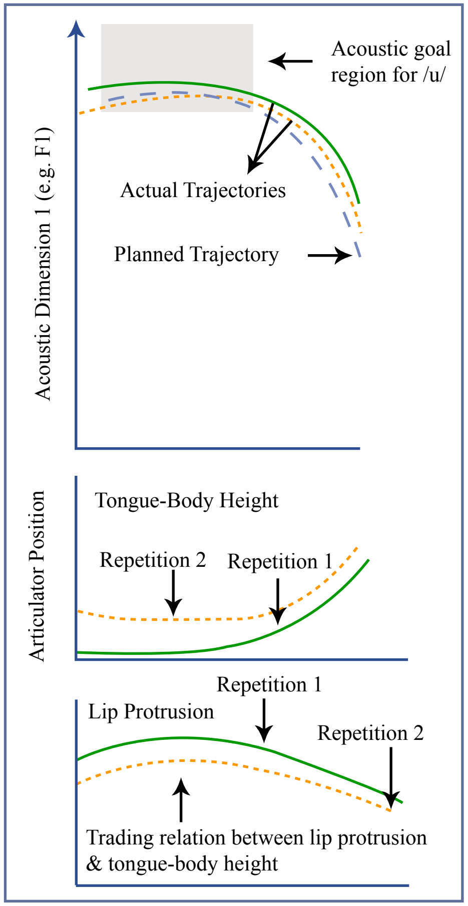
Figure by MIT OCW.



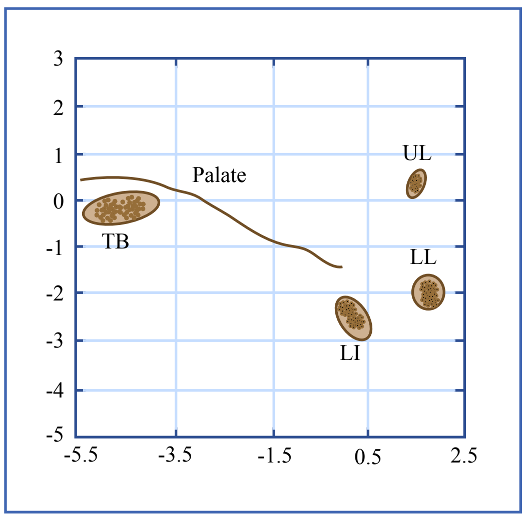
Stabilizing the sound output for the vowel /u/: Motor Equivalence







- Hypothesis: negative correlation between tongue-body raising and lip protrusion in multiple repetitions of the vowel



- Hypothesis is supported in a number of subjects
- The goal for the articulatory movements for /u/ is in an acoustic/auditory frame of reference, not a spatial one
- Strategy: Stay just within the acoustic goal region

Palatal Depth and Motor Equivalence

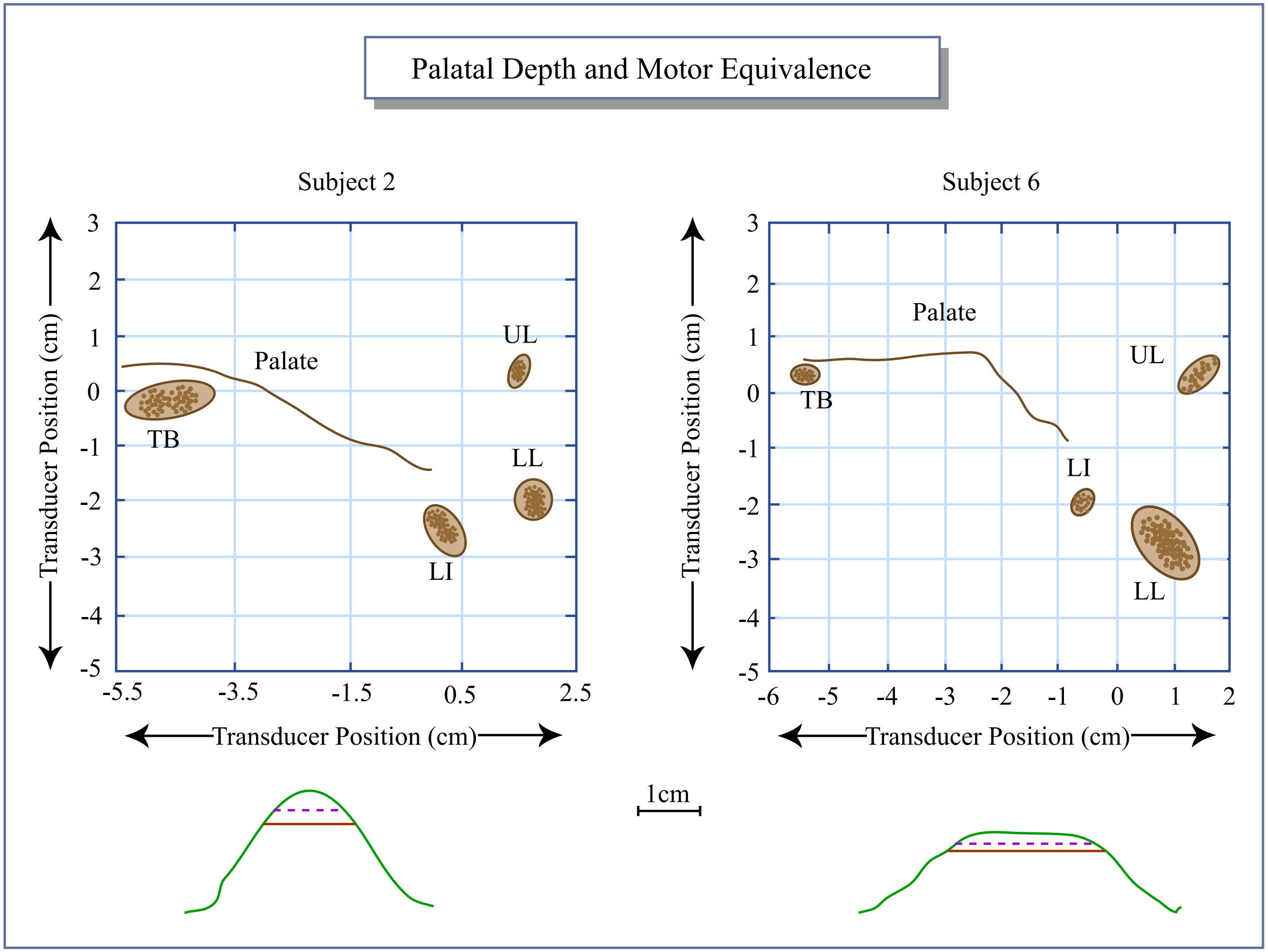
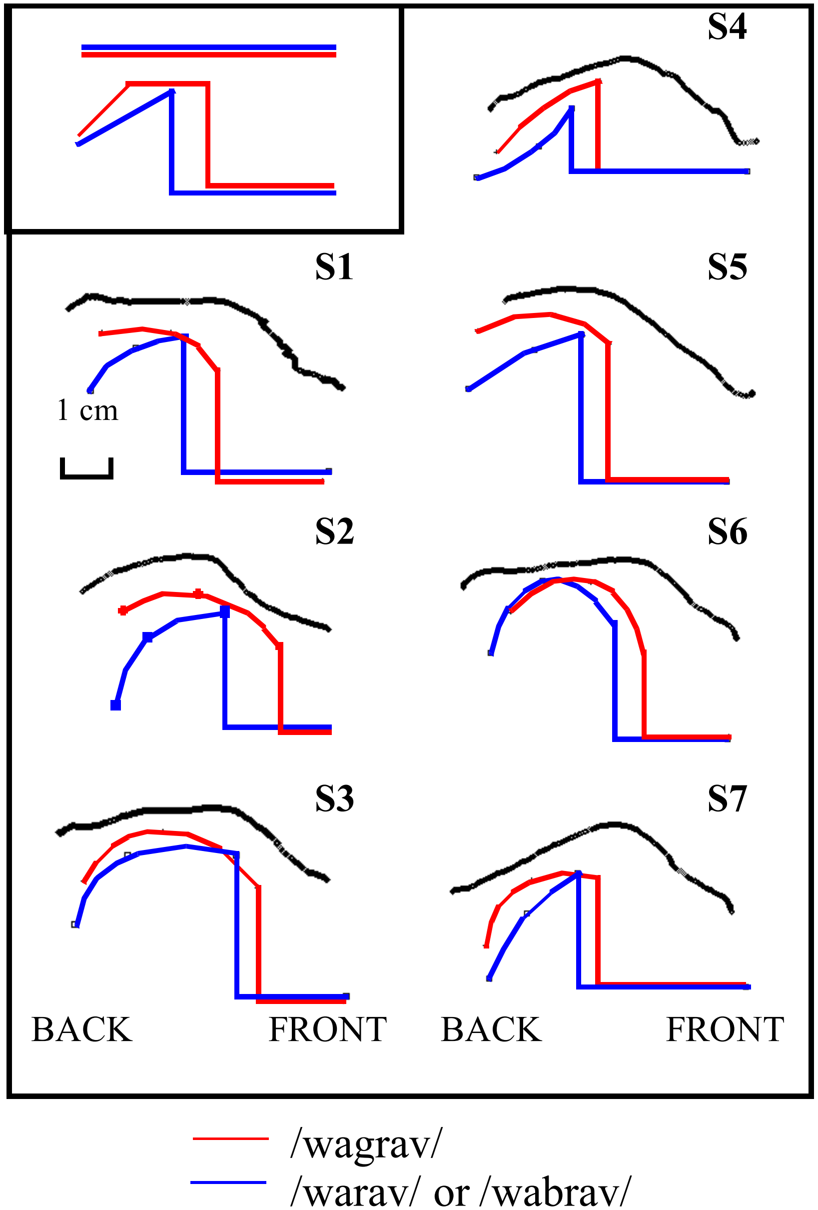


Figure by MIT OCW.

- Palatal depth can also influence
  - Variability in movement toward vowel targets



Reprinted with permission from:
Guenther, F., C. Espy-Wilson, S. Boyce, M. Matthies, M. Zandipour, and J. Perkell. "Articulatory tradeoffs reduce acoustic variability during American English /r/ production." *J Acoust Soc Am* 105 (1999): 2854-2865.

Motor Equivalence for /r/

- Speakers use similar articulatory trading relations when producing /r/ in different phonetic contexts (Guenther, Espy-Wilson, Boyce, Matthies, Perkell, and Zandipour, 1999, JASA)
- Acoustic effect of the longer front cavity of the blue outlines is compensated by the effect of the longer and narrower constriction of the red outlines (e.g., Stevens, 1998).
- F3 variability is greatly decreased by these articulatory trading relations.
- Conclusion: The movement goal for /r/ is a low value of F3 – an *auditory/acoustic goal*

Clarity vs. Economy of Effort: Another principle (continuous, as opposed to. quantal) that influences vowel categories (Lindblom, 1971)

Figures removed due to copyright reasons.

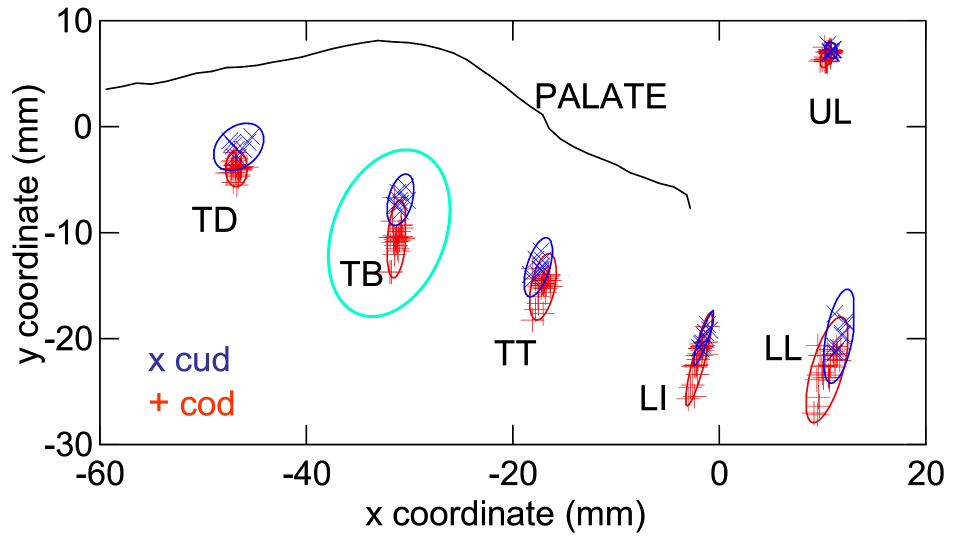
- Used an articulatory synthesizer and heuristics to estimate the location of vowels in F1, F2 space, based on
  - A compromise between "perceptual differentiation" and "articulatory ease" and
  - The number of vowels in the language
- Approximated vowel distributions for languages containing up to about 7 vowels
- Later discussed in terms of a tradeoff between *clarity* and *economy of effort,* i.e., a **relation between production and perception**

Relations between Production and Perception

- Close linkage between production and perception:
  - Speech acquisition, with and without hearing
  - Speech of Cochlear Implant users
  - Second-language learning (e.g., Bradlow et al.)
  - Focused studies of production & perception (e.g., Newman)
  - Mirror neurons – a more general action-perception link (e.g., Fadiga et al.)
- Hypothesis:
  - Speakers who discriminate well between vowel sounds with subtle acoustic differences will produce more clear-cut sound contrasts
  - Speakers who are less able to discriminate the same sound stimuli will produce less clear-cut contrasts

Production Experiment

- Data Collection
  - Subjects: 19 young-adu speakers of American English
  - For each subject:
    - Recorded articulatory movements and acoustic signal
    - Subject pronounced "Say___ hid it."; ___ = *cod, cud, who'd* or *hood*
    - *Clear, Normal* and *Fast* conditions

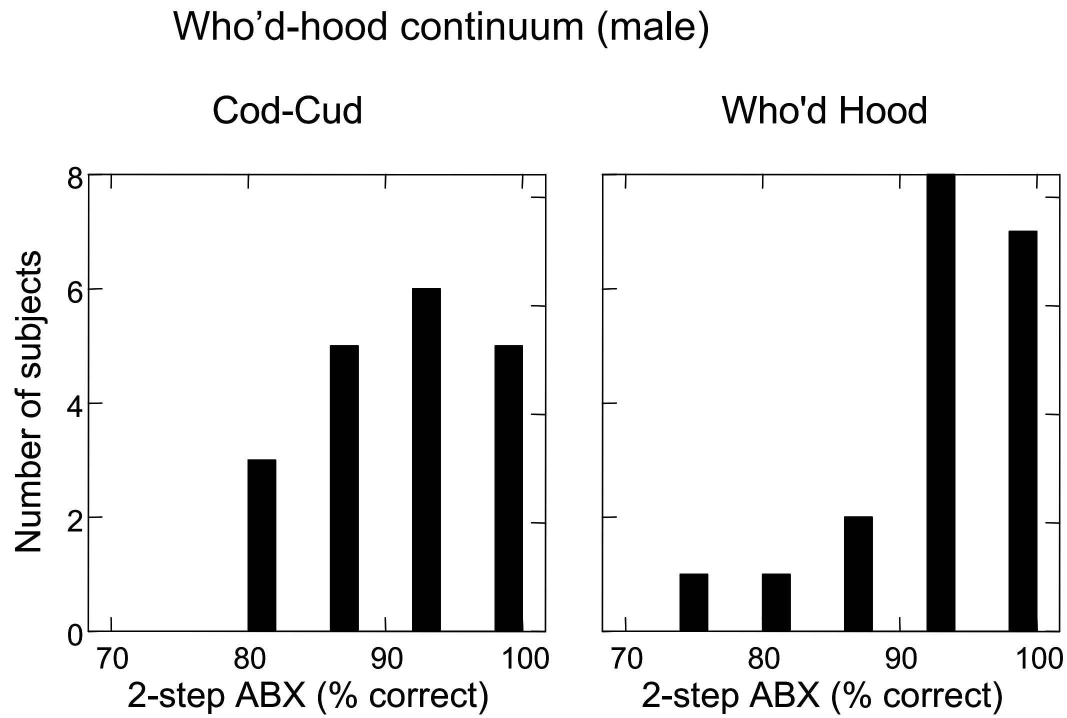


Reprinted with permission from
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, E. Stockmann, M. Tiede, M. Zandipour. "The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts." *J Acoust Soc Am* 116 (2004): 2338-44.

- Analysis
  - Calculated **contrast distance** for each vowel pair:
    - *Articulatory* (TB) *contrast distance*: distance in mm between the centroids of the *cod* and *cud* TB distributions.
    - *Acoustic contrast distance*: distance in Hz between centroids of F1, F2 distributions for *cod*, *cud*

Perception Experiment

- Methods
  - Synthesized natural-sounding stimuli in 7-step continua – for *cod-cud*, *who'd-hood*
  - Each subject: Labeling and discrimination (ABX) tasks



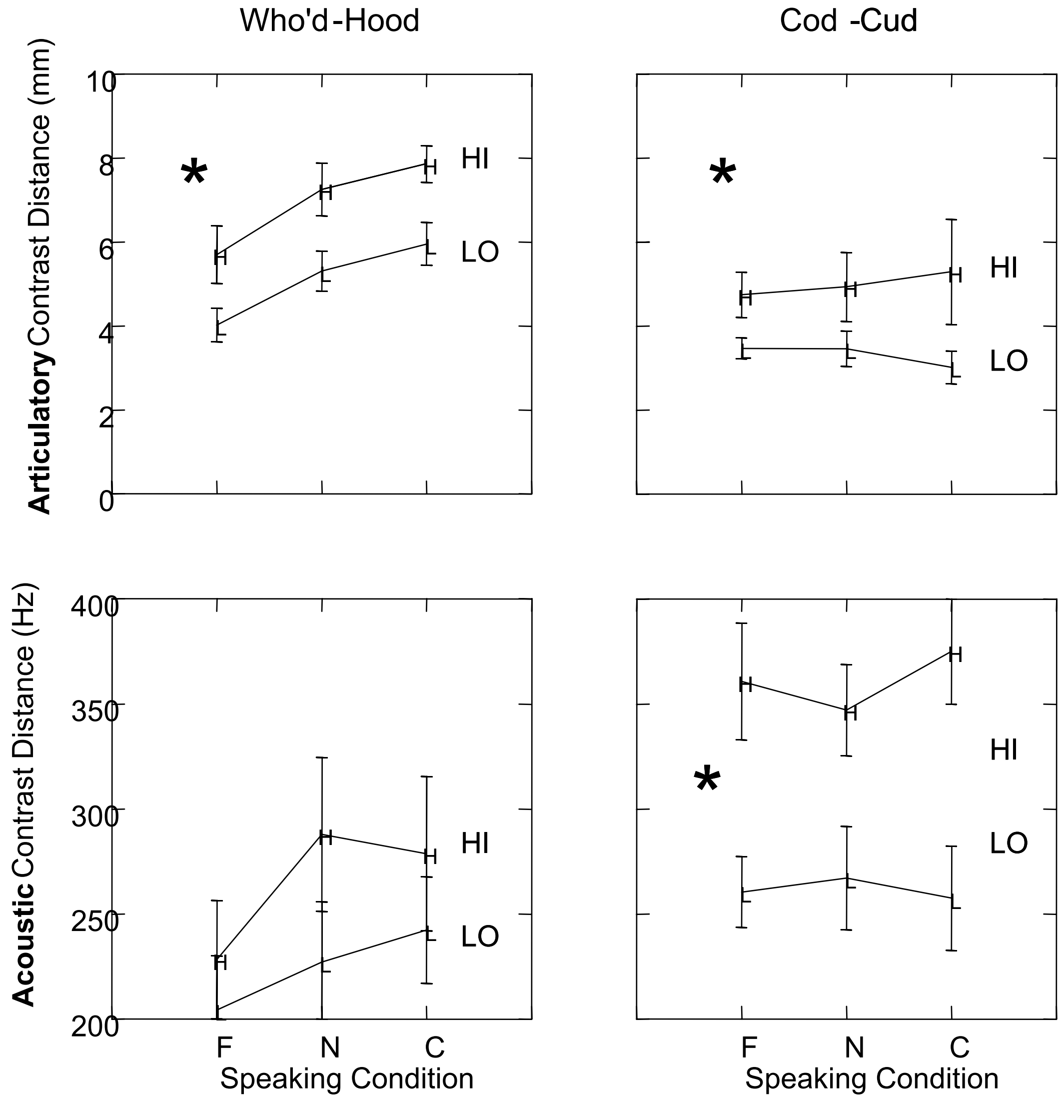
Reprinted with permission from:
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, E. Stockmann, M. Tiede, M. Zandipour. "The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts." *J Acoust Soc Am* 116 (2004): 2338-44.

- Results: ABX scores (2-step)
  - Ceiling effects: some 100% subjects probably had better discrimination than measured
  - For further analysis divide subjects into two groups:
    - **HI** discriminators - at 100% (above the median)
    - **LO** discriminators - (at median and below)

Results & Conclusions

- HI discrimination subjects produced greater contrast distance than LO discrimination subjects (measured in articulation or acoustics)
- The more accurately a speaker discriminates a vowel contrast, the more distinctly the speaker produces the contrast

* Difference between HI and LO groups is significant at $p < .001$



Reprinted with permission from:
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, E. Stockmann, M. Tiede, M. Zandipour. "The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts." *J Acoust Soc Am* 116 (2004): 2338-44.

A Possible Explanation

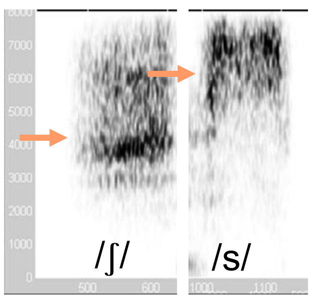
- It is advantageous to be as intelligible as possible
- Children will acquire goal regions that are as distinct as possible
  - Speakers who can perceive fine acoustic details learn *auditory goal regions* that are *smaller* and *spaced further apart* than speakers with less acute perception, because
  - The speakers with more acute perception are more likely to reject poorly produced tokens when learning the goal regions

Consonants: A saturation effect for /s/ may help define the /s-ʃ/ contrast

Figures removed due to copyright reasons.
Please see:
Figures from: Perkell, J. S., M. L. Matthies, M. Tiede, H. Lane, M. Zandipour, N. Marrone, E. Stockmann, and F. H. Guenther. "The Distinctness of Speakers' /s/–/ ʃ / Contrast is related to their auditory discrimination and use of an articulatory saturation effect." *J Speech, Language and Hearing Res* 47 (2004): 1259-69.

- Production of /ʃ/ (as in "shed")
  - Relatively long, narrow groove between tongue blade and palate
  - Sublingual space
- Production of /s/ (as in "said")
  - Short narrow groove
  - No sublingual space
- Saturation effect for /s/
  - As tongue moves forward from /ʃ/, sublingual cavity volume decreases
  - When tongue contacts lower alveolar ridge, sublingual cavity is eliminated, resonant frequency of anterior cavity increases abruptly
  - After contact, muscle activity can increase further; output is unchanged

Relations Between Production and Perception of Sibilants



Figures removed due to copyright reasons.
Please see:
Figures from: Perkell, J. S., M. L. Matthies, M. Tiede, H. Lane, M. Zandipour, N. Marrone, E. Stockmann, and F. H. Guenther. ”The Distinctness of Speakers' /s/−/ ʃ / Contrast is related to their auditory discrimination and use of an articulatory saturation effect.” *J Speech, Language and Hearing Res* 47 (2004): 1259-69.

- Hypothesis: The sibilants, /s/ and /ʃ/, have two kinds of sensory goals:
  - Auditory: particular distribution of energy in the noise spectrum
  - Somatosensory: e.g., patterns of contact of the tongue blade with the palate and teeth

- Speakers will vary in their ability to discriminate /s/ from /ʃ/
- Speakers use contact of the tongue tip with the lower alveolar ridge for /s/ to help differentiate /s/ from /ʃ/
  - This will also vary across speakers
- Across speakers, both factors, *ability to discriminate auditorily* between the two sounds and *use of contact* (a possible *somatosensory goal*), will predict the strength of the produced contrast

Methods (with the same 19 subjects as the vowel study)

- Production experiment – each subject:
  - Recorded: acoustic signal, and *contact* of the under side of the tongue tip with the lower alveolar ridge - with a custom-made sensor
  - Subject pronounced, “Say___ hid it.”; ___ = *sod, shod, said* or *shed”*
  - *Clear, Normal* and *Fast* conditions
- Analysis – calculated:
  - *Proportion of time contact was made* during the sibilant interval
  - Spectral median for /s/ and /ʃ/
  - *Acoustic contrast distance*:
    - Difference in spectral median between /s/ and /ʃ/

Figures removed due to copyright reasons.
Please see:
Figures from: Perkell, J. S., M. L. Matthies, M. Tiede, H. Lane, M. Zandipour, N. Marrone, E. Stockmann, and F. H. Guenther. ”The Distinctness of Speakers' /s/–/ ʃ / Contrast is related to their auditory discrimination and use of an articulatory saturation effect.” *J Speech, Language and Hearing Res* 47 (2004): 1259-69.

- Perception experiment - each subject:
  - Labeled and discriminated (ABX) between synthesized stimuli from a seven-step *said* to *shed* continuum

Results

Use of tongue-to-lower-ridge contact

Discrimination

–

Figures removed due to copyright reasons.
Please see:
Figures from: Perkell, J. S., M. L. Matthies, M. Tiede, H. Lane, M. Zandipour, N. Marrone, E. Stockmann, and F. H. Guenther. "The Distinctness of Speakers' /s/–/ ʃ / Contrast is related to their auditory discrimination and use of an articulatory saturation effect." *J Speech, Language and Hearing Res* 47 (2004): 1259-69.

– 12 subjects (left of vertical line) are classified as Strong (S) for use of *contact difference* (c) between /s/ and /ʃ/

– The remaining subjects are classified Weak (W) for use of *contact difference*

– Nine subjects had percent correct = 100; categorized as HI discriminators (right of line)

– 10 subjects had percent correct < 100; categorized as LO discriminators

Produced contrast distance is related to

- Ability to discriminate the contrast

* difference is significant, $p < .01$

- Use of contact difference

- Interactions
  - Speakers with good discrimination and use of contact difference: best contrasts
  - Speakers with one or the other factor: intermediate contrasts
  - Speakers with neither factor: poorest contrasts

Figures removed due to copyright reasons.
Please see:
Figures from: Perkell, J. S., M. L. Matthies, M. Tiede, H. Lane, M. Zandipour, N. Marrone, E. Stockmann, and F. H. Guenther. "The Distinctness of Speakers' /s/–/ ʃ / Contrast is related to their auditory discrimination and use of an articulatory saturation effect." *J Speech, Language and Hearing Res* 47 (2004): 1259-69.

Outline (break time?)

- Introduction
- Measuring speech production
- What are the “controlled variables” for segmental speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
  - An example utterance
  - Movements show context dependence
    - Velar movements
    - Lip rounding for /u/
  - Effects of speaking rate
  - Persistence of inaudible gestures at word boundaries
- Experiments on feedback control
- Summary

Producing Sounds in Sequences

- An example utterance
- * indicates articulations that aren't strongly constrained by communicative needs
- Articulations anticipate upcoming requirements: anticipatory coarticulation
- Coarticulation:
  - asynchronous movements of structures of differing sizes and movement time constants
  - a complicated motor coordination task

	k	æ	m	b	r
Tongue body	Rise to contact roof of mouth to achieve closure & silence	Release contact to generate a noise burst; move down to vowel position	Begin movement toward /r/ position	*	Maintain /r/ position
Tongue blade	*	Maintain contact with floor of mouth to stay out of the way	Begin retroflexion or bunching in anticipation of /r/	*	Maintain retroflexed or bunched configuration
Lips	Begin spreading for the vowel /æ/	Maintain position for vowel, then begin toward closure	Achieve & maintain closure	Maintain closure	Release rapidly & round somewhat
Mandible	Move upward to support tongue movement	Move downward to support tongue movement	Move upward to support lower lip movement	*	Move downward slightly to aid lip release
Soft palate	Maintain closure to contain pressure buildup	Begin downward movement to open velopharyngeal port for /m/	Begin closing movement toward onset of /b/	Reach closure at right instant to begin /b/, (move upward during /b/ to help expand v.t. walls - voicing)	*
Vocal-tract walls	Stiffen to contain air pressure buildup	*	*	Relax, perhaps expand actively to allow continuation of voicing for /b/	*
Vocal-fold position	Abduct maximally, with peak occurring at /k/ release	Adduct to position for voicing	Maintain position	Maintain position	Maintain position
Tension on vocal folds	Begin to raise tension to signal stress on following vowel	Achieve maximum tension for the F0 peak that signals stress	Lower tension to lower F0	Maintain tension	Maintain tension
Respiratory system	Increase subglottal air pressure to obtain a burst release for the /k/	Maintain subglottal air pressure for increased sound level to signal stress	Return to the previous value of subglottal pressure	Maintain pressure	Maintain pressure

Table by MIT OCW.

What happens when sounds are produced in sequences?

- When individuals speak to one another, additional forces are at play
  - Articulatory movements from one sound to another are influenced by dynamical factors: canonical targets are very rarely reached.
  - The speaker knows that the listener can fill in a great deal of missing information, so "reduction" takes place (see Introduction)
  - Speaking style (casual, clear, rapid, etc.) can vary
    - Amount of variation can depend on the situation and the interlocutor (a familiar speaker of the same language?)

Movements Show Context Dependence

- Coarticulation
  - At any moment in time, the current state of the vocal tract reflects the influence of preceding sounds (perservatory coarticulation) and upcoming sounds (anticipatory coarticulation)
  - Such coarticulation is a property of any kind of skilled movement (e.g., tennis, piano playing, etc.)
  - It makes it possible to produce sounds in rapid succession (up to about 15/sec), with smooth, economical movements of slowly-□ moving structures.

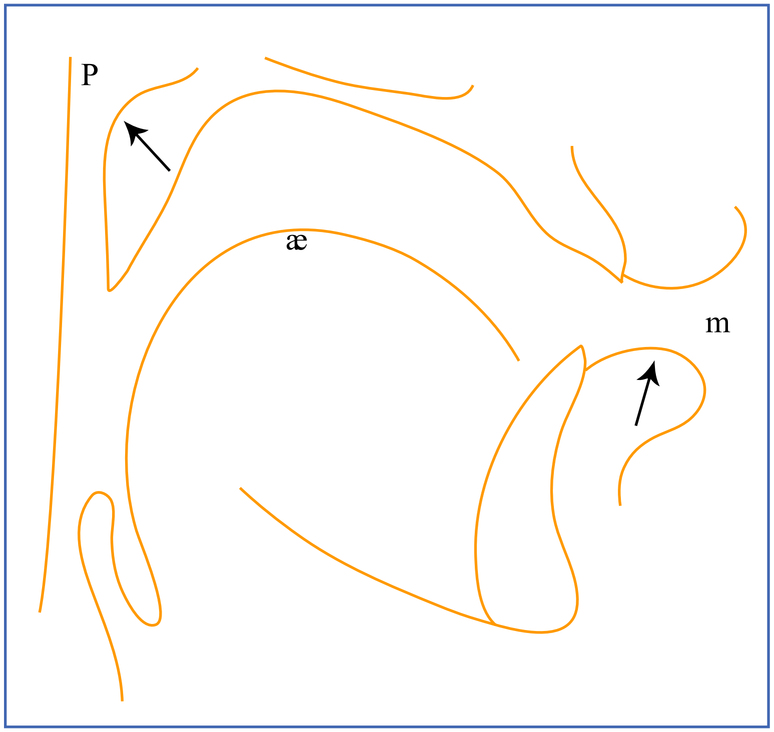
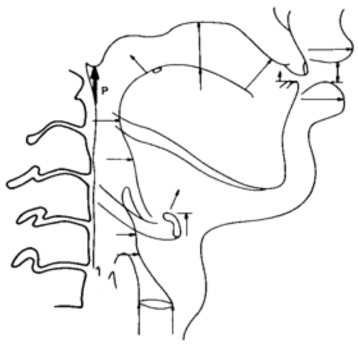
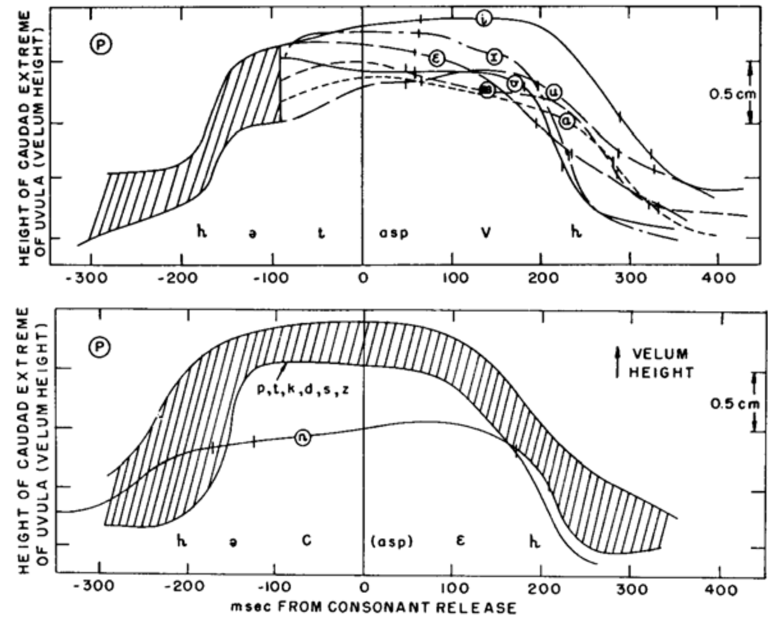


Figure by MIT OCW.

- During the /æ/ in “camping” (Kent)

Effects of Coarticulation and Speaking Rate on velar movements





From: Perkell, Joseph S. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Research Monograph No. 53. Cambridge, MA: MIT Press. 1969(c). Used with permission.

- The velum has to be raised to contain the air pressure increase of obstruent consonants
  - Its height during the /t/ is context (vowel) dependent - coarticulation
  - In the context of a nasal consonant, vowels in American English can be nasalized due to coarticulation
  - This is possible because vowel nasalization isn't contrastive in American English
- The velum (like most other vocal-tract structures) is slowly-moving
  - At higher speaking rates, its movements become attenuated

Figure removed due to copyright reasons.

Coarticulation of lip rounding for /u/

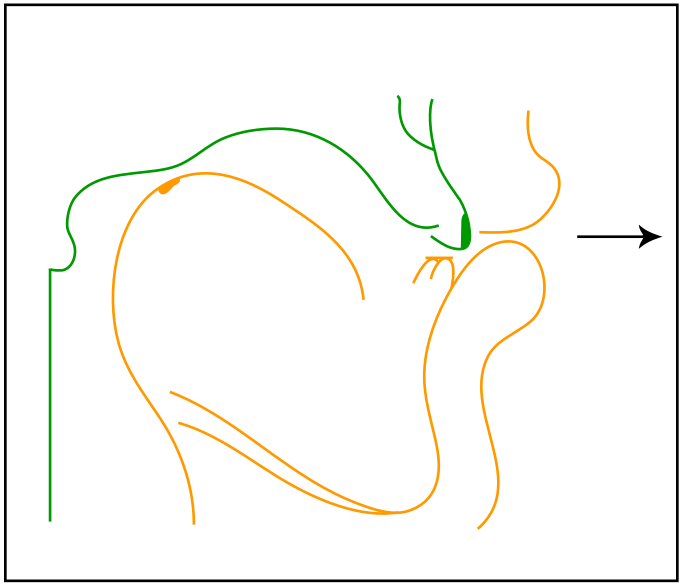
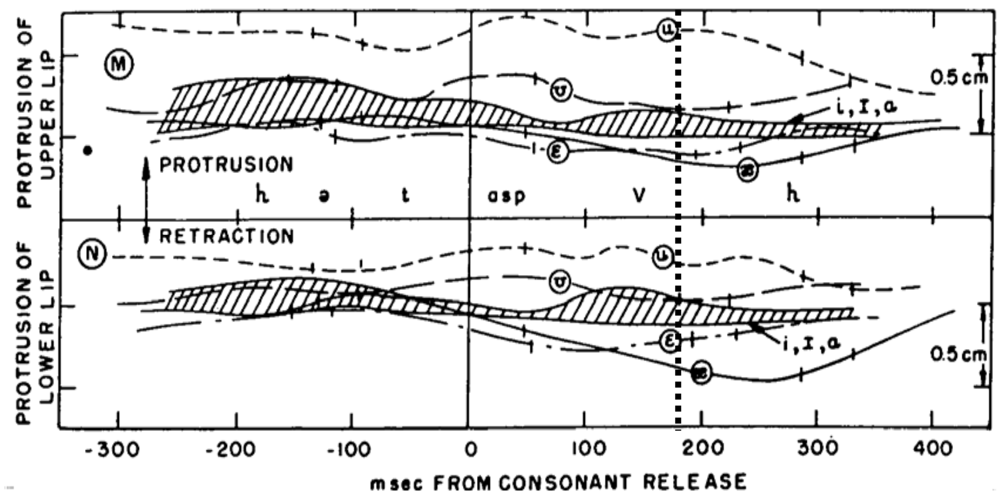


Figure by MIT OCW.



From: Perkell, Joseph S. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Research Monograph No. 53. Cambridge, MA: MIT Press. 1969(c). Used with permission.

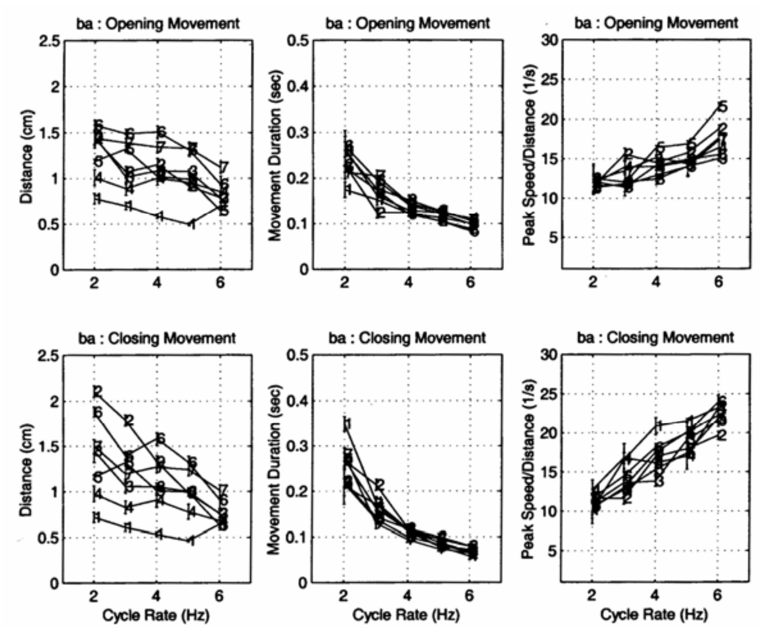
- Lip rounding in production of the vowel /u/ in /həˈtu/
  - The first three sounds in the utterance are neutral with respect to lip rounding
  - The lips are fully protruded before the utterance begins
  - Coarticulation takes place whenever it doesn't interfere with transmission of the message
  - It crosses syllable and word boundaries
  - Movements of different structures are asynchronous

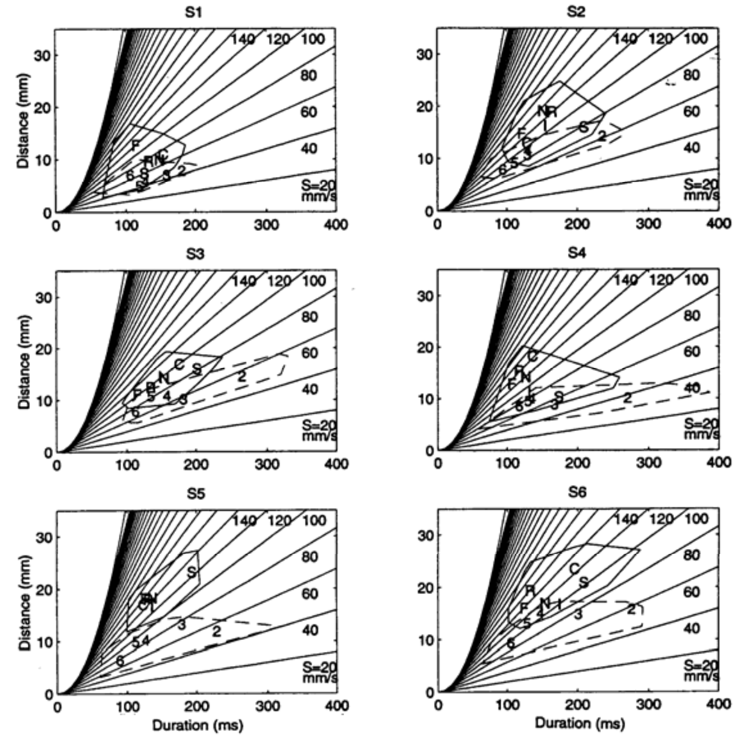
Coarticulation and Acoustic Effects (Gay, 1973, J. Phonetics 2:255-266)

Figures removed due to copyright reasons.

- Cineradiographic measurements – anticipatory and perseveratory coarticulation
  - Tongue movement from /i/ to /ɑ/ can start during the /p/ because it can't be heard and it isn't constrained physically
  - The consonants have effects only on the pellet position for the /ɑ/ (not /i/ or /u/).
  - The pellet is at an acoustically critical constriction in the vocal tract for /i/ and /u/, but not for /ɑ/.
    - Note the vertical variation for /ɑ/ (possible for constriction location - QNS).

Effects of Speaking Rate



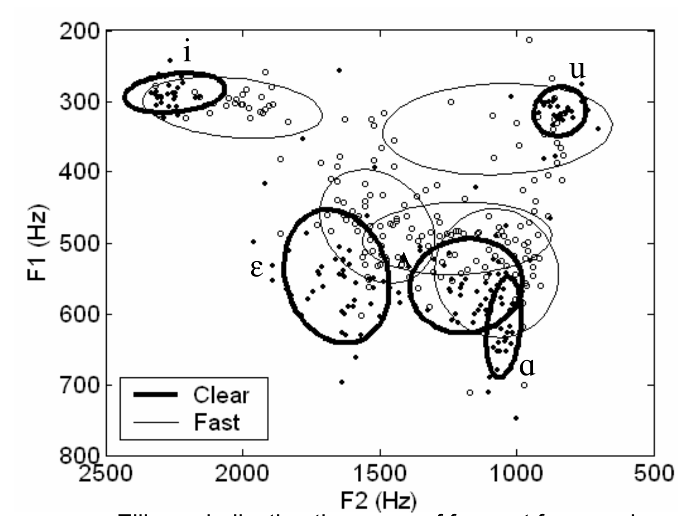


Reprinted with permission from:

Perkell, J., and M. Zandipour. "Economy of effort in different speaking conditions. II. Kinematic performance spaces for cyclical and speech movements." *J Acoust Soc Am* 112 (2002): 1642-51.

Duration (ms)

- Cyclical movements:
  - Higher rates show decreased movement durations, distances, increased speed (a measure of effort)
- Speech vs. cyclical movements:
  - Compared to cyclical, speech movements generally are faster, larger, shorter – perhaps because they have well-defined phonetic targets
- Vowels produced in fast vs. clear speech:
  - larger dispersions, goal-region edges that are closer together – less distinct from one another



- Ellipses indicating the range of formant frequencies (+/-1 s.d.) used by a speaker to produce five vowels during fast speech (light gray) and clear speech (dark gray) in a variety of phonetic contexts.

Persistence of inaudible gestures at word boundaries

U. Tokyo X-ray μ-beam
(Fujimura *et al*. 1973)

List Production

“perfect, memory”

Figure removed due to copyright reasons.

Phrasal Production

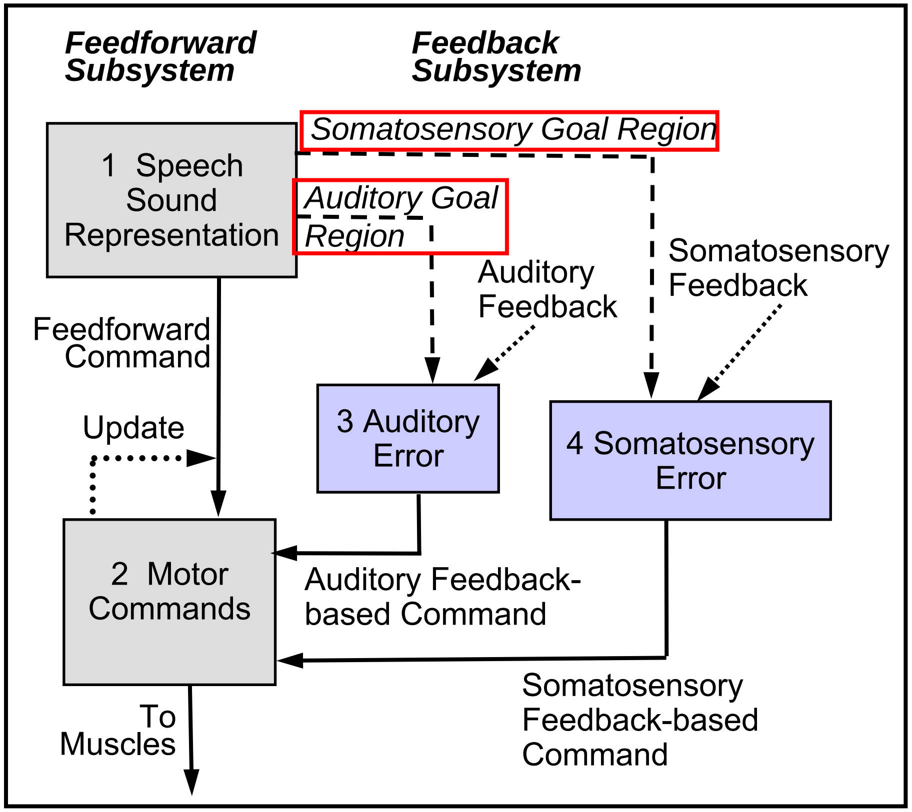
“perfec(t) memory”

- Phrasal: /m/ closure overlaps /t/ release, making it inaudible; /t/ gesture is present nevertheless (c.f. Browman & Goldstein; Saltzman & Munhall)
- Findings replicated and expanded with 21 speakers
- Explanation (DIVA): Frequently used phonemes, syllables, words become □ encoded as feedforward command sequences □

Outline

- Introduction
- Measuring speech production
- What are the "controlled variables" for segmental speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
  – DIVA: feedback & feedforward control
  – Long term effects: Hearing loss and restoration
  – An example of abrupt hearing and then motor loss
  – Responses to perturbations – auditory and articulatory
    - "Steady state" perturbations
    - Gradually increasing perturbations
    - Abrupt, unanticipated perturbations
  – Feedback vs. feedforward mechanisms in error correction
- Summary

Feedback and Feedforward Control in DIVA



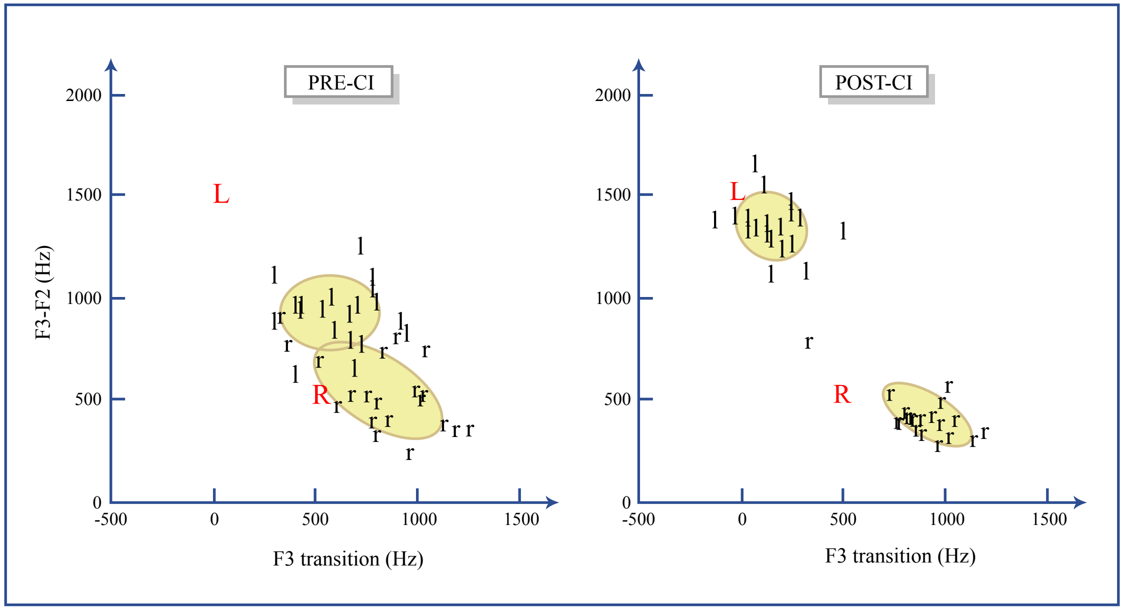
- With acquisition, control becomes *predominantly feedforward*
- Feedback control – Uses *error detection and correction* - to teach, refine and update feedforward control mechanisms
- Experiments can shed light on
  - *sensory goals*
  - *error correction*
  - *mappings between motor/sensory and acoustic/auditory parameters*

Learning and maintaining phonemic goals: Use of Auditory Feedback

- Audition is crucial for normal speech acquisition
- Postlingual deafness: Intelligible speech, but with some abnormalities
- Regain some hearing with a Cochlear Implant (CI):
  - Usually show parallel improvements in perception, production and intelligibility

Acoustic measures of contrast between /l/ and /r/

6 months after receiving a CI



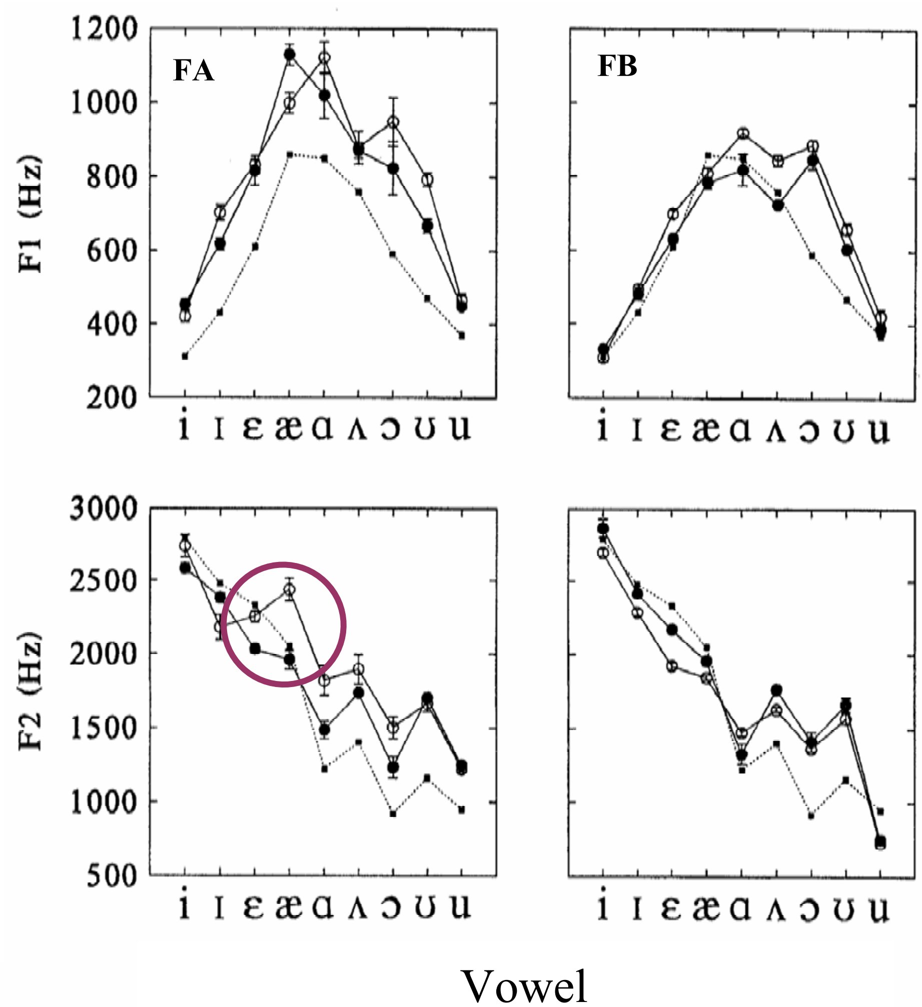
Figures by MIT OCW.

- Phonemic contrast is enhanced pre- to post-implant – typical for CI users, many of whom have somewhat diminished contrasts pre-implant

Long-term stability of auditory-phonemic goals for vowels

- Typical pre- (○) and post- (●) implant formant patterns: generally congruent with normative data (▪)
  - FA: some irregularity of F2 pre-implant (18 years after onset of profound hearing loss)
  - One year post-implant: F2 values are more like normative ones
- Phonemic identity doesn't change; degree of contrast can
- Goals and feedforward commands for vowels generally are stable
  - If they degrade from hearing loss, can be recalibrated with hearing from a CI

Data from 2 cochlear implant users

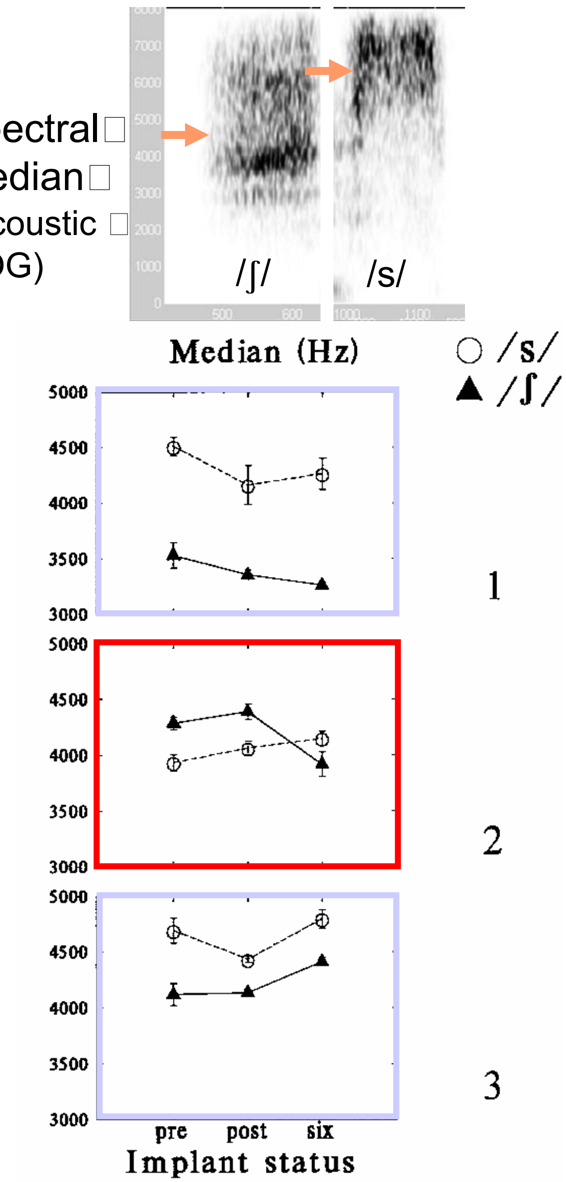


Reprinted with permission from:
Perkell, J., H. Lane, M. Svirsky, and J. Webster. "Speech of cochlear implant patients: A longitudinal study of vowel production." *J Acoust Soc Am* 91 (1992): 2961-2979.

Long-term stability of phonemic goals for sibilants in CI users

- Subjects 1 and 3: good distinctions between /s/ and /ʃ/ pre-implant –
  - Typical, decades following onset o hearing loss
- Subject 2: reversed values and distorted productions pre-implant
  - After about 6 months of implant use, sibilant productions improved
- These precisely differentiated articulations are usually maintained for years without hearing
  - Possibly because of the use of *somatosensory goals* – e.g. pattern of contact between tongue, teeth and palate

Spectral median (Acoustic COG)



Reprinted with permission from:
Matthies, M. L., M. A. Svirsky, H. Lane, and J. S. Perkell. "A preliminary study of the effects of cochlear implants on the production of sibilants." *J Acoust Soc Am* 96 (1994): 1367-1373.

Responses to abrupt changes in hearing and motor innervation

An NF2 patient with sudden hearing loss, followed by some motor loss

- Two surgical interventions
  - OHL: Onset of a significant hearing loss (especially spectral) from removal of an acoustic neuroma
  - Hypoglossal nerve transposition surgery → Some tongue weakness
- /s-ʃ/ contrast: Good until second surgery, when contrast collapsed
- Hypothesis: Feedforward mappings invalidated by transposition surgery
  - Without spectral auditory feedback, compensatory adaptation (relearning) was impossible – as might be possible with hearing
  - Somatosensory goal deteriorated without auditory reinforcement

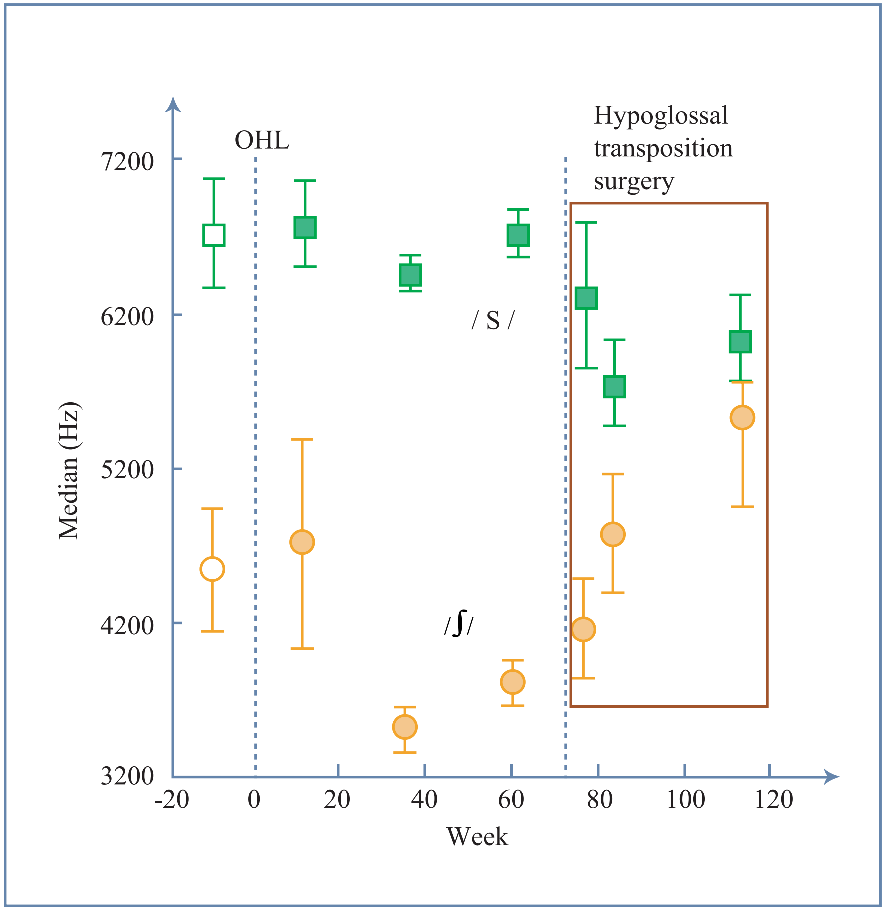


Figure by MIT OCW.

Spectral median for /s/ and /ʃ/ vs. weeks in an NF2 patient

Vowel Contrasts and Hearing Status

- Compare English with Spanish CI users, CI processor OFF and ON
- Previous findings: Contrasts increase with hearing, decrease without
- Hypothesis: Because of the more crowded vowel space in English, turning the CI processor OFF and ON will produce more consistent decreases and increases in vowel contrasts in English than in Spanish

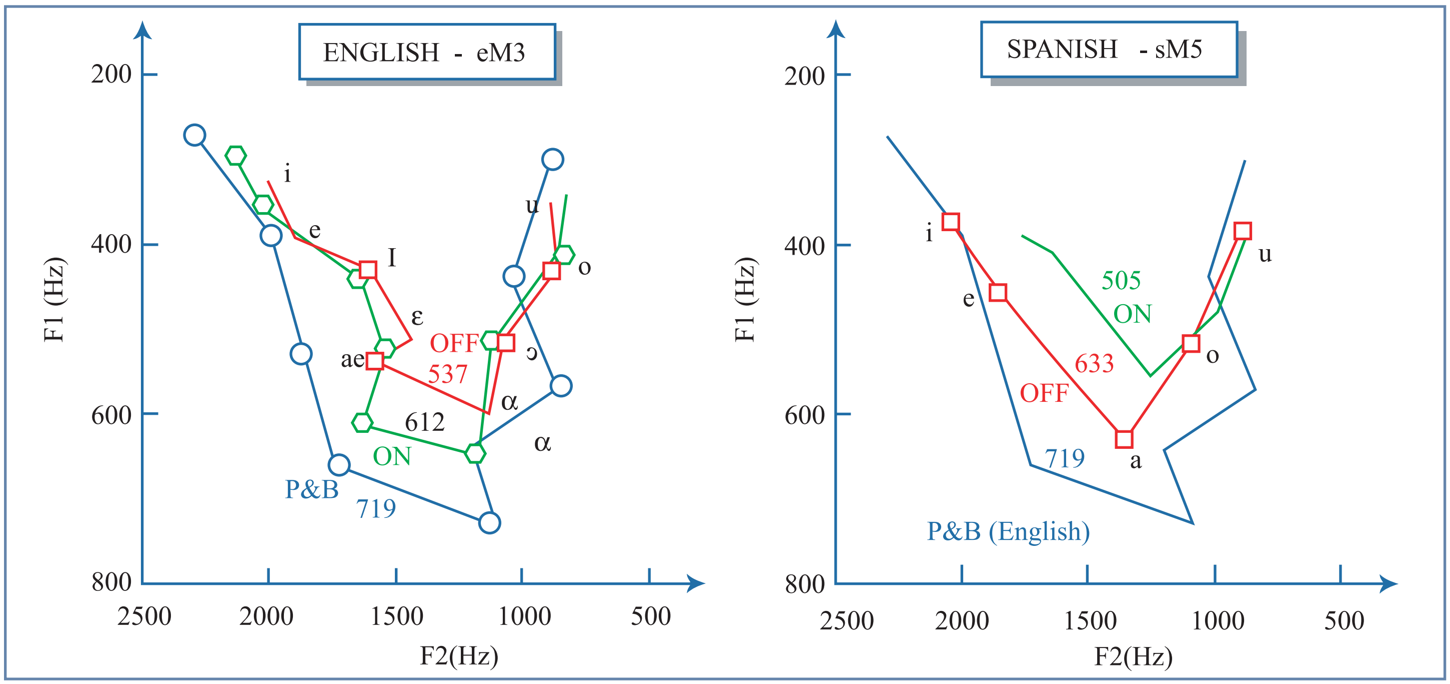


Figure by MIT OCW.

- Average Vowel Spacing (AVS) – a measure of overall vowel contrast
- Change of AVS from processor ON to processor OFF (for 24 hours)
  - AVS: decreases for the English speaker, increases for the Spanish speaker

AVS – by subject

- Prediction: AVS increases with the CI processor (hearing)
- Changes follow the predicted pattern more consistently for English than Spanish speakers

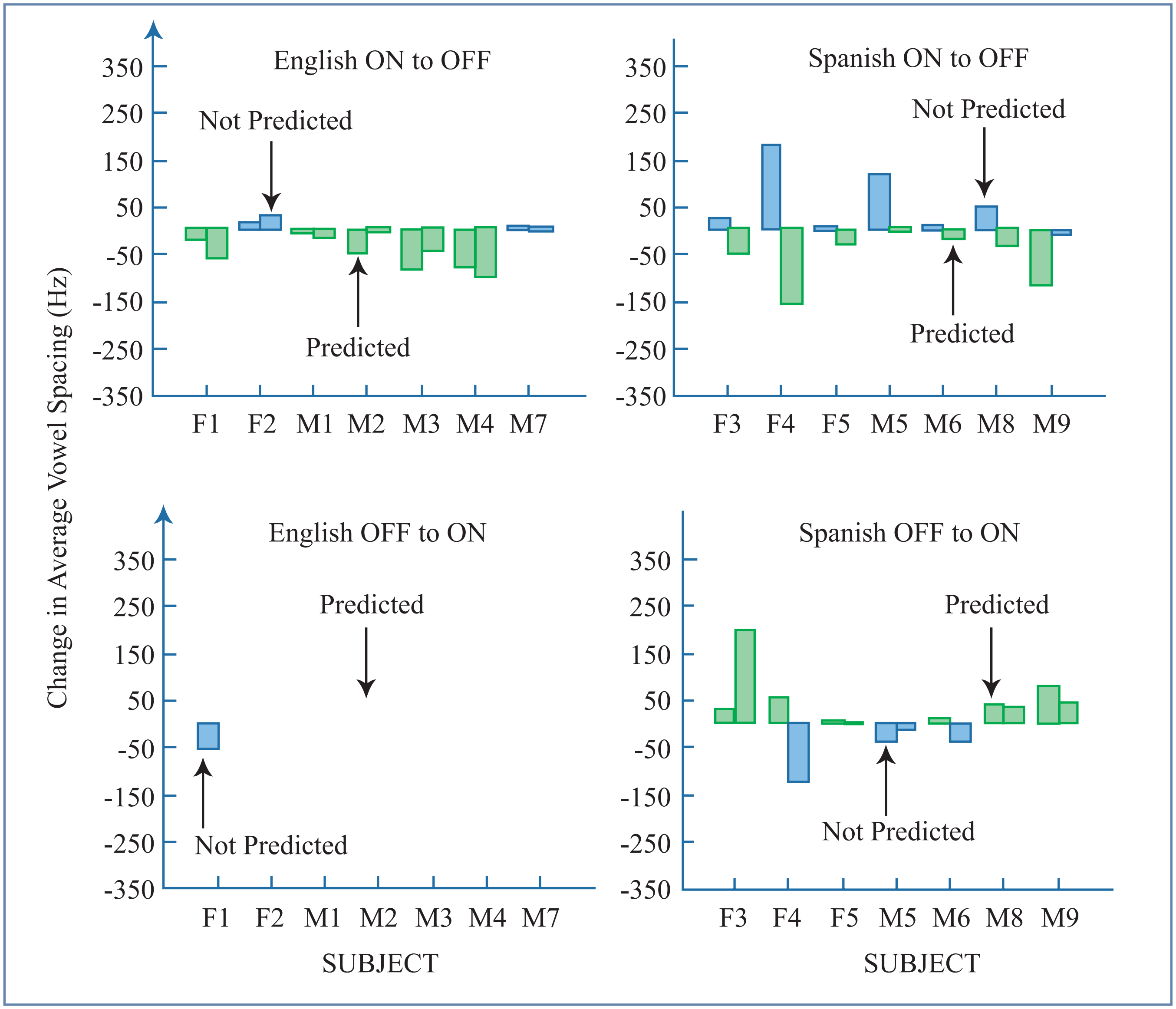
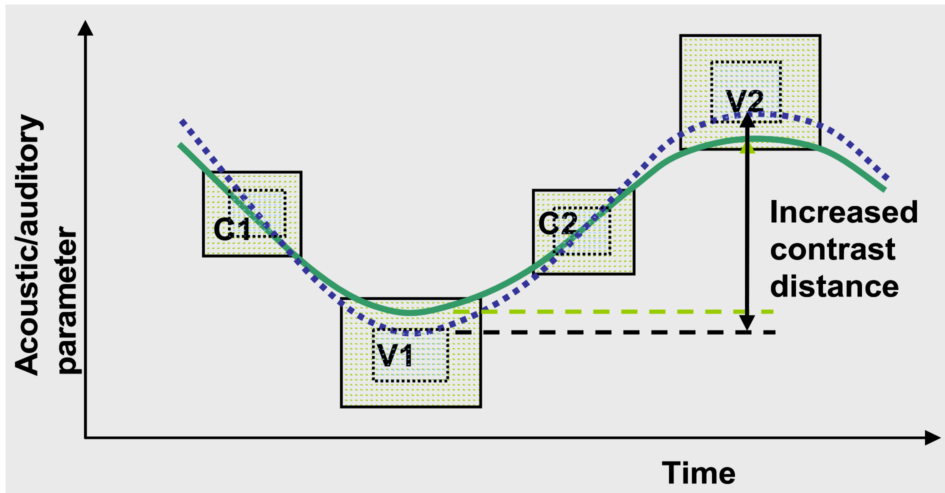


Figure by MIT OCW.

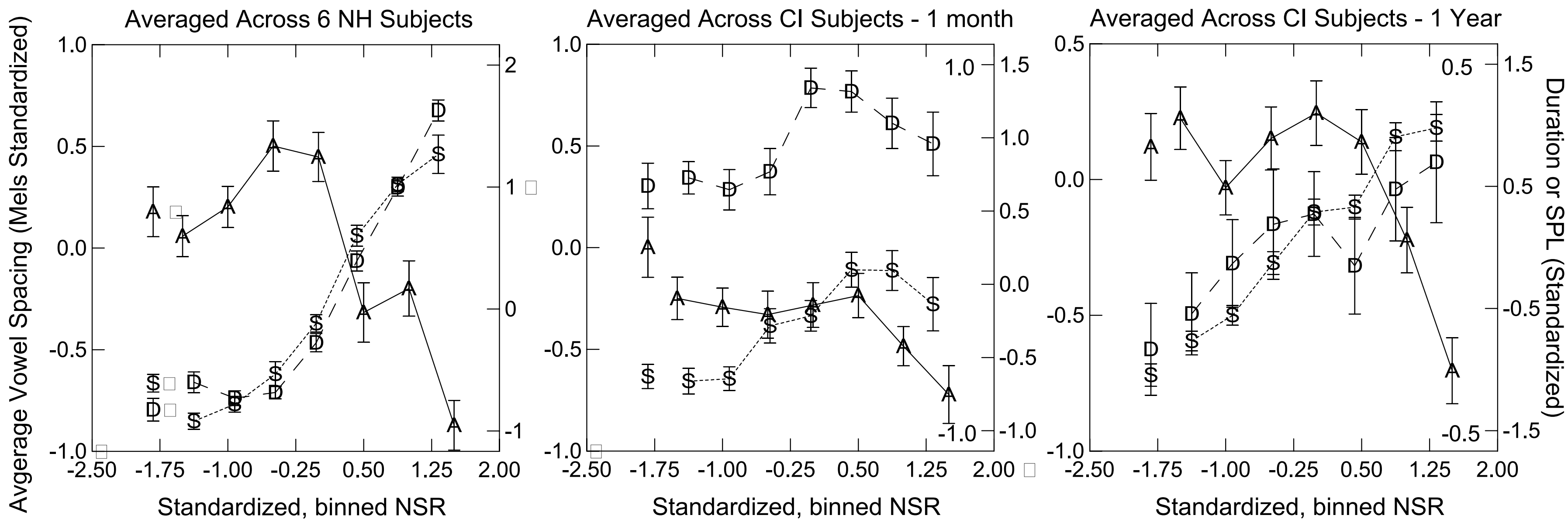
Modeling Contrast Changes: Clarity vs. Economy of Effort

- DIVA contains a parameter that changes sizes of all goal regions simultaneously – to control speaking rate and clarity (e.g., AVS)



- Shrinking goal region size – like what English speakers do with hearing
  - Produces *increased clarity* (contrast distance), *decreased dispersion*
  - Without hearing, *economy of effort* dominates
- With fewer vowels in Spanish, clarity demands aren't as stringent
  - Acceptable contrasts may be produced regardless of hearing status, without changing goal region size

Effect of Varying S/N in Auditory Feedback



- Normal-hearing and CI subjects heard their own vowel productions mixed with increasing amounts of noise
- In general, AVS increased, then decreased with increasing N/S
- Possible explanation: With increasing N/S
  - If auditory feedback is sufficient, clarity is increased
  - As feedback becomes less useful, economy of effort predominates
- Similar result for /s-ʃ/ contrast, but with peak at lower NSR

Bite block experiments

Figures removed due to copyright reasons.

- Speakers compensate fairly well with the mandible held at unusual degrees of opening
- Compensations may be better for quantal vowels (with better-defined articulatory targets)
- Presumably, the speakers mappings are not as accurate for the perturbed condition
- Compensation continues to improve, possibly with the help of auditory feedback

Mappings can be Temporarily Modified: Auditory Feedback
(Houde and Jordan)

- Sensorimotor Adaptation
- Methods:
  - Fed-back (whispered) vowel formants were gradually shifted
  - 16 msec delay
  - Subjects were unaware of shift

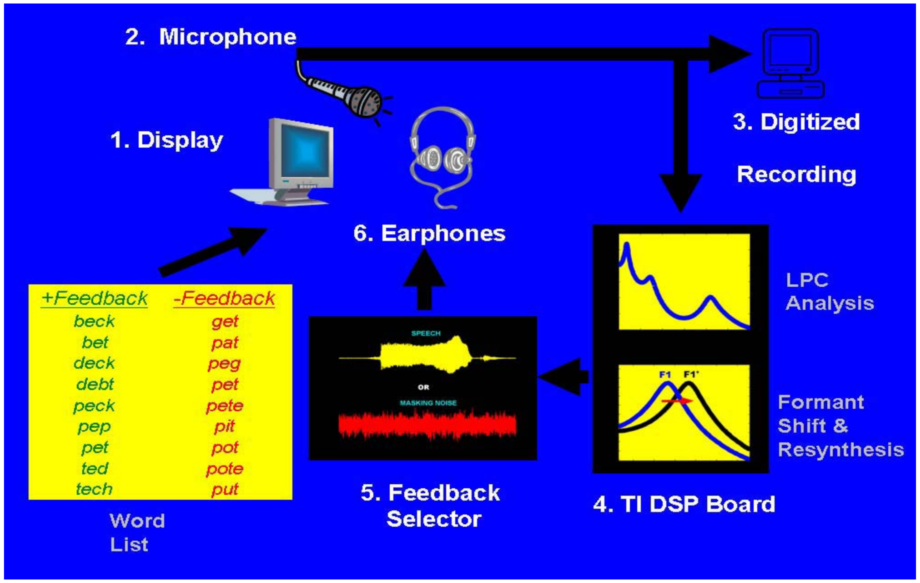
Figures removed due to copyright reasons.
Please see:
Figures from: Houde, J. F. and M. I. Jordan. "Sensorimotor adaptation of speech I: Compensation and adaptation." *J Speech, Language, Hearing Research* 45 (2002): 295-310.

- Results:
  - Subjects adapted for shift by modifying productions in the opposite direction

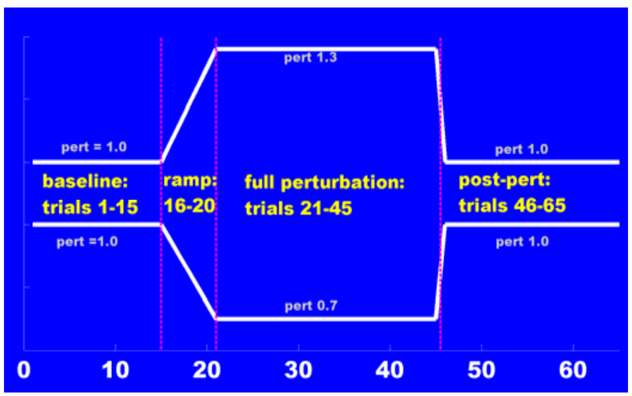
Figures removed due to copyright reasons.
Please see:
Figure from: Houde, John Francis. "Sensorimotor adaptation in speech production." Thesis (Ph. D.)--Massachusetts Institute of Technology, Dept. of Brain and Cognitive Sciences, 1997.

  - Effect generalized to other consonant environments and to other vowels
  - Effect persisted in the presence of masking noise: “Adaptation”
  - Adaptation was exhibited later, simply by putting subjects in the apparatus (no shift)
- Speakers use auditory goals and auditory-motor mappings.

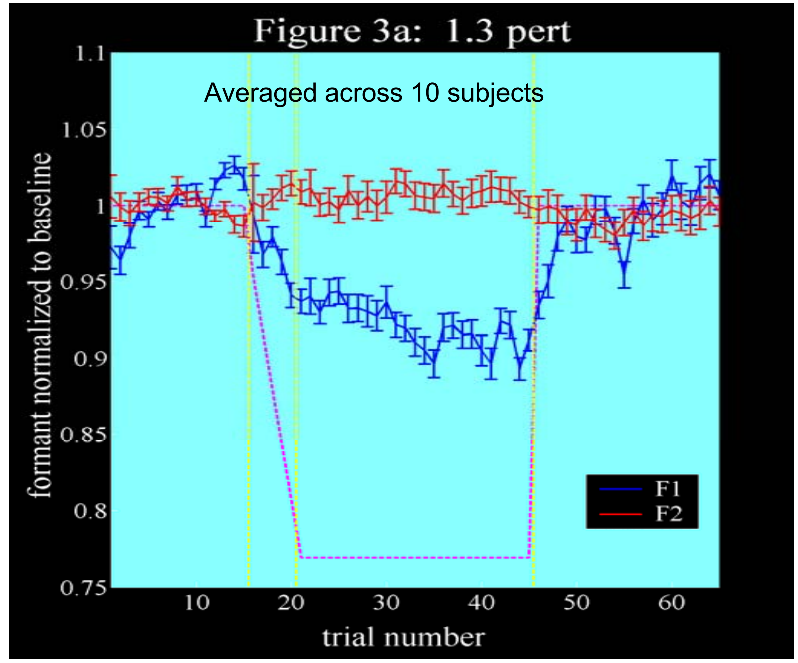
Sensorimotor adaptation



- Thesis project of Virgilio Villacorta
  - Based on work of Houde & Jordan, but with voiced vowels

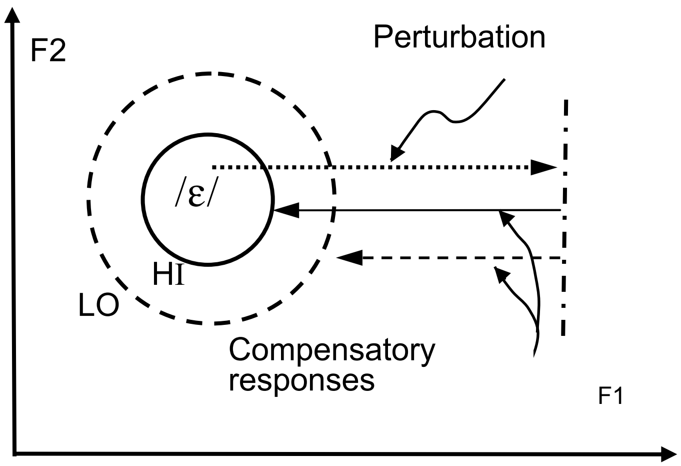


- Subjects hear own vowels with F1 perturbed, are unaware of perturbation

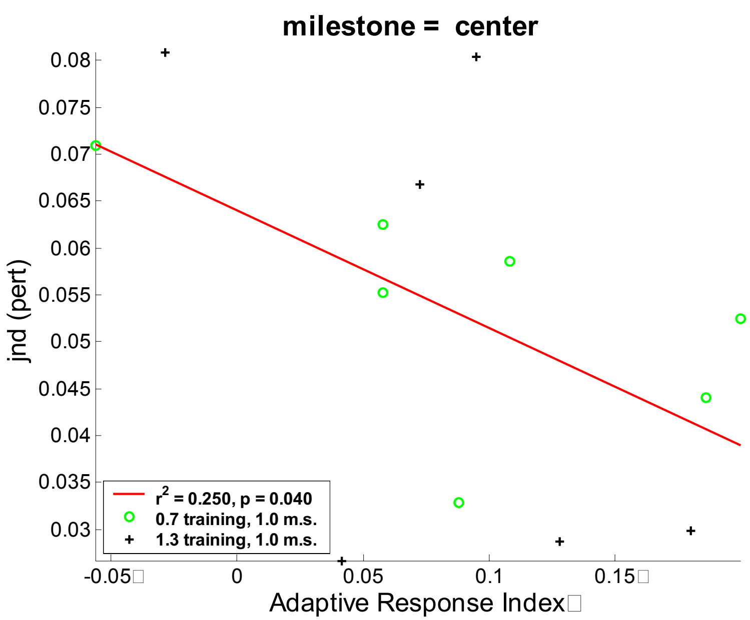


- Subjects partially compensate by shifting F1 in opposite direction; Shift is formant-specific
- Mismatch between expected and produced auditory sensations → Error correction
- 20 subjects: Varied in amount of compensation
- Is there a relation between perceptual acuity and amount of compensation?

Relation Between Adaptation and Auditory Discrimination



Hypothetical compensatory responses to F1 perturbation by High- and Low-Acuity speakers

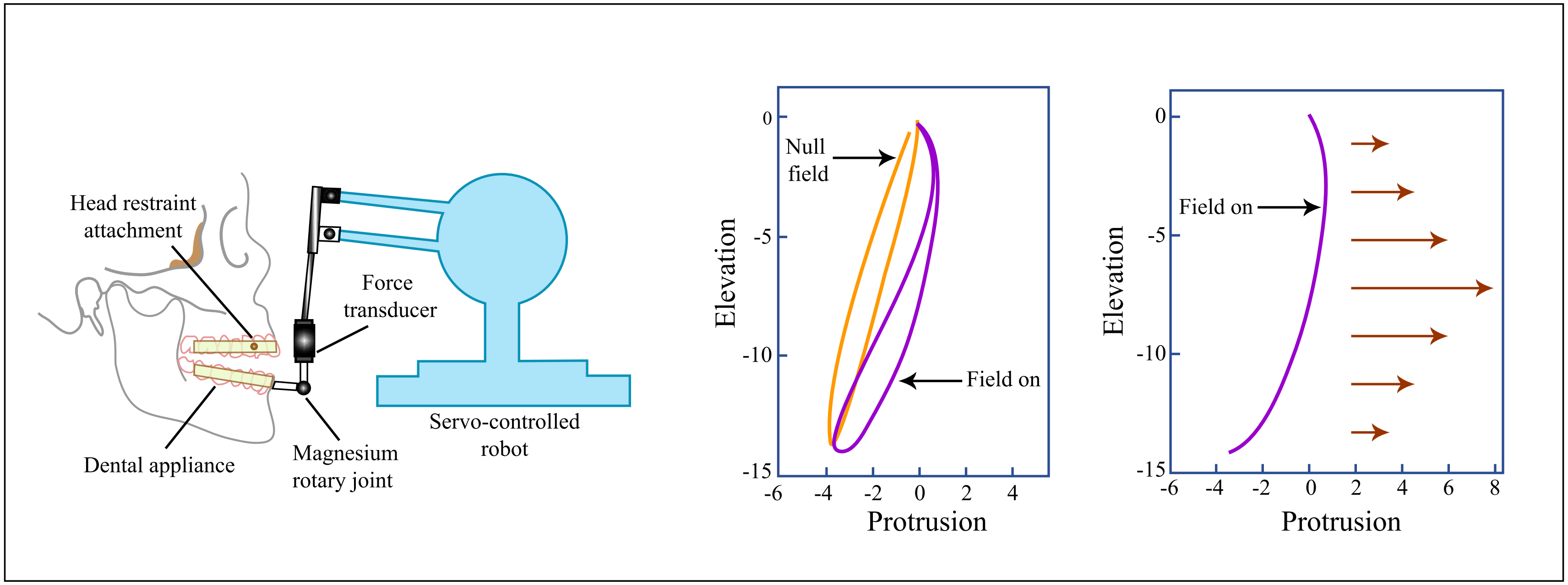


- DIVA and previous studies: Production goals for vowels are primarily regions in auditory space
  - Speakers with more acute auditory discrimination have smaller goal regions, spaced further apart
- Hypothesis:
  - More acute perceivers will adapt more to perturbation
- Measure of auditory acuity: JNDs on pairs of synthetic vowel stimuli
- Result: Hypothesis is supported

	first order $r_{xy\|\|z}$		
	jnd, ARI||F1 sep	*F1 sep, ARI||jnd*	*jnd, F1 sep||ARI*
r	**-0.717**	**0.661**	**0.656**
r^2	**0.514**	**0.436**	**0.430**
p-score	**0.004**	**0.010**	**0.021***

Modifying a Somatosensory-to-Motor Mapping

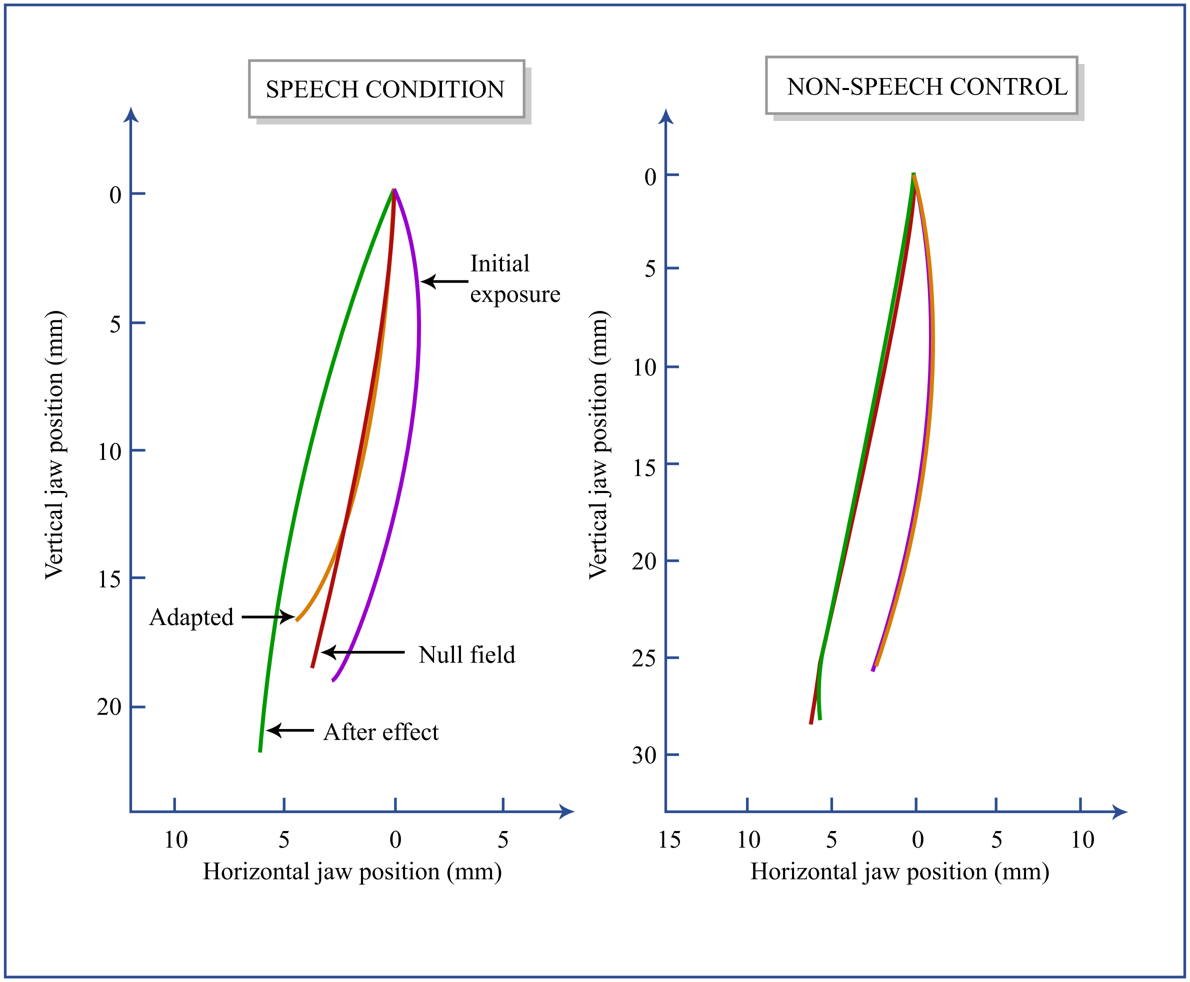
A "Force Field Adaptation" experiment (Ostry & colleagues)



Figures by MIT OCW.

- Methods:
  - Velocity-dependent forces applied (gradually) by a robotic device act to protrude the jaw: proportional to instantaneous jaw lowering or raising velocity
  - Jaw motion path over large number of repetitions (700) is used to assess adaptation, which may be evidence of:
    - Modification of somatosensory-motor mappings
    - Incorporation of information about dynamics in speech movement planning (Ostry's interpretation)

Results



Figures by MIT OCW.

Summary and interpretation

- Subjects adapt to a motion dependent force field applied to the jaw during speech production
- Kinesthetic feedback alone is not sufficient for adaptation; have to be in a "speech mode"
- Control signals (mappings) are updated based on differences between expected and actual feedback
- Information about dynamics is incorporated in speech motor planning

Rapid drift of spectral median for /ʃ/

- A CI "on-off" experiment

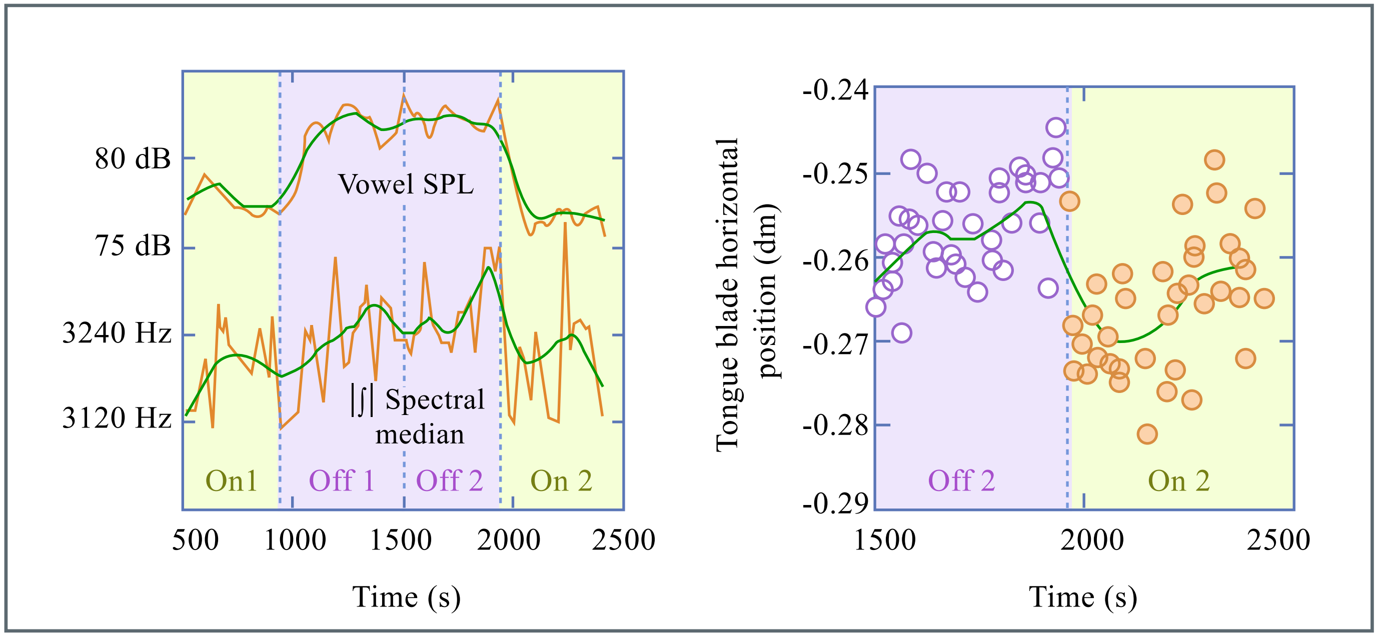
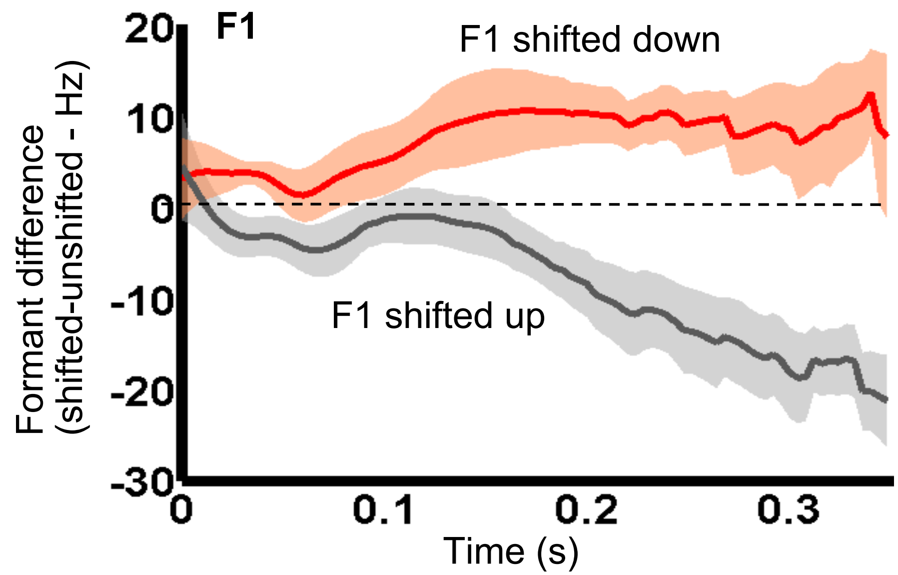


Figure by MIT OCW.

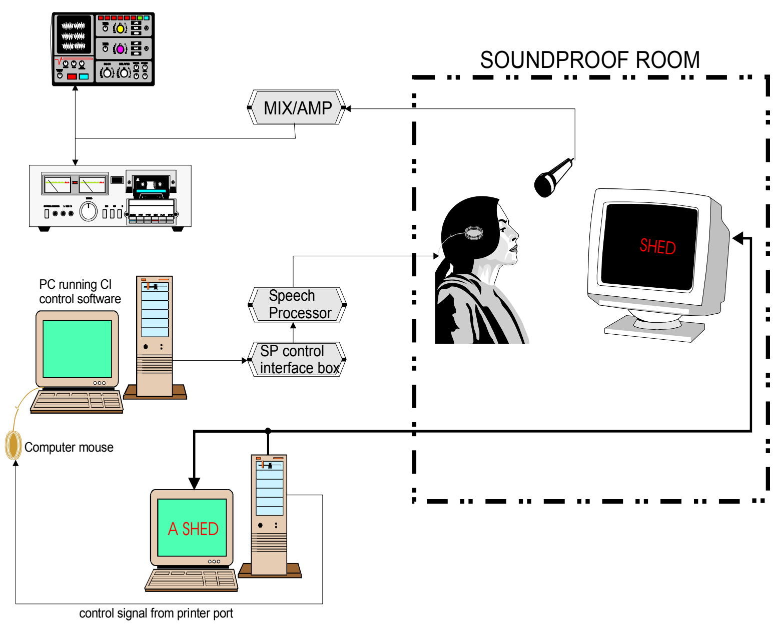
- Observations
  - Vowel SPL increased rapidly with CI processor off, decreased with processor on
  - Spectral median drifted upward toward /s/ during the 1000 seconds with processor off – Surprising, since the goals are usually stable
  - Hearing one aberrant utterance when the processor was turned on, speaker overcompensated to restore an appropriate /ʃ/
- Extremely narrow dental arches (and movement transducer coil on tongue) may have made it difficult for speaker to rely on somatosensory goal
- He may have had to rely predominantly on auditory feedback to maintain feedforward control on an utterance-to-utterance basis

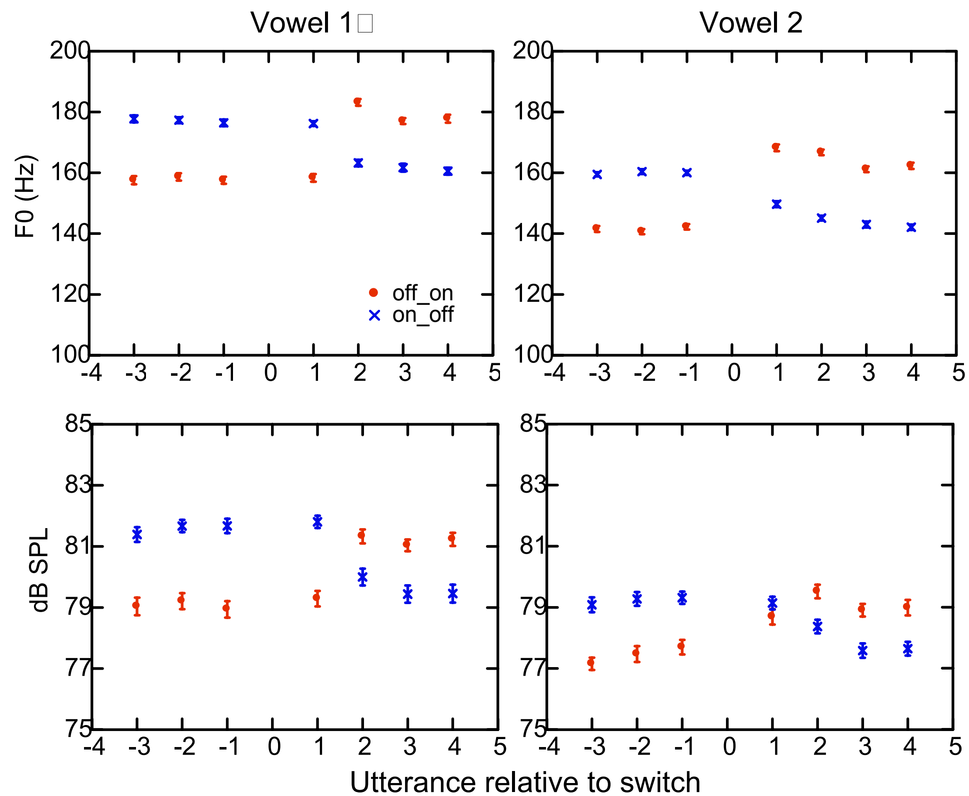
Unanticipated Acoustic Perturbations (Tourville et al., 2005)



- Methods – like sensorimotor adaptation, but with sudden, unanticipated shift of F1
  - Subjects pronounced /CεC/ words with auditory feedback through a DSP board
  - In 1 of 4 trials, F1 was shifted up toward /æ/ or down toward /ɪ/
- Results (averaged across 11 subjects)
  - Subjects produced compensatory modification of F1, in direction opposite to shift
  - Delay of about 150 ms. – compatible with other results, in which F0 was shifted.
- Result is compatible with error detection and correction mechanisms in DIVA

How long does it take for parameters to change when hearing is turned on or off?





- Subject pronounced a large number of repetitions of four 2-syllable utterances (e.g., *done shed, don said*; quasi-random order).
- CI processor state (hearing) was switched between on and off unexpectedly

- Example results for one subject
  – Changes not evident until second vowel
  – Change may be more gradual for SPL than for F0
- Results varied among parameter and subject
  – Perhaps related to subject acuity?

Unanticipated Movement perturbation – Motor Responses

Please see:
Abbs, J. H., and V. L. Gracco. "Control of complex motor gestures -orofacial muscle responses to load perturbations of lip during speech." *Journal of Neurophysiology* 51 (1984): 705-723.

- Abbs et al.; others (1980s)
  - In response to downward perturbation of lower lip in closure toward a /p/,
  - Upper lip responds with increased downward displacement, accompanied by EMG and velocity increases
  - The response is phoneme-specific

Further observations and interpretation (Abbs et al.)

Figures removed due to copyright reasons.

- Motor equivalence at the muscle and movement levels

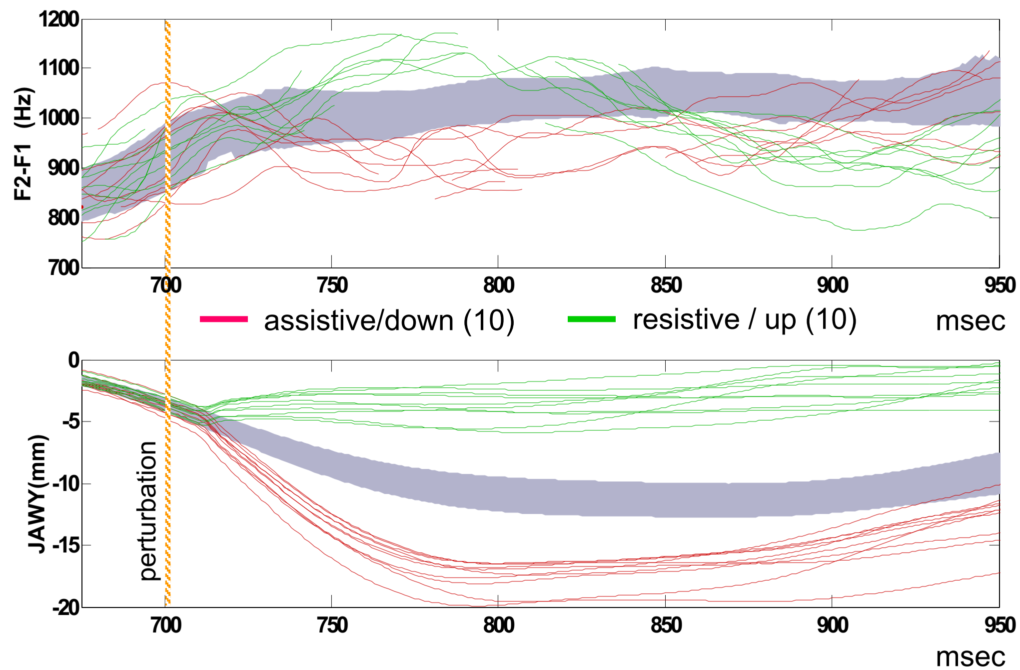
- Coordinated speech gestures are performed by “synergisms” –
  - Temporarily recruited combinations of neural and muscular elements that convert a simple input into a relatively complex set of motor commands
- There are alternative interpretations (Gomi, et al.)

Motor and acoustic responses to unanticipated jaw perturbations

(In collaboration with David Ostry)

"see red" N = 100 (two 50 rep trials)

- Robot used to perturb jaw movements
  - Triggered by downward movement
- 50 repetitions/utterance e.g., "see red"
  - 5 perturbed upward (resistive)
  - 5 downward (assistive)



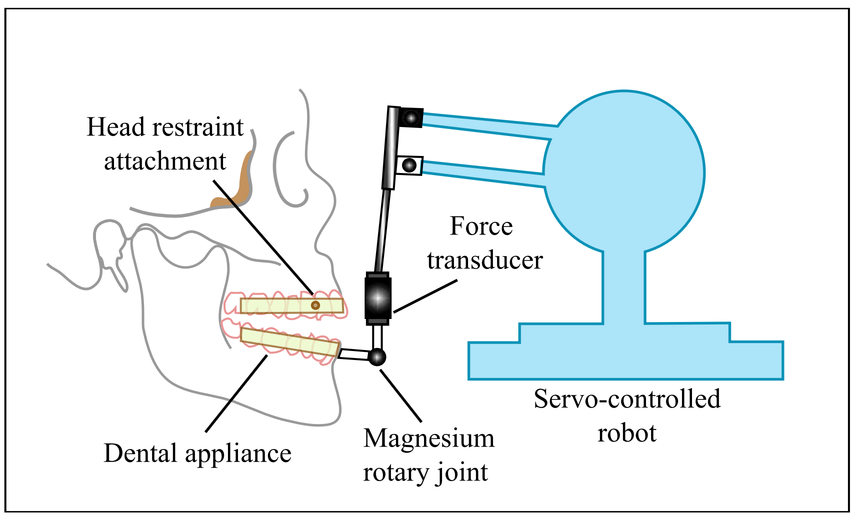


Figure by MIT OCW.

- Formants begin to recover 60-90 ms after perturbation; jaw does not
- Two other subjects were similar
- Evidence of within-movement, closed-loop error correction

Compensatory Responses to Unexpected Palatal Perturbation

(Honda , Fujino & Murano)

- A subject pronounced phrase: /iɑ ʃɑ ʃɑ ʃɑ ʃɑ ʃɑ ʃɑ ʃɑ ʃɑ/
- Movements and acoustic signal were recorded
- Palatal configuration was perturbed by inflation of a small balloon on 20% of trials (randomly determined)
- Feedback conditions:
  - Feedback not blocked
  - Auditory feedback blocked with masking noise
  - Tactile feedback blocked with topical anesthesia
  - Both types of feedback blocked
- Measures
  - Articulatory compensations
  - Listener judgments of distorted sibilants

Results

- Perturbation caused distortions in /ʃ/ production
  - Compensation and feedback:
    - With feedback not blocked, speaker compensated within about 2 syllables
    - With auditory or tactile feedback blocked, speaker was much less able to compensate
    - With both forms of feedback blocked, compensation was worst
- Results are compatible with
  - Sensory goals as basic units
  - Use of mismatches between expected and actual sensory consequences to correct feedforward commands

Figures removed due to copyright reasons.

Mean error score for fricative consonant identification (%)

Syllable No.	1	2	3	4	5	6	7	8
Normal auditory-feedback								
Steady-state deflated	0	0	0	0	0	0	0	0
Inflation	83	14	0	0	0	0	0	0
Deflation	8	0	0	0	0	0	0	0
Steady-state inflated	0	0	0	0	0	0	0	0
Masked auditory-feedback								
Steady-state deflated	0	0	0	0	0	14	11	11
Inflation	72	39	39	44	28	42	33	50
Deflation	0	0	0	3	3	11	11	17
Steady-state inflated	47	39	44	50	44	44	44	44

Feedforward vs. Feedback Control and Error Correction

- In DIVA, feedback and feedforward control operate simultaneously; feedforward usually predominates
- Feedback control intervenes when there is a large enough mismatch between expected and produced sensory consequences (sensorimotor adaptation results)
- Timing of correction:
  - With a long enough movement, correction is expressed (closed loop) during the movement (e.g., “see red”)
  - Otherwise, correction is expressed in the feedforward control of following movements (e.g., /ʃ/ spectrum, vowel SPL, F0 when CI turned on)
  - Correction to an auditory perturbation takes longer than to a somatosensory perturbation (presumably due to different processing times)
- Additional Examples of error correction
  - Closed-loop responses to perturbations (see Abbs, others)
  - Feedforward error correction with, e.g., dental appliances
  - Responses to combined perturbations (cf. Honda & Murano)
  - All are compatible with DIVA’s use of feedback

Outline

- Introduction
- Measuring speech production
- What are the “controlled variables” for segmental speech movements?
- Segmental motor programming goals
- Producing speech sounds in sequences
- Experiments on feedback control
- Summary

Summary of Main Points

- Highest level control variables for phonemic movements
  - Auditory-temporal and somatosensory-temporal goal regions
- Goal regions encoded in CNS
  - Projections (mappings) from premotor to sensory cortex: □ Expected sensory consequences of producing speech sounds□
- Goal regions defined partly by articulatory and acoustic saturation effects that are properties of vocal-tract anatomy and acoustics
  - Most vowels: goals primarily auditory; saturation effects, acoustic
  - Consonants: both auditory and somatosensory goals; saturation effects, primarily articulatory (e.g., any consonant closure)
- Articulatory-to-acoustic motor equivalence (/u/, /r/)
  - Help stabilize output of certain acoustic cues
  - Evidence that goals are auditory

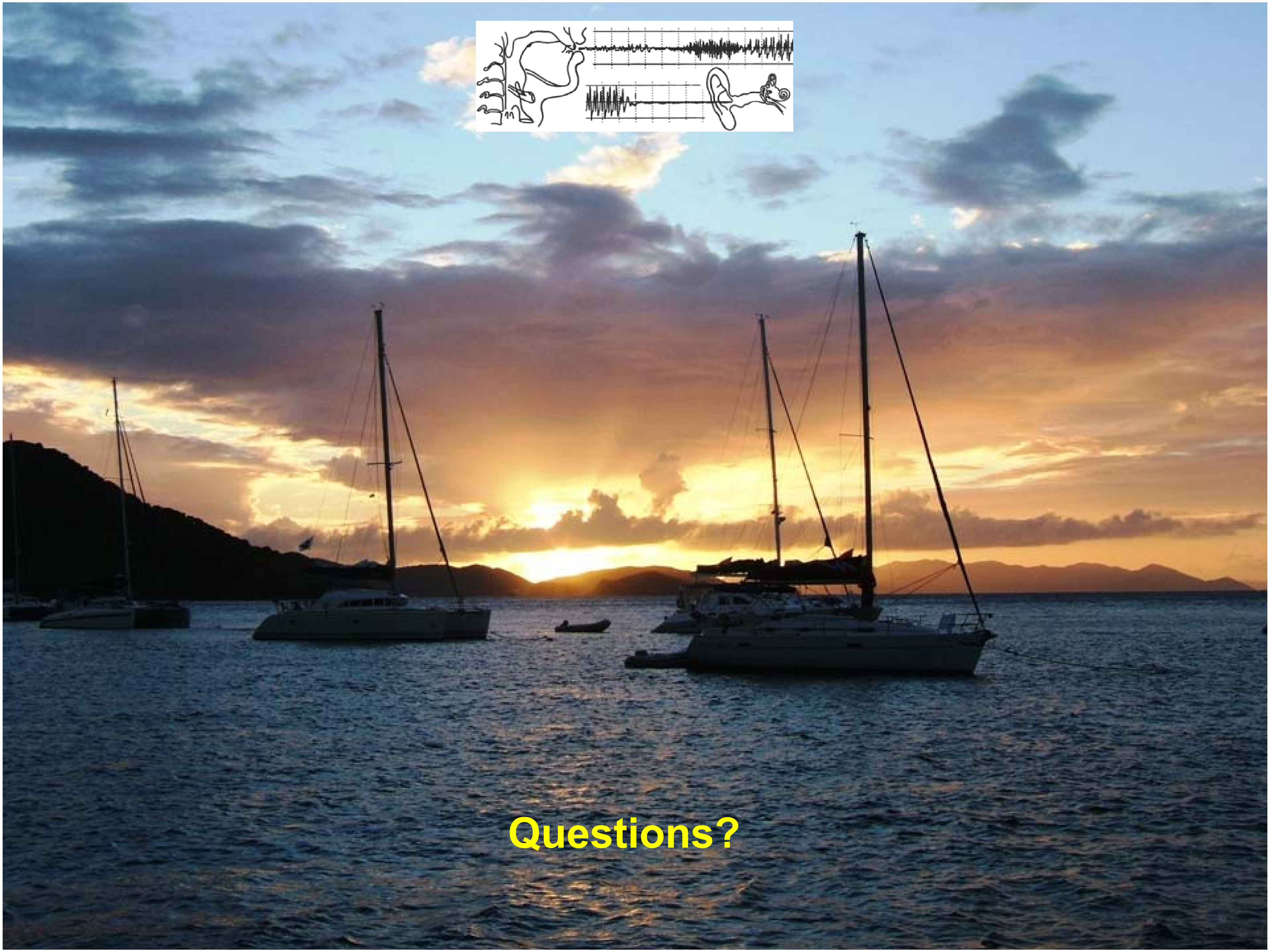
Summary (continued)

- Auditory feedback (CI users)
  - Used to acquire goals and feedforward commands
  - Needed to maintain appropriate motor commands with vocal-tract growth, perturbations
- Goals and feedforward commands are usually stable, even with hearing loss
- Clarity vs. economy of effort
  - Tradeoff evident when hearing (CI) is turned on, off, in presence of noise
- Relations between production and perception
  - Better discriminators produce more distinct sound contrasts
  - Better discriminators may learn smaller, more distinct goal regions
- Feedback and feedforward control
  - Frequently used sounds (syllables, words) are encoded as feedforward commands
  - Responses to perturbations: intra-gesture are closed loop; inter-gesture are via adjustments to subsequent feedforward commands

DIVA (Next lecture)

Figures removed due to copyright reasons.

- Components have hypothesized correlates in cortical activation
- Hypotheses can be tested with brain imaging
- Can quantify relations among *phonemic specifications*, *cortical activity*, *movement* and the *speech sound output*


Questions?

SOME REFERENCES

- Abbs, J.H. and Gracco, V.L. (1984). Control of complex motor gestures - orofacial muscle responses to load perturbations of lip during speech. *Journal of Neurophysiology*, *51*, 705-723.
- Bradlow, A.R., Pisoni, D.B., Akahane-Yamada, R. and Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production, *J. Acoust. Soc. Am.* 101, 2299-2310.
- Browman, C.P. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*, 201-251.
- Fujimura, O. & Kakita, Y. (1979). Remarks on quantitative description of lingual articulation. In B. Lindblom and S. Öhman (eds.) Frontiers of Speech Communication Research, Academic Press.
- Fujimura, O., Kiritani,S. Ishida,H. (1973). Computer controlled radiography for observation of movements of articulatory and other human organs, *Comput. Biol. Med.* 3, 371-384
- Guenther, F.H., Espy-Wilson, C., Boyce, S.E., Matthies, M.L., Zandipour, M. and Perkell, J.S. (1999) Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *J. Acoust. Soc. Am.*, *105*, 2854-2865.
- Honda, M. & Murano, E.Z. (2003). Effects of tactile and auditory feedback on compensatory articulatory response to an unexpected palatal perturbation, Proceedings of the 6th International Seminar on Speech Production, Sydney, December 7 to 10.
- Houde, J.F. and Jordan, M.I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *J. Speech, Language, Hearing Research*, 45, 295-310.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (eds.), Speech production and speech modeling. Dordrecht: Kluwer. 403-439.
- Newman, R.S. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. *J. Acoust. Soc. Am. 113*, 2850-2860.
- Perkell, J.S. and Nelson, W.L. (1985). Variability in production of the vowels /i/ and /a/, *J. Acoust. Soc. Am. 77*, 1889-1895.
- Saltzman, E.L. and Munhall, K.G. (1989). A dynamical approach to gestural patterning in speech production, *Ecological Psychology, 1*, 333-382.
- Stevens, K.N. (1989). On the quantal nature of speech. *J. Phonetics*,*17*, 3-45.

Slide 6 (right figure):
Denes, Peter B., and Elliot N. Pinson. *The Speech Chain: The Physics and Biology of Spoken Language.* New York, NY: W.H. Freeman, 1993. ISBN: 0716723441.

Slide 16 (left figure):
Grant, J. C. Boileau (John Charles Boileau). *Grant's Atlas of Anatomy*. 5th ed. Baltimore, MD: Williams & Wilkins, 1962.

Slide 18, Slide 27 (right three), Slide 28, Slide 55:
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico, and M. Zandipour. "A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss." *J Phonetics* 28 (2000): 233-372.

Slide 20-21:
Stevens, K. N. "On the quantal nature of speech." *J Phonetics* 17 (1989): 3-45.

Slide 23 (left):
Fujimura, O., and Y. Kakita. "Remarks on quantitative description of lingual articulation." *Frontiers of Speech Communication Research*. Edited by B. Lindblom and S. Öhman. Academic Press, 1979.

Slide 23 (mid/lower 2):
Perkell, J. S. "Properties of the tongue help to define vowel categories: Hypotheses based on physiologically-oriented modeling." *Journal of Phonetics* 24 (1996): 3-22.

Slide 24:
Denes, P. B., and E. N. Pinson. *The Speech Chain*. New York, NY: W.H. Freeman, 1993, p. 68. ISBN: 0716722569.

Slides 26 (left set):
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico, and M. Zandipour. “A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss.” *J Phonetics* 28 (2000): 233-372.

Slide 26 (top right):
Perkell, J. S. “Properties of the tongue help to define vowel categories: Hypotheses based on physiologically-oriented modeling.” *Journal of Phonetics* 24 (1996): 3-22.

Slide 27:
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico, and M. Zandipour. "A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss." *J Phonetics* 28 (2000): 233-372.

Slide 42:
Perkell, J. S. “Articulatory Processes.” *The Handbook of Phonetic Sciences*. Edited by W. Hardcastle and J. Laver. Blackwell, 1997, pp. 333-370.

Slide 56, Slide 57:
Perkell, J., W. Numa, J. Vick, H. Lane, T. Balkany, and J. Gould. “Language-specific, hearing-related changes in vowel spaces: A preliminary study of English- and Spanish-speaking cochlear implant users.” *Ear and Hearing* 22 (2001): 461-470.

Slides 66:
Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, P. Perrier, J. Vick, R. Wilhelms-Tricarico, and M. Zandipour. “A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss.” *J Phonetics* 28 (2000): 233-372.