24.963 Linguistic Phonetics
Fall 2005

24.963

Linguistic Phonetics

Speech Production

Reading for week 11 TBA

Assignment: Experiment 2 - Voicing effects on formants

- Finish any uncompleted assignments

Experiment 2

- Voicing of a following obstruent has long been known to affect the realization of /aɪ/ in many (most?) dialects of English.
- Most dramatic case is 'Canadian raising', more subtle differences are typical.
- Most consistent difference seems to be that F2 is higher and F1 is lower in the offglide before voiceless Cs.
- Explanation has been elusive - e.g. shorter duration before voiceless might lead us to expect truncation of the offglide, but we see the reverse.

Experiment 2

- Kwong and Stevens looked at a variant of this effect, the effects of underlying voicing of flaps on preceding /aɪ/ in the infamous pair *writer* vs. *rider*.

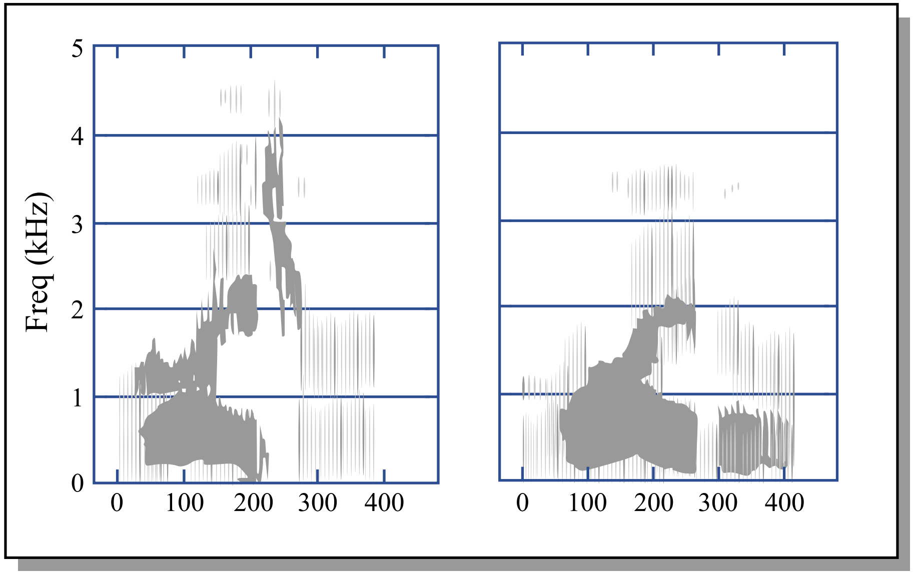


Image by MIT OpenCourseWare. Adapted from Kwong, K. W., and K. N. Stevens. "On the Voiced/Voiceless Distinction for Write/Rider." *Speech Communication Group Working Papers* XI (1999): 1-20. Research Laboratory of Electronics, MIT.

Experiment 2

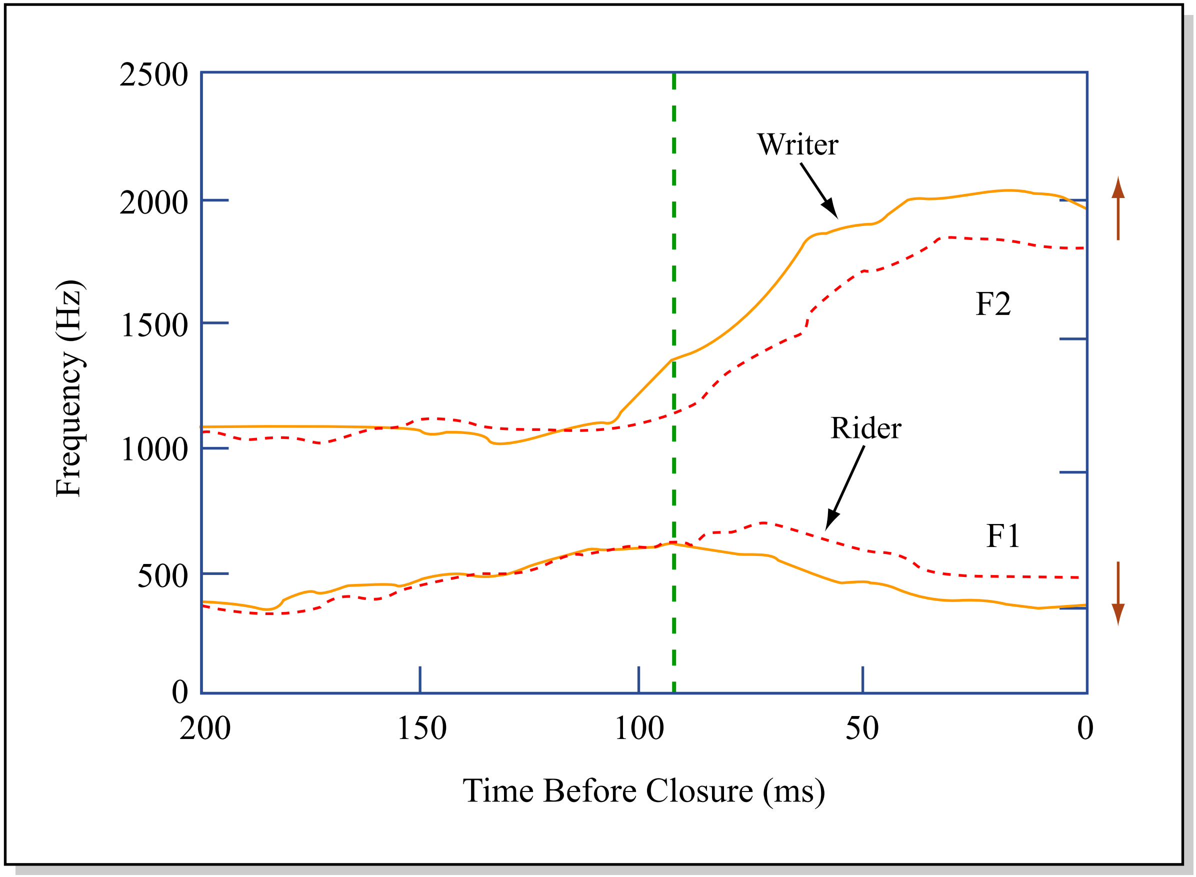


Image by MIT OpenCourseWare. Adapted from Kwong, K. W., and K. N. Stevens. "On the Voiced/Voiceless Distinction for Write/Rider." *Speech Communication Group Working Papers* XI (1999): 1-20. Research Laboratory of Electronics, MIT.

Experiment 2

- Kwong and Stevens propose the following explanation for the difference:
- Voicing in stops is difficult to sustain because the build-up in oral pressure tends to make a sufficient pressure drop across the glottis impossible.
- Oral pressure increase can be slowed by expanding the vocal tract during closure.
- The pharynx is a key site for expansion.
- If pharynx is already fully expanded, as in [i], then no further expansion is possible (also stiffens pharynx walls).
- So pharyngeal expansion should be reduced before voiced stops (to allow for expansion during the stop) but should be mximized before voiceless stops to assist in suppressing voicing.

Experiment 2

- Predicts higher F2 and lower F1 in high front offglides before voiceless.
- Lower F1 and lower F2 in high back offglides.
- No effect on low or lax vowels.

- Let’s test these predictions.

Speech Production

- Speaking is a very complex motor task, involving the coordination of many articulators.
- Degrees of freedom: the speech production system has a large number of degrees of freedom (individual muscle lengths) - more than is required to achieve speech movements.
  - Motor control involves coordination these degrees of freedom to achieve goals.
  - Excess degrees of freedom allow for flexibility, but result in challenges in movement planning and motor learning.

Speech Production

- So one of the central questions is ‘What are the control parameters in speech production?’
  - i.e. the actions of multiple muscles are assumed to be controlled by a single higher level parameter, reducing the degrees of freedom in the system.
  - What is the nature of these control parameters? Articulatory positions? Trajectories? Auditory targets?
- Timing/coordination: Speaking involves coordinating movements in time.
  - How are the control parameters varied over time?
  - How are changesin control parameters coordinated?

Muscles of the speech production mechanism

Muscles of Respiration	
Muscles of inhalation	Muscles of exhalation
Diaphragm	Internal intercostals
External intercostals	Internal obliques
(Pectoralis major)	(External obliques)
(Pectoralis minor)	Rectus abdominis
Scalenes	(Subcostals)
	(Transversus thoracic)

The Muscles of Speech Production

Muscle	Function
Larynx	
A. Intrinsic	
1. Thyroarytenoid (TA)	Vocal cord tensor, forms body of vocal cord;is active during f_o change. Acts to change thickness of vocal cord for register changes; may also act to change overall tension of vocal cord for phonation in different registers.
2. Posterior cricoarytenoid (PCA)	Opens the glottis for either breathing or the production of *-voiced* sounds.
3. Lateral cricoarytenoid (LCA)	Adducts the vocal cords; applies medial compression; is active during f_o changes, always active in onset of phonation when it adducts vocal cords, setting phonation neutral position.
4. Cricothyroid (CT)	Applies longitudinal tension to vocal cords; is active during f_o changes.
5. Interarytenoid	Adducts the vocal cords; applies medial compression. May be active in setting phonation neutral position.

B. Extrinsic	
1. Sternohyoid (SH)	Lowers the hyoid *if* muscles that go from hyoid to skull and mandible are slack. Also stabilizes hyoid when muscles like digastric tense to open mandible. May be active in initiating phonation register shifts.
2. Thyrohyoid (TH)	Decreases distance between thyroid cartilage and hyoid bone.
3. Sternothyroid (ST)	Lowers the thyroid cartilage.
Pharynx	
1. Superior constrictor (SC)	Constrict the pharynx; active during swallowing and in the production of sounds like the vowel [a].
2. Medial constrictor (MC)	
2. Inferior constrictor (IC)	
2. Platopharyngeus	Constricts the pharynx; also can lower the soft palate.
Soft palate	
1. Levator palatin:	Raises soft palate, sealing nasal

Image by MIT OpenCourseWare. Adapted from Lieberman, Philip, and Sheila E. Blumstein. "Speech Physiology, Speech Perception, and Acoustic Phonetics." *Cambridge Studies in Speech Science and Communication*. New York, NY: Cambridge University Press, 1988.

	cavity in the production of oral sounds. The SC also is active in some speakers when they seal their nasal cavity.
2. Palatoglossus (PG)	Raises tongue body or lowers soft palate.
Tongue	
A. Intrinsic	
1. Superior longitudinal (SL)	Turns up the tip of tongue.
2. Inferior longitudinal (IL)	Turns down the tip of tongue.
3. Transverse MI)*	Narrows the tip of tongue.
4. Vertical (MI)*	Flattens the tip of tongue.
B. Extrinsic	
1. Genioglossus (GC)	Pulls tongue body forward; depresses the tongue body; can elevate the hyoid. Is active in production of sounds like [i] or [u], where pharynx is widened by tongue body moving forward.

Image by MIT OpenCourseWare. Adapted from Lieberman, Philip, and Sheila E. Blumstein. "Speech Physiology, Speech Perception, and Acoustic Phonetics." *Cambridge Studies in Speech Science and Communication*. New York, NY: Cambridge University Press, 1988.

2. Styloglossus	Pulls tongue body towards styloid process. Is probably active in production of sounds like [u] and velar consonants.
Suprahyoid	
1. Anterior belly of digastric (AD)	Opens the jaw if the hyoid is stabilized by tensioning muscles that connect hyoid to sternum; raises hyoid otherwise. Can be used in the production of sounds like [a].
2. Geniohyoid (GH)	Opens jaw if hyoid is stabilized; raises hyoid and pulls it forward.
3. Mylohyoid (MH)	Raises tongue body.
Mandible (lower jaw)	
1. Masseter (MAS)	Closes the jaw.
2. Temporalis (TEM)	Closes the jaw; pulls lower jaw backwards.

3. Internal pterygoid (IP)	Closes the jaw.
Lips and face	
1. Orbicularis oris (OO)	Closes the mouth; puckers the lips; acts to close and round lips in sounds like [u].
2. Depressor labii inferior (DLI)	Opens and retracts lips. Active in the release of sounds like [p] and [b].
3. Levator labii superior	Opens lips; sometimes active in release of sounds like [p] and [b].

Image by MIT OpenCourseWare. Adapted from Lieberman, Philip, and Sheila E. Blumstein. "Speech Physiology, Speech Perception, and Acoustic Phonetics." *Cambridge Studies in Speech Science and Communication*. New York, NY: Cambridge University Press, 1988.

Levels of control - example

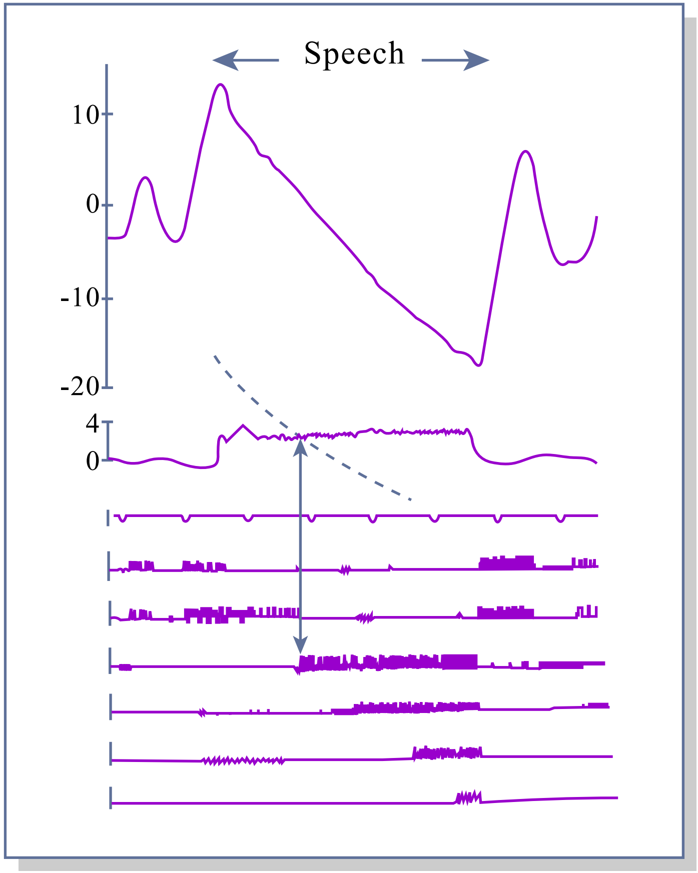
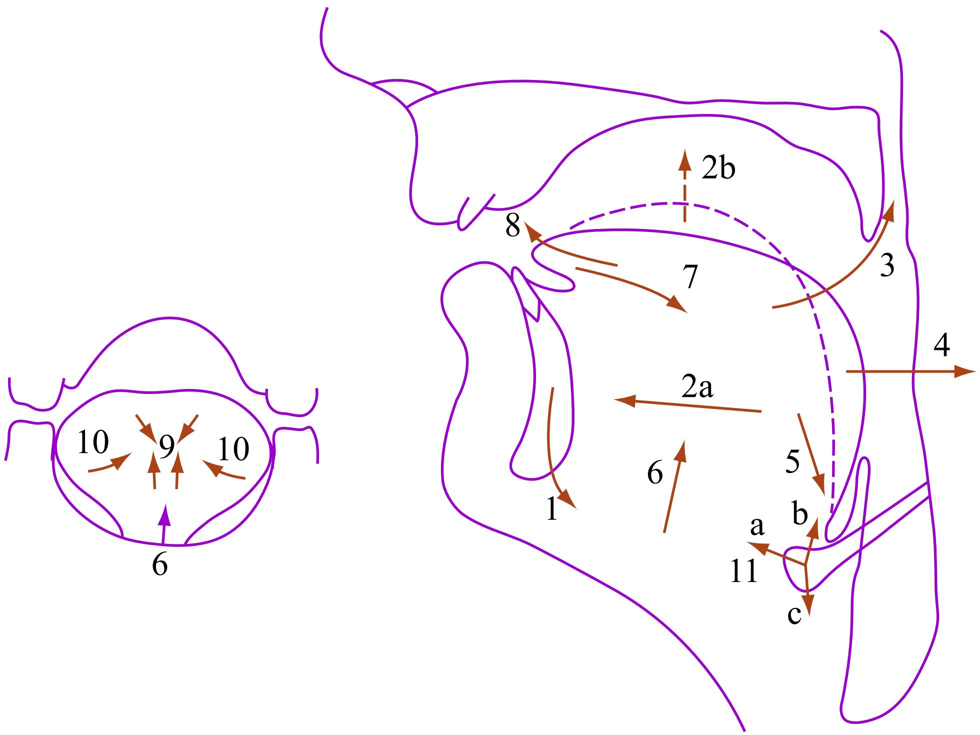


Image by MIT OpenCourseWare. Adapted from Draper, Ladefoged, and Whitteridge. J*ournal of Speech and Hearing Research* 2 (1959).

Excess degrees of freedom



The major forces affecting the position of the tongue.

(1) Jaw opening

(2a) M. Genioglossum

(2b) Affect of (2a)

(3) M. Styloglossus

(4) M. Glossopharyngeus

(5) M. Hyoglossus

(6) M. Mylohyoideus

(7) M. Longitudinalis inferior

(8) M. Longitudinals superior

(9) M. Verticalis

(10) M. Lateralis

(11) Forces affecting the hyoid:
- (a) M. Geniohyoideus
- (b) M. Stylohyoideus
- (c) M. Thyrohyoideus

Image by MIT OpenCourseWare. Adapted from Ladefoged, DeClerk, Lindau, and Papçun. "An Auditory-motor Theory of Speech Production." *UCLA Working Papers in Phonetics* 22 (1972).

A simple model of speech production: the ‘beads on a string’ model

- Idea: Speech production involves concatenating a temporal sequence of targets corresponding to phonological segments.
- Targets are vocal tract shapes.
- Speech production involves concatenating a sequence of vocal tract shapes in time, and coordinating the muscles to move between these shapes.

Coarticulation

- The influence of segmental context on the articulatory/acoustic realization of a target segment.

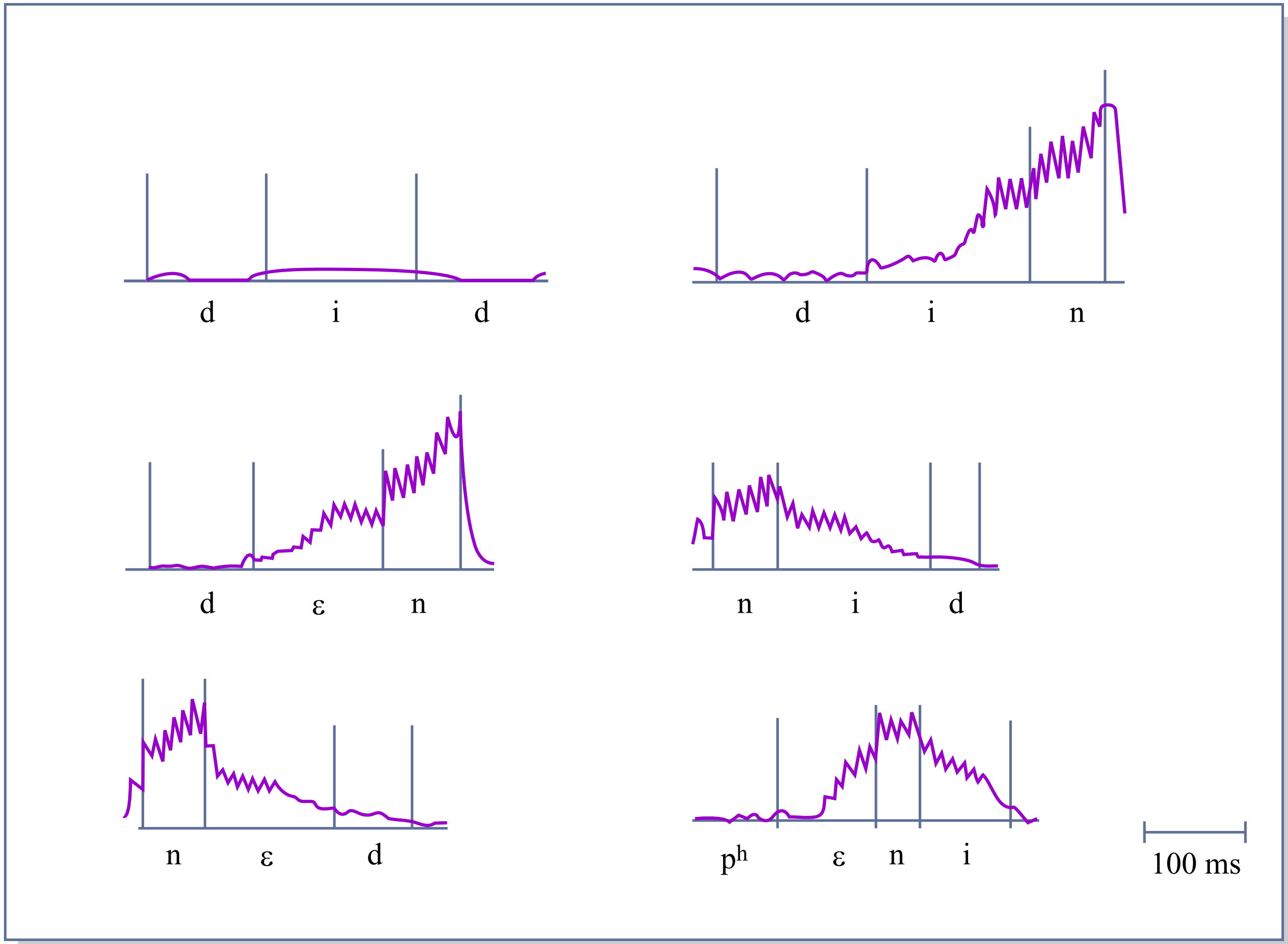


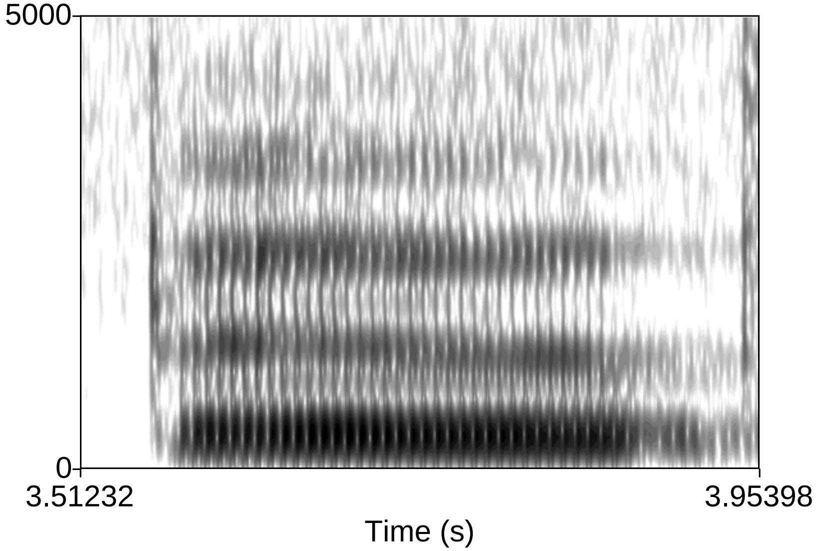
Image by MIT OpenCourseWare. Adapted from Cohn, A. "Nasalization in English: Phonology or Phonetics?" *Phonology* 10 (1993): 43-81.

Coarticulation

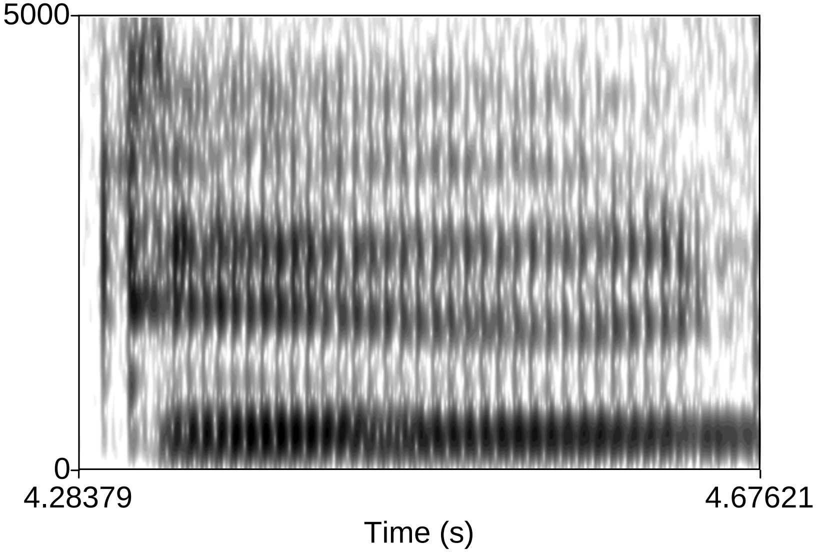
- Data on coarticulatory variation have been important in the development of models of speech production.
- We need to account for the types of influence that one segment has on another, and for the temporal extent of the influence of a segment on its neighbours.
- The simplest 'beads on a string' model leads us to expect that coarticulatory variation results solely from the transitions between segments (cf. Delattre et al's (1955) theory of acoustic loci for consonants, Liberman 1957).
- In fact coarticulation is considerably more complex than this.
  - Long range coarticulation effects.
  - Variation in targets as well as transitions.

Target variation

- Simple 'beads on a string' model implies that segment targets are invariant - variation is restricted to transitions.
- In a CV sequence, F2 at the consonant varies according to the following vowel (locus equation), and F2 in the vowel varies according to the adjacent consonants (undershoot).



[bub]



[dud]

Target variation

- In a CV sequence, F2 at the consonant varies according to the following vowel (locus equation), and F2 in the vowel varies according to the adjacent consonants (undershoot).
- Hillenbrand, Clark & Nearey 2001

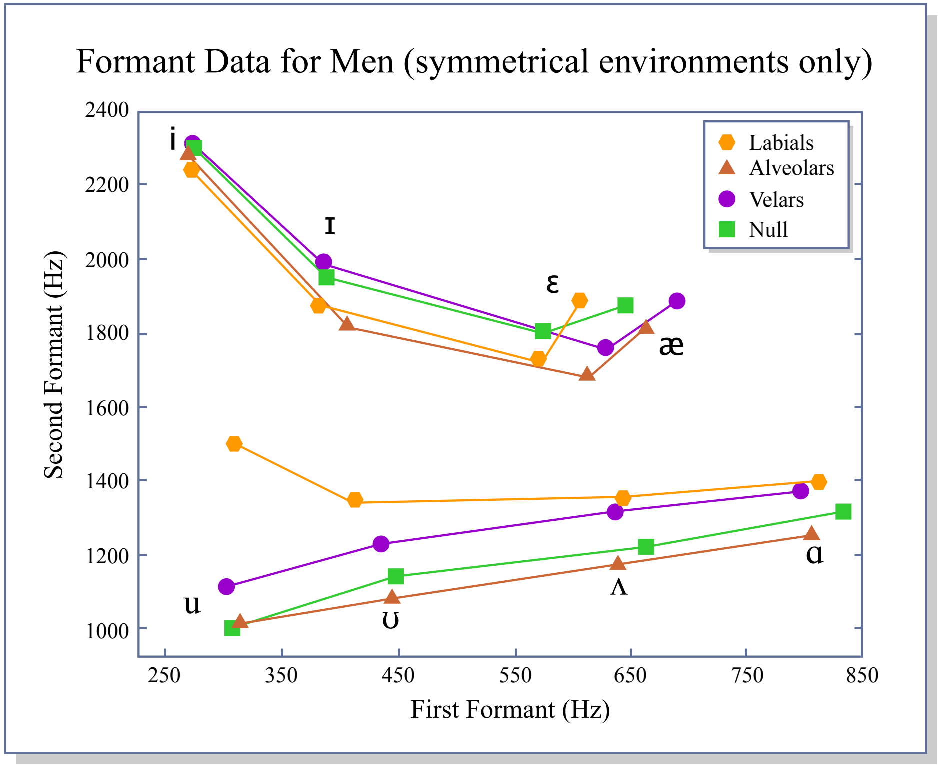


Image by MIT OpenCourseWare. Adapted from Hillenbrand, Clark, and Nearey. "Effects of Consonant Environment on Vowel Formant Patterns." *The Journal of the Acoustical Society of America* 109, no. 2 (February 2001): 748-763.

Target variation

- /t/ may be partially or wholly dental when followed by a dental fricative.
- Target variation suggests that we need a less rigid notion of a target, e.g. a range (Keating's windows) or a violable target (Lindblom 1963, Flemming 2001, Browman and Goldstein).

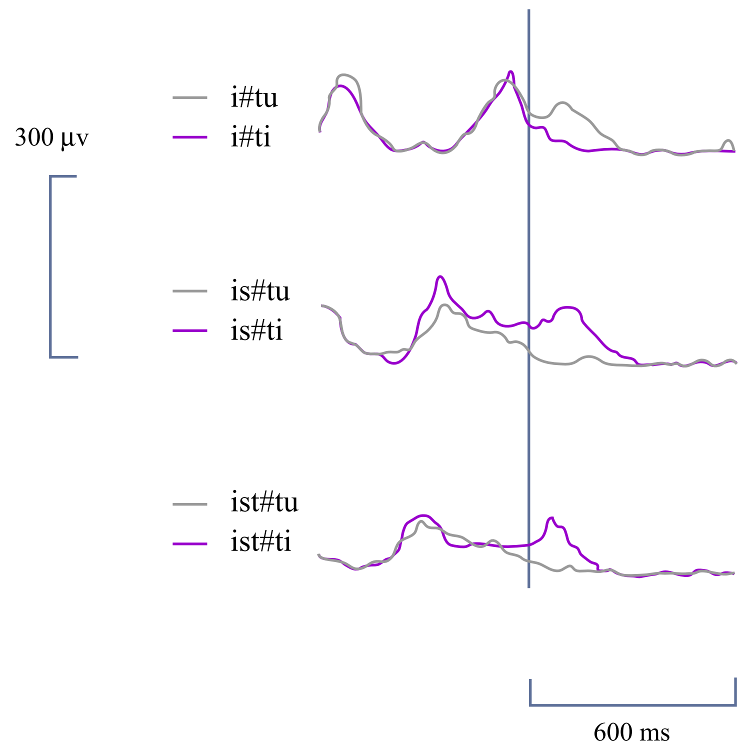
Coarticulation between non-adjacent segments

Lip-rounding: Lip-rounding for rounded vowels has been reported to begin substantially before the onset of the vowel itself:

- ‘Coarticulation of lip protrusion extends over as many as four consonants preceding the vowel /u/’ (Daniloff and Moll 1968) - e.g. [sku], [ist#tu].
- Benguerel and Cowan (1974) report coarticulation of lip-rounding across seven segments.
- Perkell (1969) reports that protrusion starts at the beginning of English nonsense words like [hətu].
- But Boyce et al (1990) found that lip-rounding precedes rounded vowel onset by a relatively fixed duration.

Coarticulation between non-adjacent segments

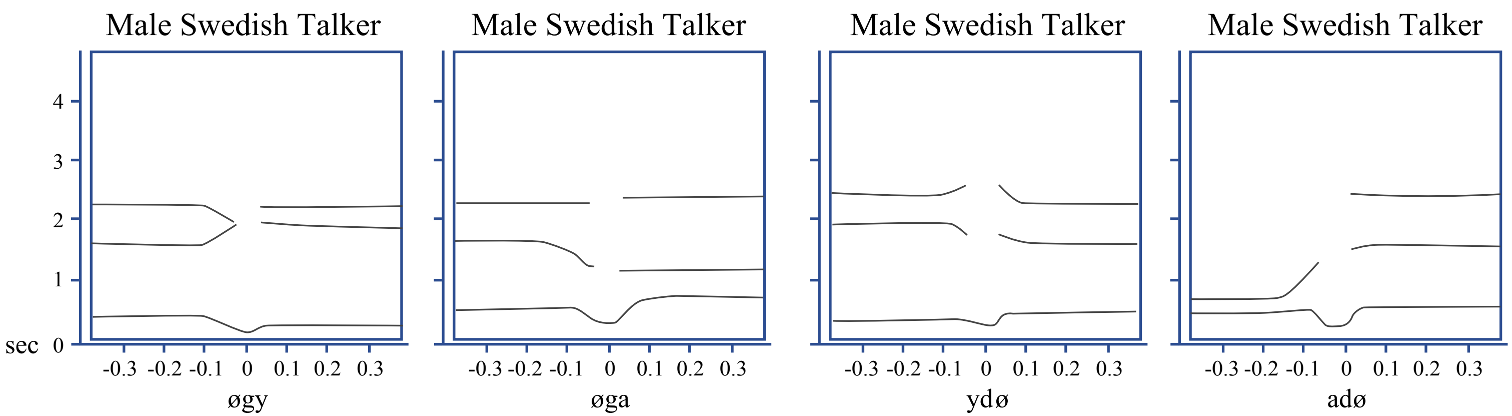
- Boyce et al (1990)



Ensemble-averaged EMG activity for 15 to 20 tokens of 3 pairs of utterances with increasing numbers of “neutral” consonants. The vertical mark indicates the lineup point for averaging tokens, at the acoustic onset of the second vowel.

Coarticulation between non-adjacent segments

- Coarticulation between vowels across intervening consonants has been well-known since Öhman (1966).



Formant transitions in both the initial (first two graphs) and final (latter two graphs) vowel are different in the two cases, due to the influence of one on the other.

Image by MIT OpenCourseWare. Adapted from Öhman, S. E. G. "Coarticulation in VCV Utterances: Spectrographic Measurements." *Journal of the Acoustical Society of America* 39 (1966): 151–168.

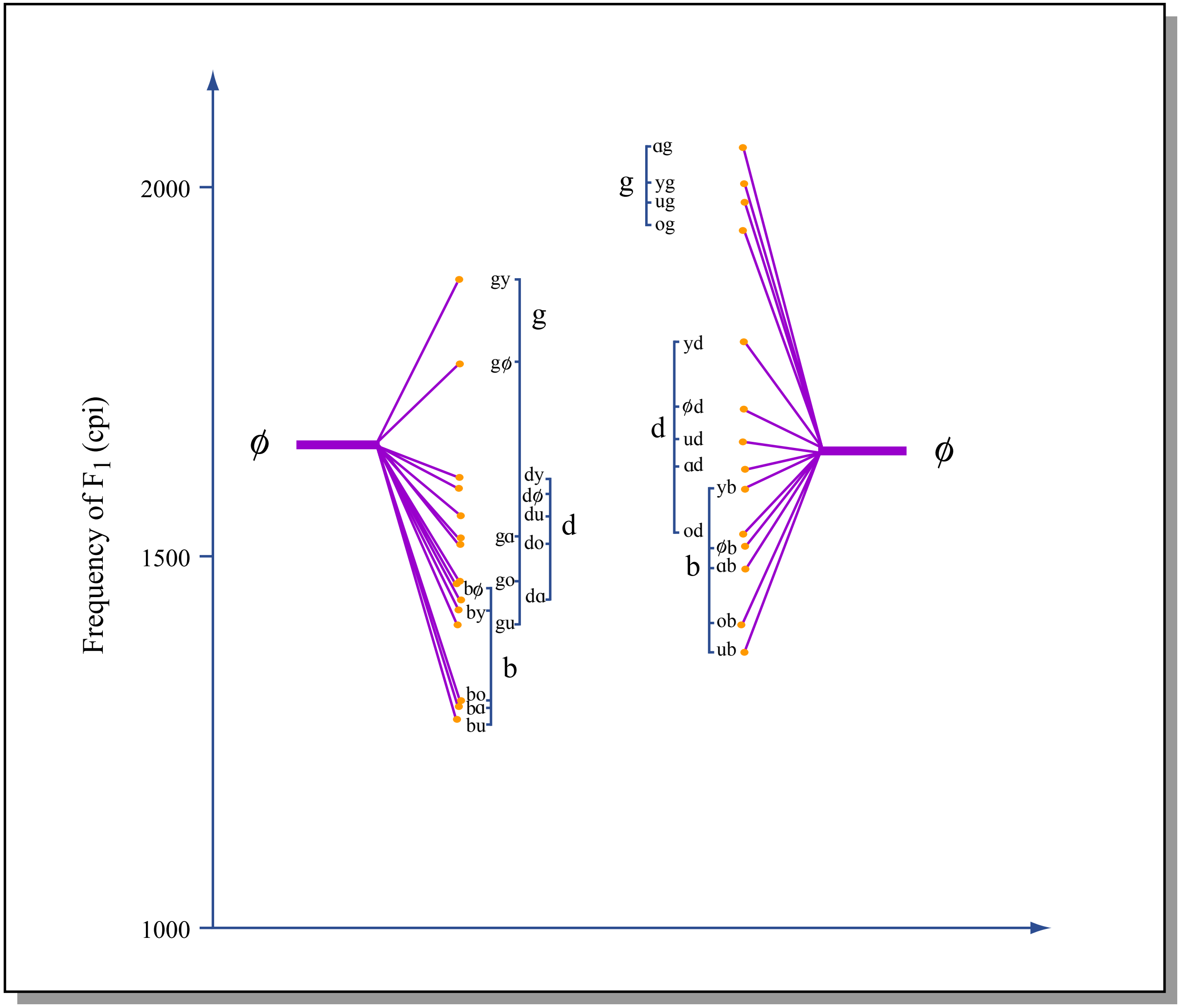


Image by MIT OpenCourseWare. Adapted from Öhman, S. E. G. "Coarticulation in VCV Utterances: Spectrographic Measurements." *Journal of the Acoustical Society of America* 39 (1966): 151–168.

Coarticulation between non-adjacent segments

- Öhman (1966)

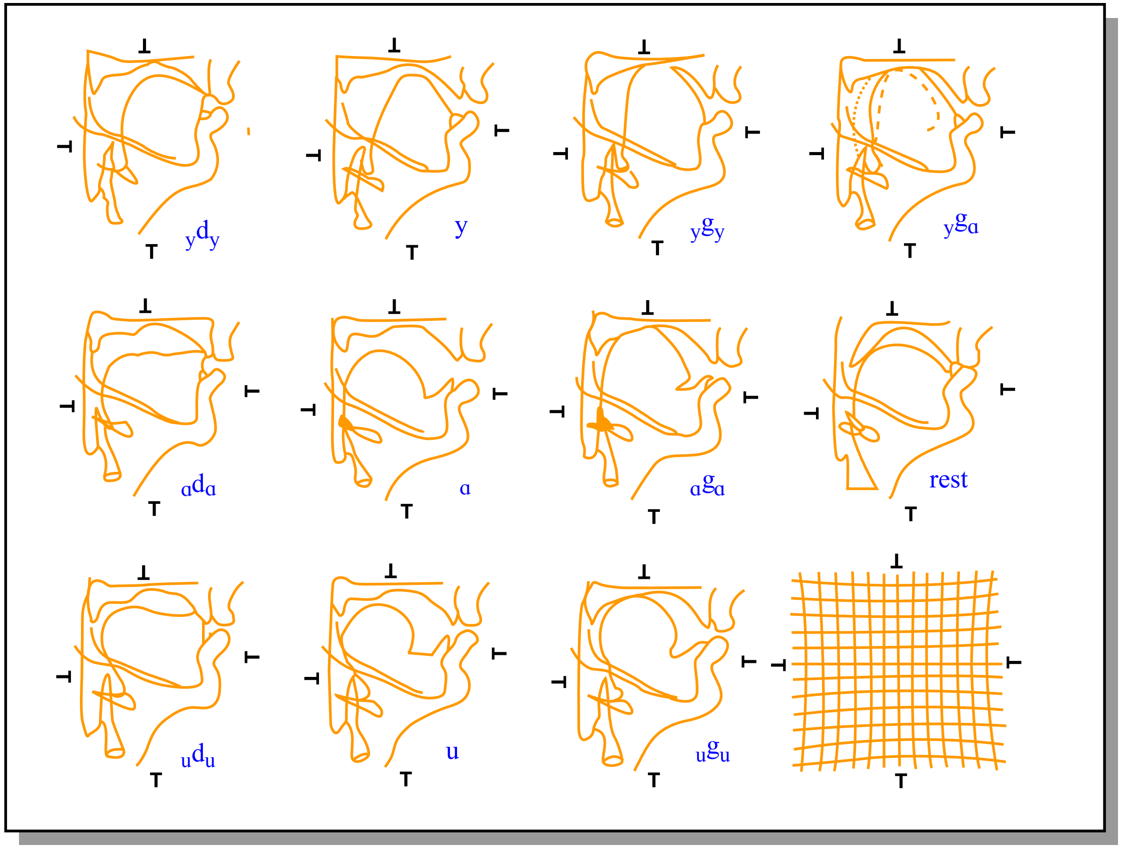


Image by MIT OpenCourseWare. Adapted from Öhman, S. E. G. "Coarticulation in VCV Utterances: Spectrographic Measurements." *Journal of the Acoustical Society of America* 39 (1966): 151–168.

Models of coarticulation

- Articulatory Phonology (Browman and Goldstein)
- Window Model (Keating)

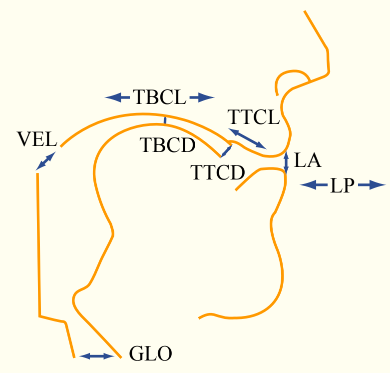
Articulatory Phonology

- Theory developed by Browman and Goldstein (1986, 1987, 1989 etc).
- Not a theory of phonology.
- The basic unit of articulatory control is the **gesture**.
- A gesture specifies the formation of a linguistically significant constriction.
- Defined within the framework of Task Dynamics (Saltzmann and Munhall 1989).

Articulatory Phonology

- A gesture specifies the formation of a linguistically significant constriction.
- The goals of gestures are defined in terms of tract variables (e.g. lip aperture).
- Movement towards a particular value of a tract variable is typically achieved by a set of articulators.
- A gesture takes a tract variable from its current value towards the target value.

Tract variable		Articulators involved
LP	lip protrusion	upper and lower lips, jaw
LA	lip aperture	upper and lower lips, jaw
TTCL	tongue-tip constriction location	tongue-tip, tongue-body, jaw
TTCD	tongue-tip constriction degree	tongue-tip, tongue-body, jaw
TBCL	tongue-body constriction location	tongue-body, jaw
TBCD	tongue-body constriction degree	tongue-body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottis



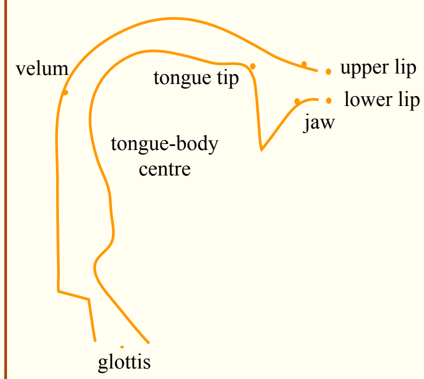


Image by MIT OpenCourseWare. Adapted from Browman, and Goldstein. *Journal of Phonetics* 18 (1990): 299-320.

Articulatory Phonology

- Since a gesture involves the formation of a constriction it is usually specified by:
  - constriction degree
  - (constriction location)
  - (constriction shape)
  - stiffness
- In the Task Dynamic model, movement along a tract variable is modeled as a spring-mass system.
- In Browman and Goldstein's model critical damping is assumed, so articulators move towards the target position on the tract variable in a non-linear, assymptoting motion.

Articulatory Phonology

- Gestures are coordinated together to produce utterances (represented in the 'gestural score' format).

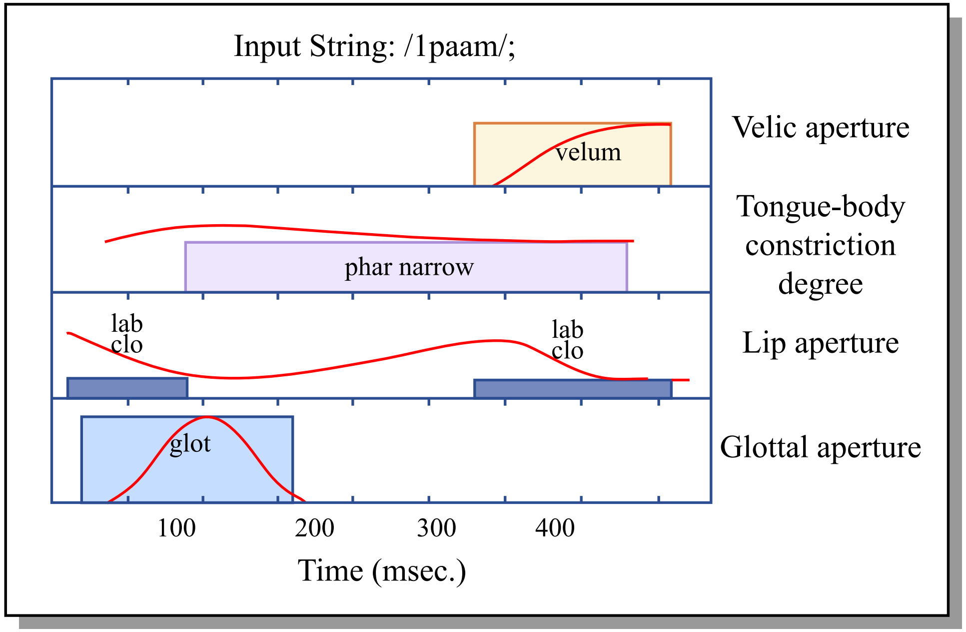


Image by MIT OpenCourseWare. Adapted from Browman, and Goldstein. *Journal of Phonetics* 18 (1990): 299-320.

Articulatory Phonology

- This model tackles the ‘degrees of freedom’ problem: articulator movements are derived from control of a limited set of tract variables and stiffness parameters.
- Gestures specify dynamic movements, but are defined in terms of static parameters.

Modeling coarticulation

- Overlap is the basic mechanism for modeling coarticulation - coarticulation as coproduction (Fowler 1980).
  - E.g. vowel gestures will typically overlap with consonant gestures.
- When two gestures involve the same tract variables (e.g. vowels and velars, two vowels), blending results (a compromise between the demands of the two simultaneously active gestures).
- Coarticulatory effects will also result from the fact that gestures specify movement from the current location to form a particular constriction, so the articulator movements resulting from a given gesture will depend on the initial state of the articulators.

Modeling coarticulation

- Note that the degree of underspecification in gestural scores is probably exaggerated in the examples given in the papers (cf. Mattingly 1990).
- E.g. rounded vowels involve larynx lowering, velum position varies with vowel height.
- Casual speech - Browman and Goldstein argue that gestural overlap also provides a good account of casual speech processes.

Keating’s Window Model

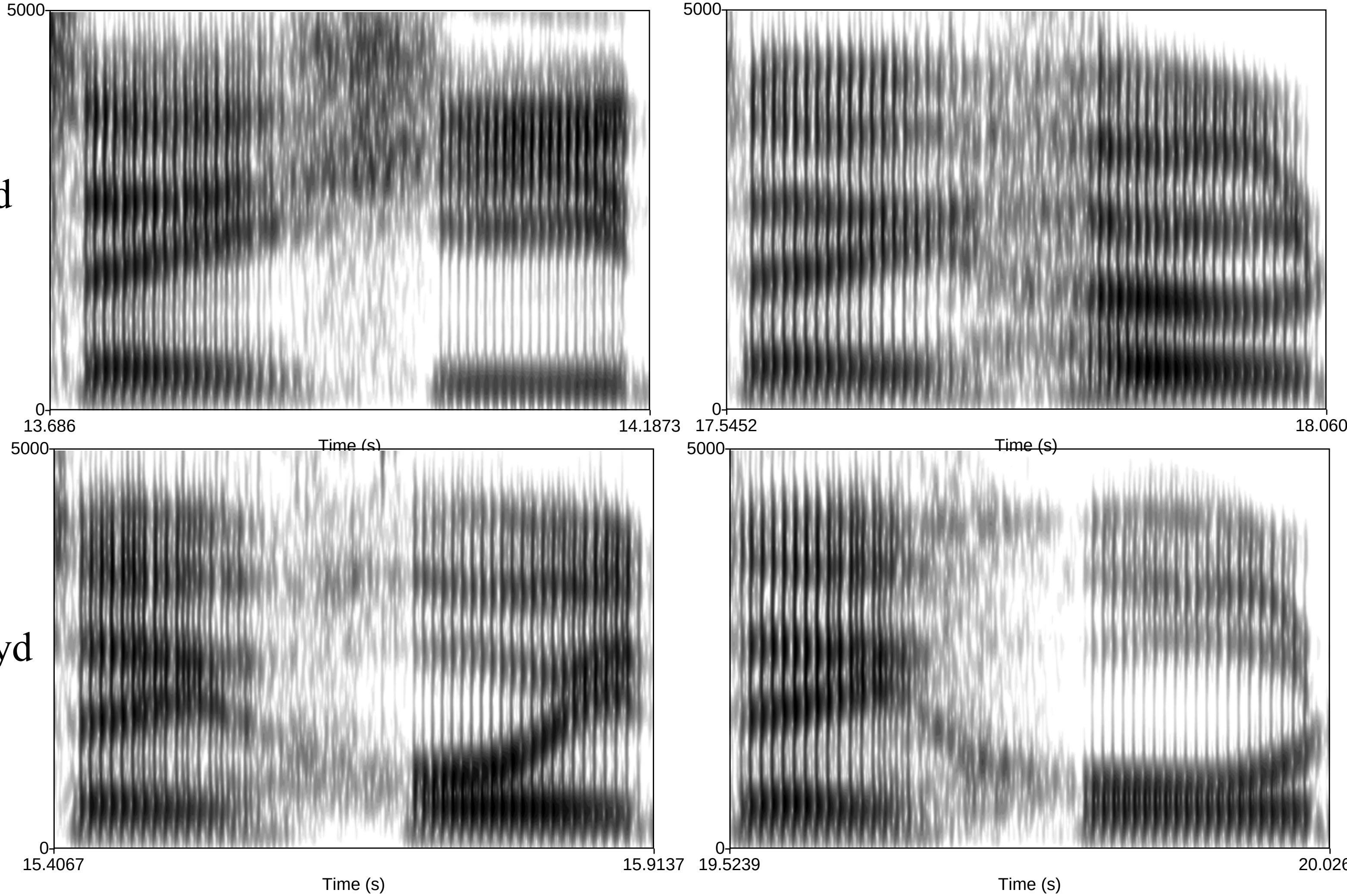
- Keating’s window model of coarticulation is a development of the ‘targets and interpolation’ approach to speech production:
- Segments can be specified for targets on a number of parameters (e.g. velum height, jaw height).
- Segments need not have targets on all parameters (underspecification).
- Targets for a given segment need not be simultaneous.

Keating (1988)

- Example of underspecification: Argues that [h] lacks specifications for oral features, based on data like the following:

heed

hoed



Hoyd

hoard

Keating's (1990) 'Windows' model

- Phonetic underspecification á la Keating (1988) allows only inviolable targets on a parameter, or no target at all (freely variable).
- Keating (1990) argues that this is too simplistic - targets may vary in degree of specificity.
- Implemented by replacing point targets with 'windows' specifying a range of acceptable values on a parameter.

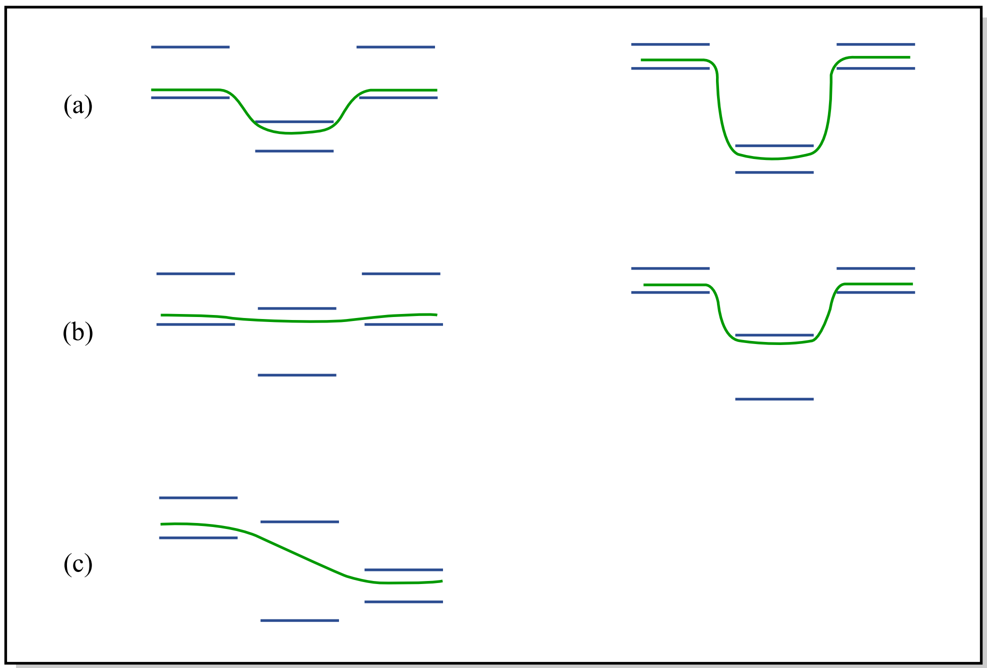


Image by MIT OpenCourseWare. Adapted from Keating, P. A. "The Window Model of Coarticulation: Articulatory Evidence." In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Edited by John Kingston and Mary E. Beckman. New York, NY: Cambridge University Press, 1990, pp. 451-470. ISBN: 9780521368087.

Keating's (1990) 'Windows' model

- Motivated by evidence for segments that are exhibit substantial, but bounded, contextual variability on a parameter. E.g. velum height in English vowels:

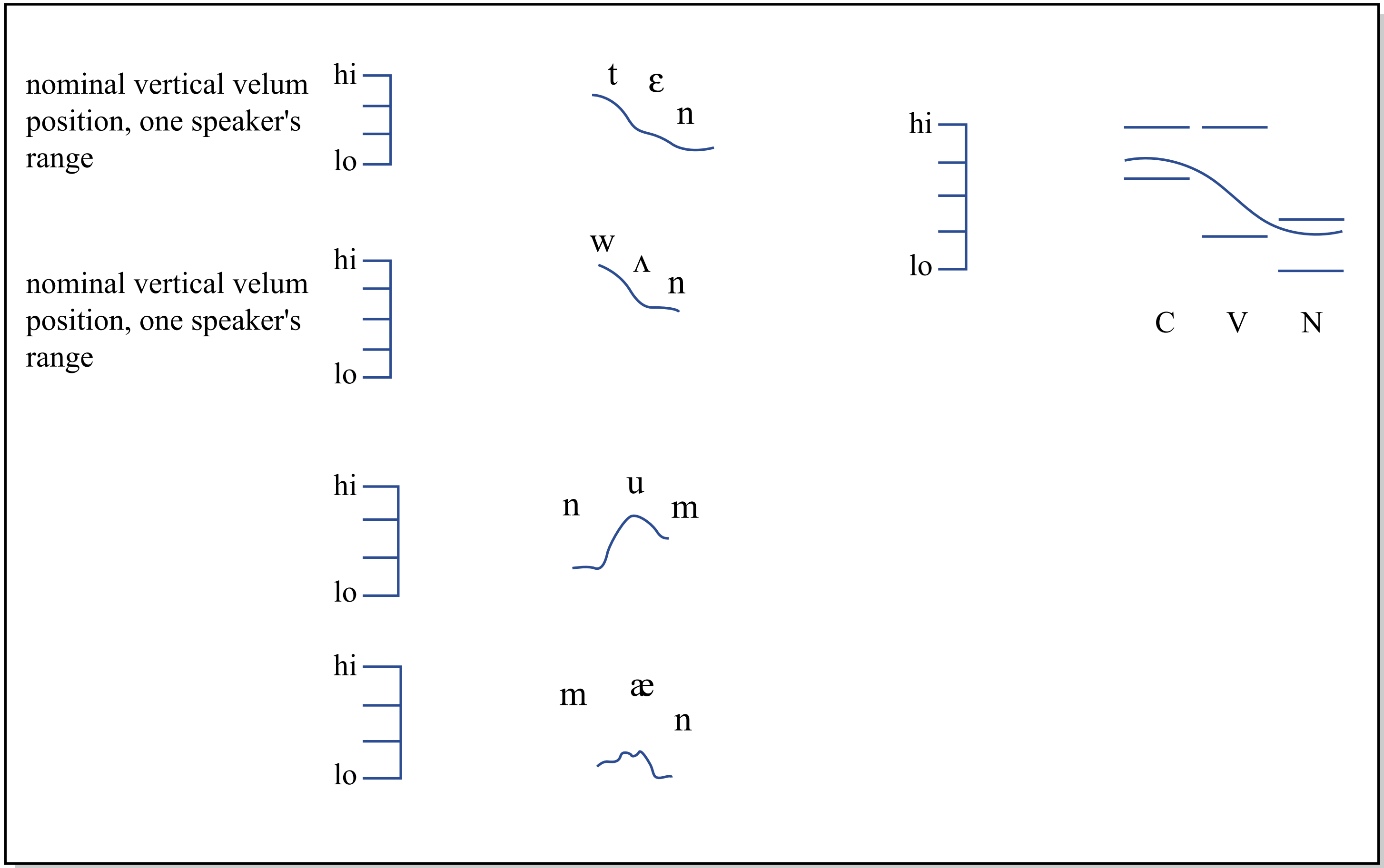


Image by MIT OpenCourseWare. Adapted from Keating, P. A. "The Window Model of Coarticulation: Articulatory Evidence." In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Edited by John Kingston and Mary E. Beckman. New York, NY: Cambridge University Press, 1990, pp. 451-470. ISBN: 9780521368087.

Distinguishing the models: Duration of anticipatory coarticulation

- Gestural/coproduction models tend to predict fixed duration of anticipatory coarticulation relative to onset of segment, whereas targets-and-interpolation models tend to predict interpolation across any number of unspecified segments.
- Studies have yielded mixed results (e.g. above).
- Perkell and Matthies (1992) argue for a ‘hybrid’ model on the basis of data on anticpatory labial coarticulation in English.

Distinguishing the models: Duration of anticipatory coarticulation

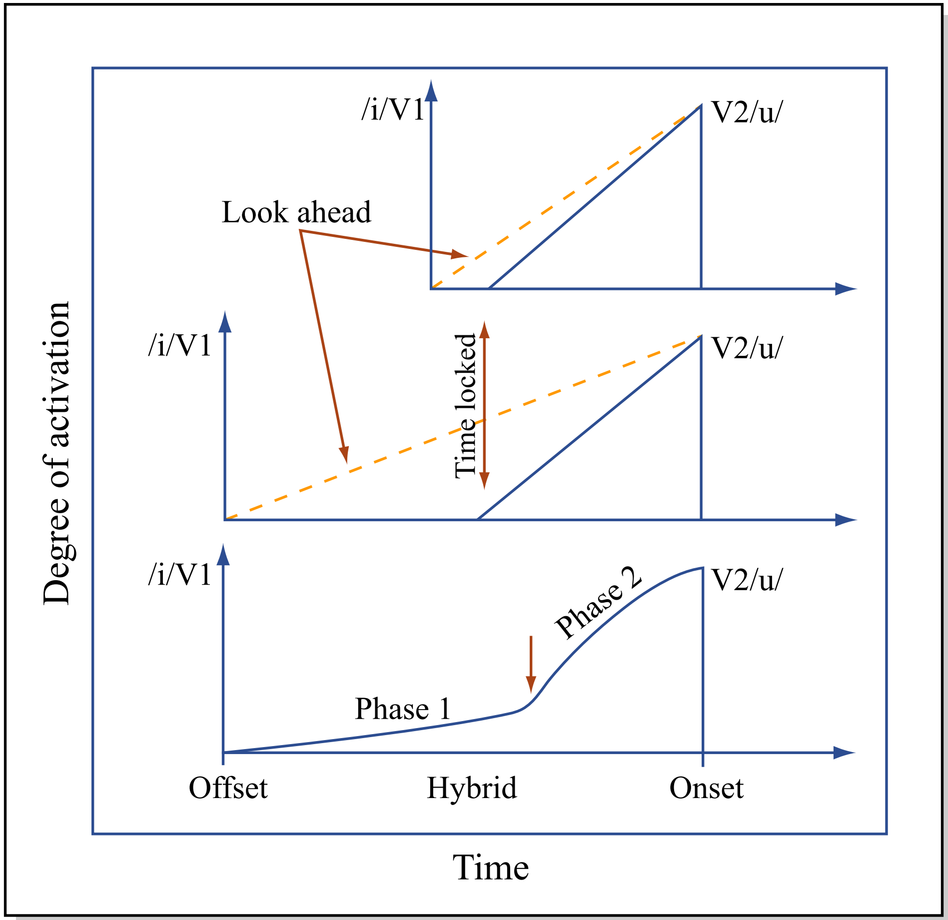


Image by MIT OpenCourseWare. Adapted from Perkell, J. S., and Matthies, L. M. "Temporal Measures of Anticipatory Labial Coarticulation for the Vowel /u/: Within- and Cross-Subject Variability." *The Journal of the Acoustical Society of America* 91, no. 5 (1992): 2911-2925.

Distinguishing the models: Duration of anticipatory coarticulation

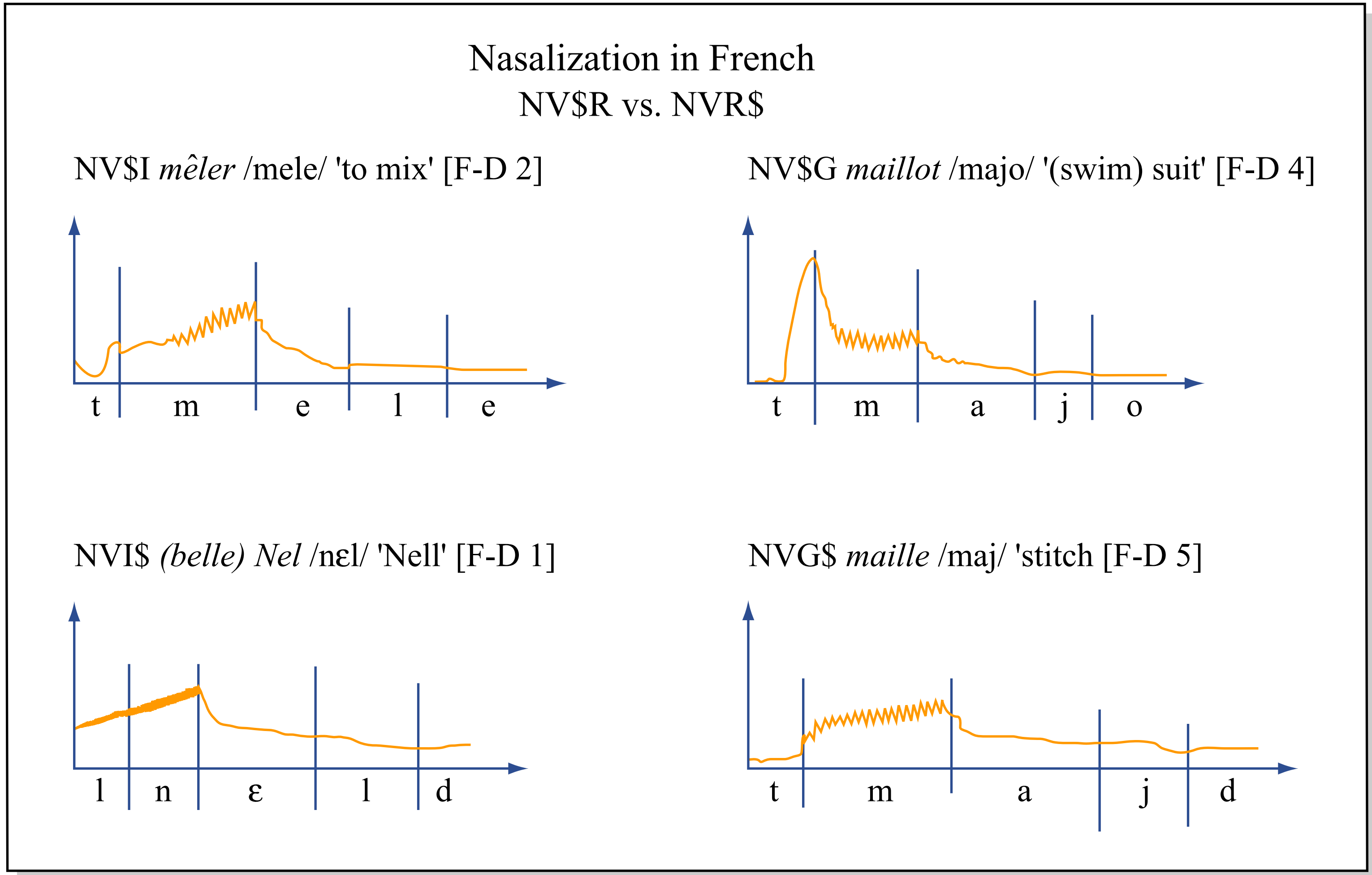
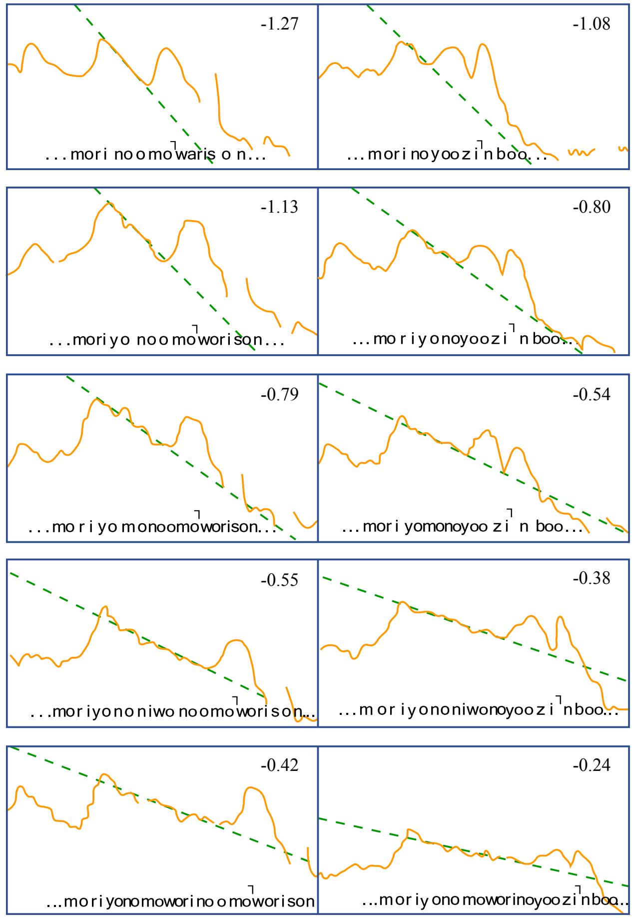


Image by MIT OpenCourseWare. Adapted from Cohn, Abigail. "Phonetic and Phonological Rules of Nasalization." Ph.D. dissertation, University of California, Los Angeles, 1990.

Pierrehumbert and Beckman (1988) - Fundamental frequency in Tokyo Japanese.



Unaccented phrases with 3, 4, 5, 7, and 8 morae before accentual-phrase boundaries with two different allophones of L%. The dashed line in each panel is a regression line fitted to all f_0 values between the peak for the phrasal **H** in the first phrase and the minimum for the interphrasal L%. The number in the upper right-hand corner is the slope of the regression line. Here a right corner in the transcription indicates the location of an accent.

References

- Benguerel, A.P. and H.A. Cowan (1974) Coarticulation of upper lip protrusion in French. *Phonetica* 30, 41-55.
- Boyce, S.E., R.A. Krakow, F. Bell-Berti and C.E. Gelfer (1990) Converging sources of evidence for dissecting articulatory movements into core gestures. *Journal of Phonetics* 18, 173-188.
- Browman, C. and L. Goldstein (1986). Toward an articulatory phonology. *Phonology yearbook* 3, 219-252.
- Browman, C. and L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6, 201-252.
- Cohn, A. (1990) *Phonetic and Phonological Rules of Nasalization*. UCLA Working Papers in Phonetics 76. (Ph.D. dissertation, UCLA).
- Daniloff and Moll (1968). Coarticulation of lip rounding. *Journal of Speech and Hearing Research* 11, 707-721.
- Delattre, Liberman and Cooper (1955). Acoustic loci and transitional cues for consonants. *JASA* 27, 769-773.
- Dickson, D.R. and W. Maue-Dickson (1982). *Anatomical and Physiological Bases of Speech*. Little, Brown and Co., Boston.
- Hardcastle, W.J. (1976). *Physiology of Speech Production*. Academic Press, London.
- Keating, P. (1988). Underspecification in Phonetics. *Phonology* 5.

References

- Liberman, A.M. (1957). Some results of research on speech perception. *JASA* 29, 117-23.
- Lieberman, P. and S. Blumstein (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics*. CUP, Cambridge.
- Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements.
- Perkell, J. (1969). *Physiology of speech production*. MIT Press, Cambridge.
- UCLA Phonetics Lab. *Dissection of the Speech Production Mechanism*. *UCLA Working Papers in Phonetics* 77.
- Zemlin, W. 91981). *Speech and Hearing Science: Anatomy and Physiology*. 2nd edn. Prentice-Hall, Englewood Cliffs.