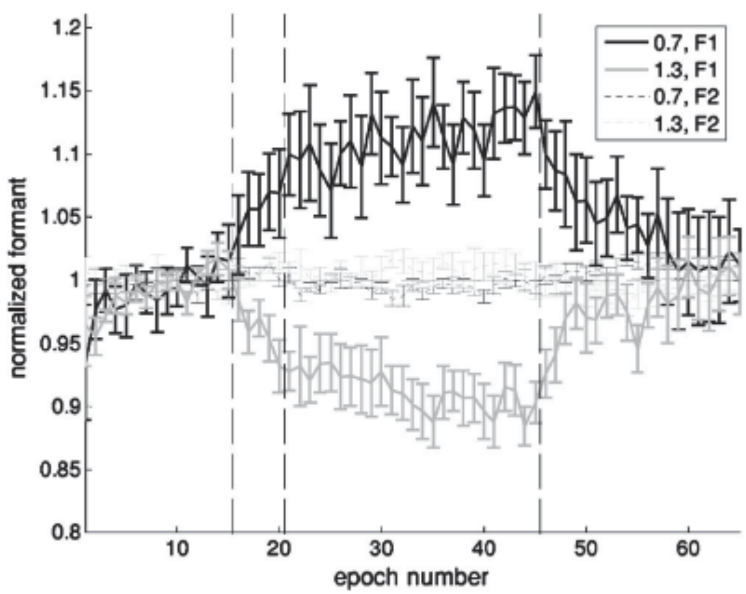
24.915/24.963 2

Linguistic Phonetics

Targets in speech production



- Reading: Reading: Mitterer, H. (2006). On the causes of compensation for coarticulation: Evidence for phonological mediation. *Perception & Psychophysics*, *68*, 1227-1240.
- Assignment: Write up a report on our experiment studying voicing contrasts among English affricates (due 12/1).
  - Project is due 12/11

Targets in Speech Production

- What are the control parameters in speech production?
  - muscle tensions
  - lengths and shortening velocities of muscles
  - vocal tract shape
  - acoustic/perceptual properties
  - all of the above?
- Ultimately speech production involves control of individual muscles, so in what sense can the control parameters lie at any other level?
  - Positing control parameters at a particular level helps to make sense of variability in the production of the same unit across contexts, or utterance to utterance.

Example

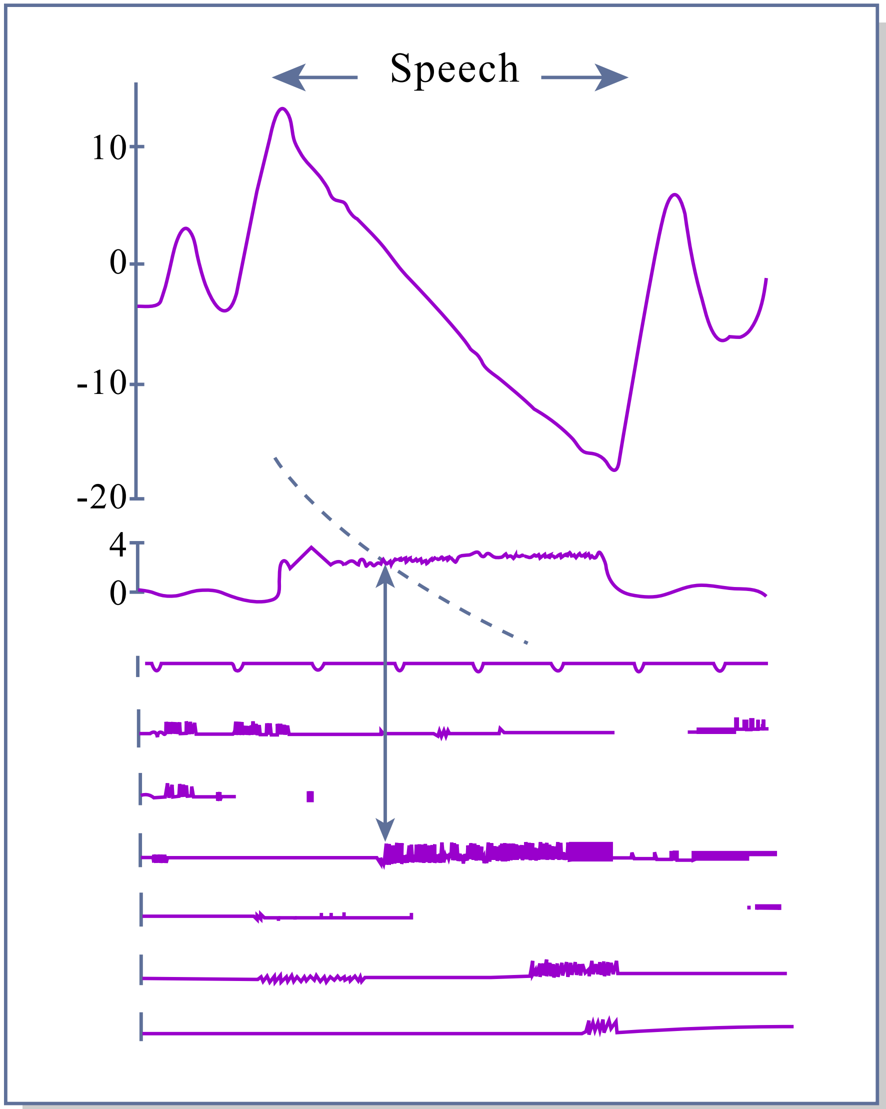


FIGURE 2. Upper part of figure, a reproduction of a record of the variations in the volume of air in the lung and the oesophageal pressure during respiration and speech (counting from one to 32 at a conversational loudness). Lower part of figure, a diagrammatic representation of the muscular activity which was observed to accompany such pressure and volume changes. The dashed line which has been superimposed on the pressure record indicates the relaxation pressure associated with the corresponding volume of air in the lungs. It is equal to zero when the amount of air in the lungs is the same as that at the end of a normal breath. The arrows indicate the moment when the relaxation pressure is no longer greater than the mean pressure below the vocal cords. At this moment the external intercostal activity ceases, and that of the internal intercostals commences.

Draper, Ladefoged and Whitteridge 1959

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

Variability in articulator positions - bite block experiments

- Speakers can produce normal vowels with a bite block held between the teeth, preventing normal jaw movement (Lindblom et al 1979).
  - Vowel is normal from onset of speech.
- Although jaw-raising normally contributes to labial closure, speakers can produce /p, b/ with a bite block fixing the jaw in a low position (Lindblom et al 1987).

➢ Targets are not articulator positions.

➢ Could be:

  - Constriction locations (e.g. Articulatory Phonology).
  - Acoustic/perceptual targets.

Feedback, feedforward and internal models

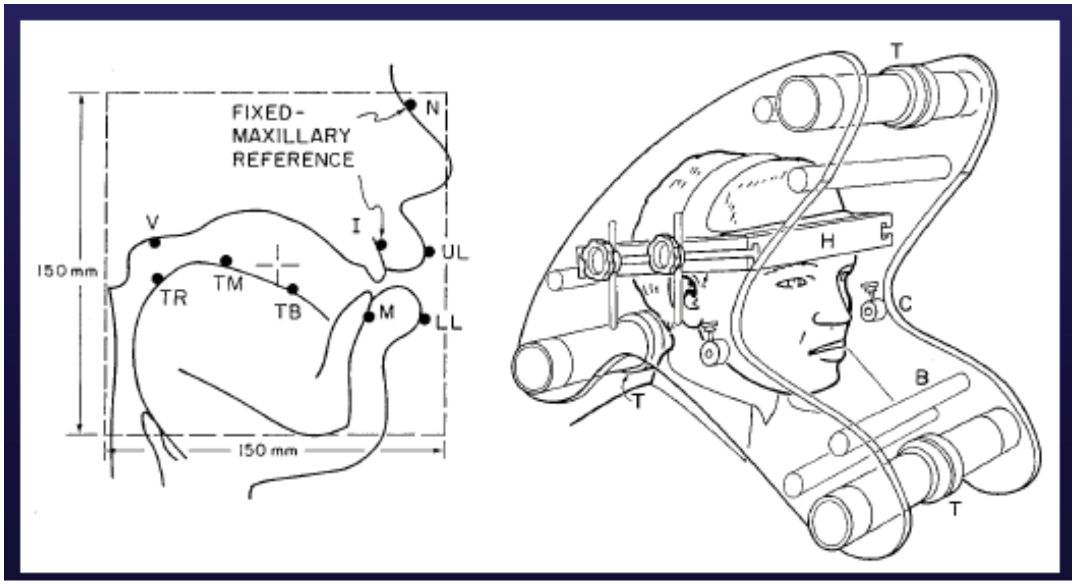
- A problem for perceptual targets: they are insufficient to regulate on-line speech production because feedback via the auditory system would be too slow to account for rapid compensation for perturbations (e.g. bite blocks, jaw loading).
  - Auditory processing+correction to motor commands+delay between muscle activation and generation of tension 'would probably be longer than some brief speech movements' (Perkell et al 1997).
- Solution: Speakers have an internal model of the relation between motor commands, orosensory feedback, and acoustic/perceptual consequences (Guenther 1992, Jordan 1992, etc).
- Internal model is used to determine motor commands on the basis of perceptual targets.
  - flexible - can compensate for bite blocks, etc.
- Is used to predict auditory consequences of current articulations, based on current motor commands and orosensory feedback concerning current state of the vocal tract (e.g. Guenther & Perkell 2001)
  - This information can be used for rapid error correction.

Evidence for acoustic/perceptual targets

- According to this model, targets are fundamentally perceptual, but are used to construct motor plans based on an internal model.
- What is the evidence for acoustic targets?
  - Argued against articulator positions as targets based on 'motor equivalence' - many combinations of articulator movements are used to achieve the same linguistically relevant goal - e.g. closing the lips.
  - One type of argument for acoustic targets is based on pushing the idea of motor equivalence into the acoustic/perceptual domain: variation in the articulation of the 'same sound' can be understood as achieving equivalence at a perceptual level, not at an articulatory level.
  - Specifically: compensatory articulations - two independent articulators are traded off to achieve a constant acoustic target.

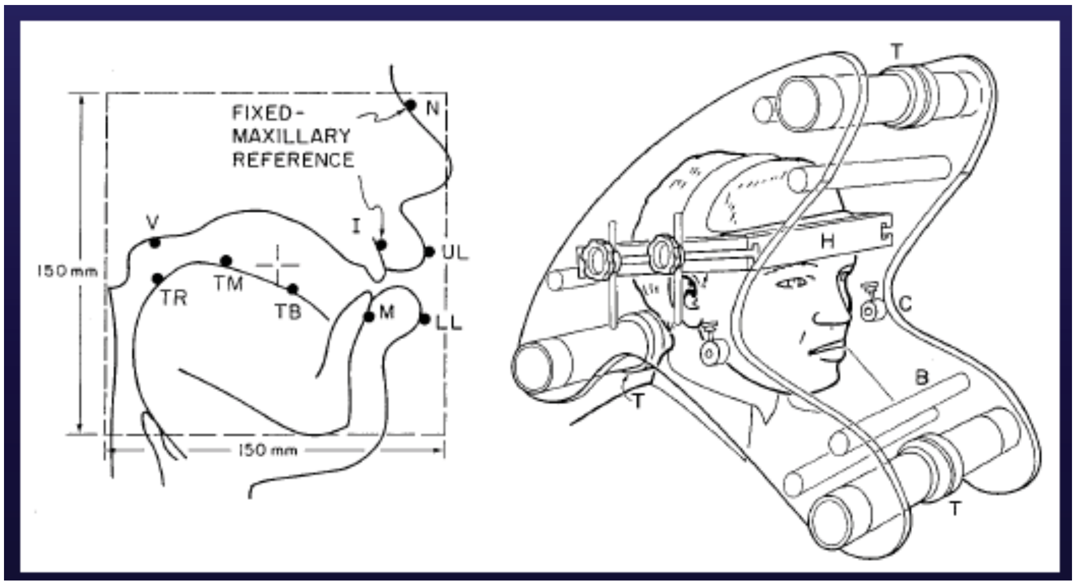
Evidence for acoustic targets - motor equivalence

- E.g. target for [u] is low F2.
  - Low F2 is achieved by a combination of tongue body backing and lip-rounding.
  - Smaller lip constriction could compensate for less tongue body backing, and vice versa, in achieving a particular low F2 target.
  - If speakers exploit this motor equivalence, it would suggest that their target is fundamentally acoustic, since lip rounding and tongue body position are not linked articulatorily.



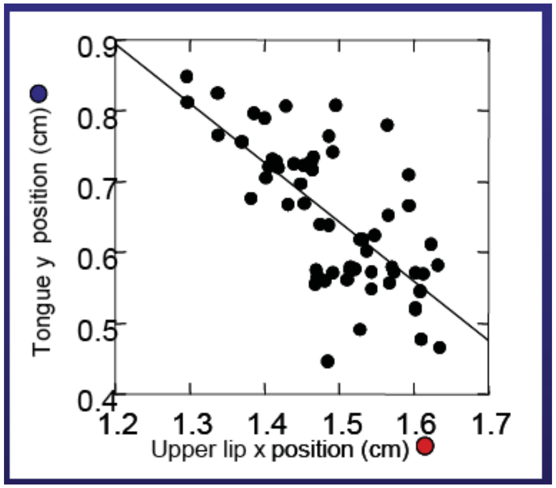
Evidence for acoustic targets - motor equivalence

- Perkell et al (1993): tracked articulator movements using an Electro-Magnetic Midsagittal Articulometer (EMMA) system.
- Three transmitter coils positioned around the speaker's head emit alternating magnetic fields.
- Induce electric currents in transducer coils glued to articulatory structures (tongue, lips, etc).
- Calculate position of transducers based on induced currents.



Perkell et al (1993)

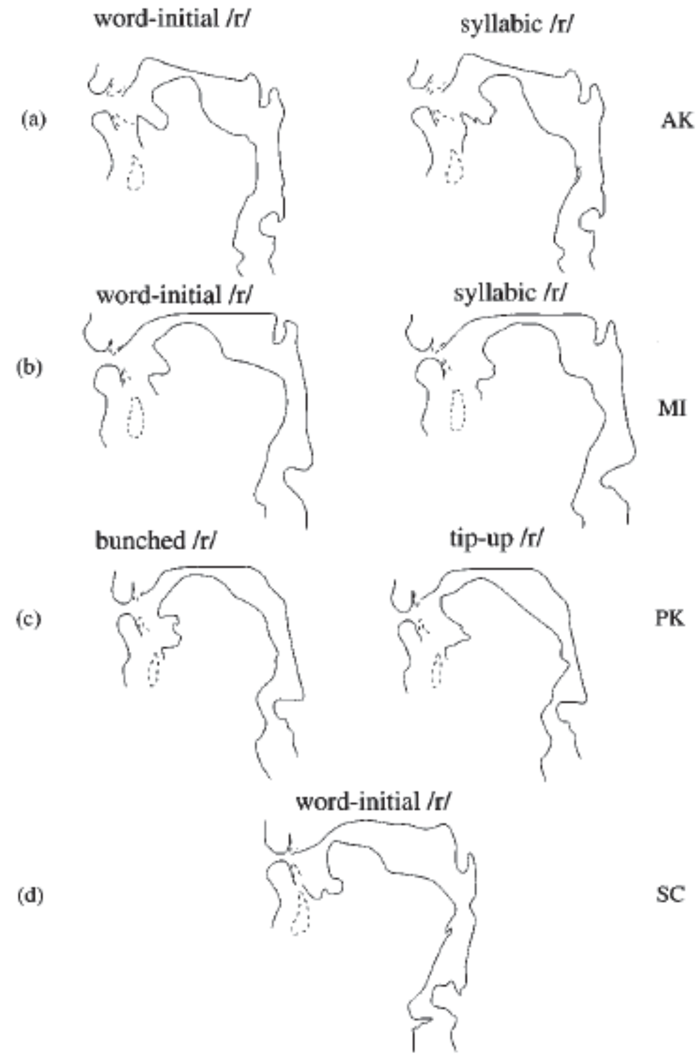
- Tracked articulator positions during production of /hu/ *who* in a variety of contexts.
- Predict inverse correlation between tongue body position and lip constriction.
- Weak inverse correlations for 3/4 subjects.
- Variation in clarity of speech could result in correlated variation in tongue body position and lip constriction.



Reproduced from Perkell, Joseph S., Melanie L. Matthies, Mario A. Svirsky, and Michael I. Jordan. "Trading relations between tongue - body raising and lip rounding in production of the vowel/u: A pilot "motor equivalence"study." The Journal of the Acoustical Society of America 93, no. 5 (1993): 2948-2961. with the permission of the Acoustical Society of America.

Guenther et al (1999) - English [ɹ]

- American English [ɹ] is produced in a variety of ways - ‘buched’, ‘retroflex’, ‘tip up’ etc.
- All of these articulatory configurations produce extremely low F3 (<2000 Hz).
- F3 is (generally) a front cavity resonance
- Three ways of lowering F3:
  - lengthen front cavity
  - lengthen the constriction
  - decrease the cross-sectional area of 2 the constriction
  - (also: constriction at lips).
- Guenther et al (1999) provide evidence for trade-offs between these strategies across contexts.
  - /warav, wabrav, wadrav, wagrav, wavrav/



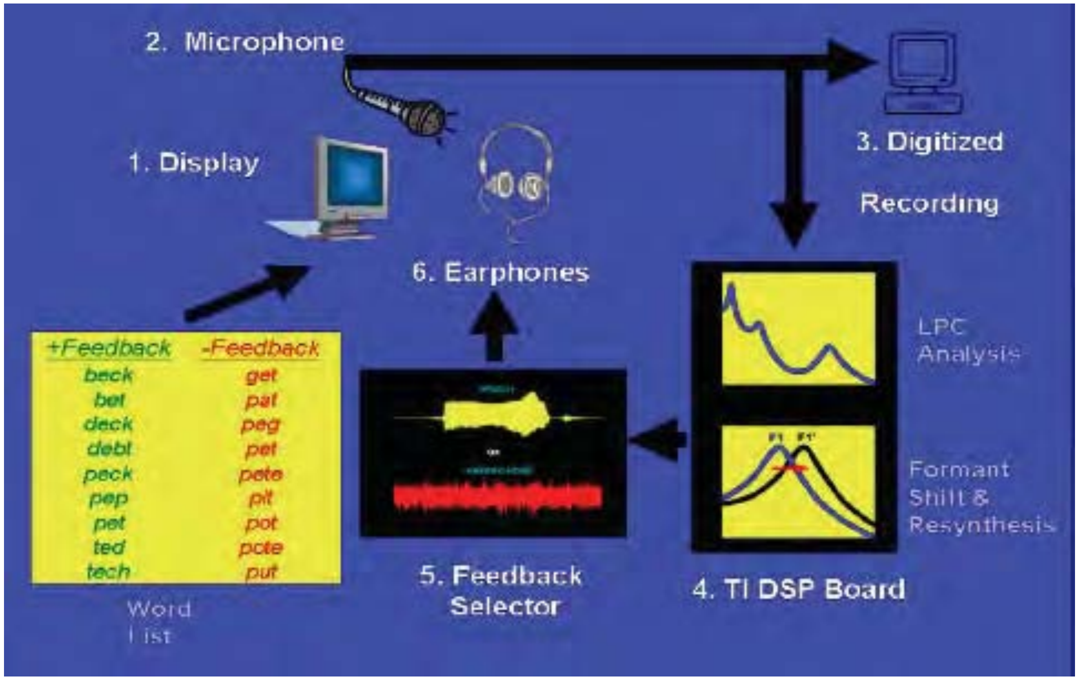
Reproduced from Narayanan, Shrikanth S., Abeer A. Alwan, and Katherine Haker. "Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals." The Journal of the Acoustical Society of America 101, no. 2 (1997): 1064-1077. doi: https://doi.org/10.1121/1.418030, with the permission of the Acoustical Society of America.

The role of auditory feedback

- Auditory feedback is required to learn internal models of the relations between articulation and acoustics.
- Auditory feedback is also required to maintain internal models - speech of post-lingually deafened adults deteriorates.
  - Production and perception improve together after cochlear implant surgery (Perkell 2012).
- Evidence that auditory feedback is used for rapid adaptation of internal models: modified auditory feedback (Villacorta et al 2007).

The role of auditory feedback - Villacorta et al

- Real time modification of auditory feedback:
  - Process subject’s speech to shift F1,
  - Play modified speech back to subject through headphones.
  - Some words masked by noise (no feedback)
  - Feedback only on [ε], other vowels tested (with noise) for generalization.



Source: Villacorta, Virgilio M., Joseph S. Perkell, and Frank H. Guenther. "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception." The Journal of the Acoustical Society of America 122, no. 4 (2007): 2306-2319.

The role of auditory feedback - Villacorta et al

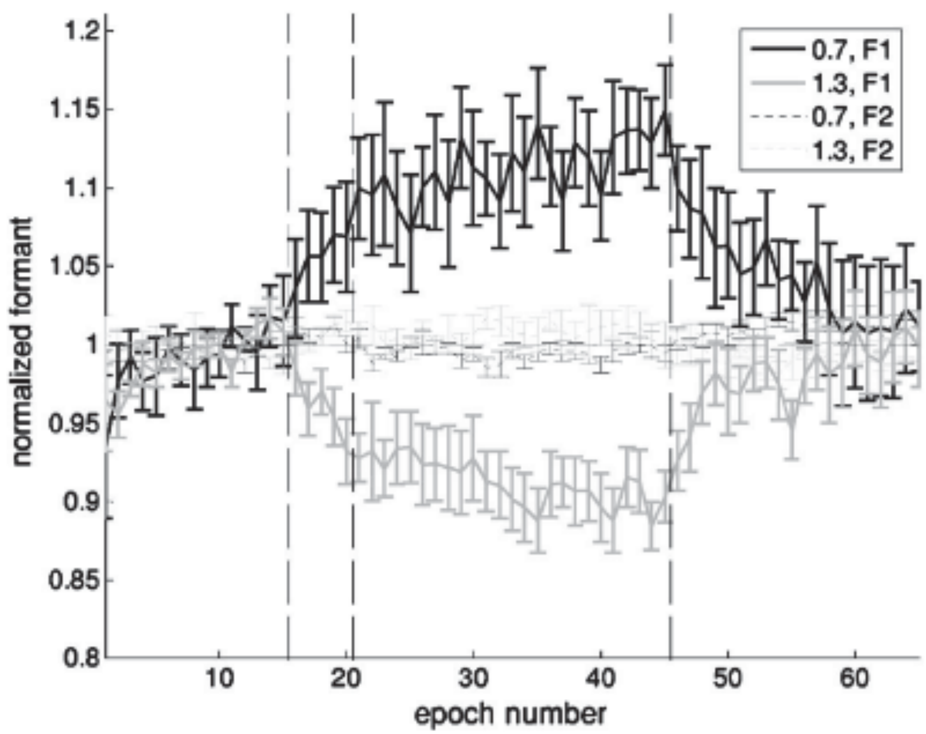
- Formant perturbation was ramped on gradually.
- Normal unmodified feedback was restored in the last 20 repetitions of the word set.

Figure removed due to copyright restrictions.
Source: Figure 2, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.

Results - [ɛ] words with feedback

- Partial compensation in F1, no change in F2.
- Full compensation for F1*1.3 would be reduction to 0.77.
- Full compensation for F1*0.7 would be increase to 1.43.
- Hypothesize that partial feedback results because orosensory feedback is not changed, so large articulatory modifications would be perceived as erroneous.
  - auditory and orosensory targets.

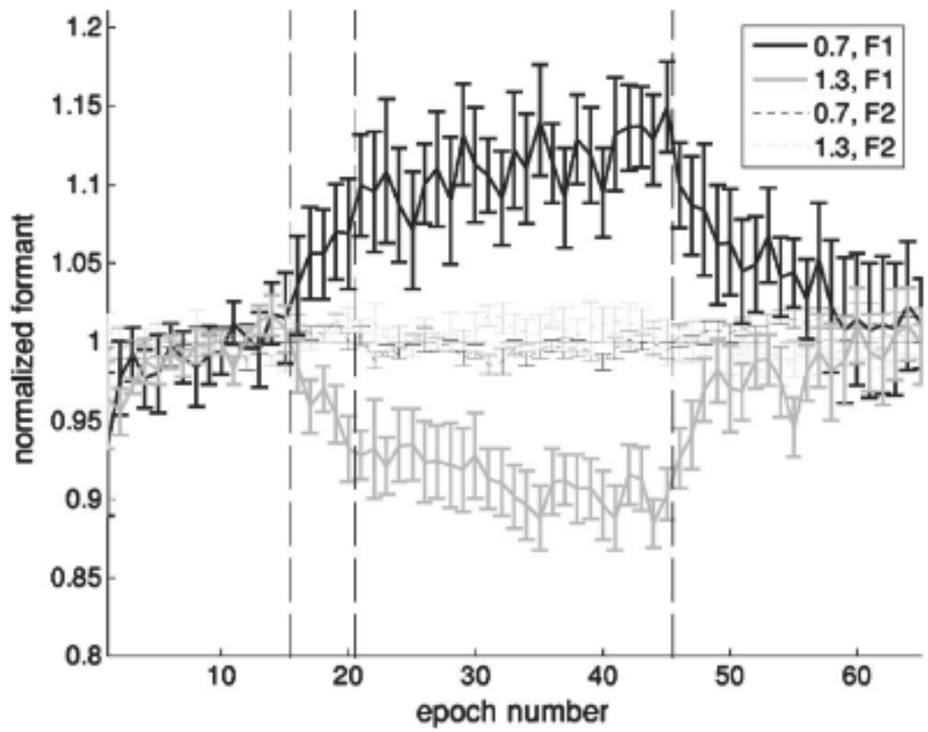
Figure removed due to copyright restrictions.
Source: Figure 2, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.



Results - [ɛ] words with feedback

- Compensation gradually declines after feedback returns to normal.
- Suggests speakers have modified their forward models and have to shift them back to normal based on further auditory feedback.

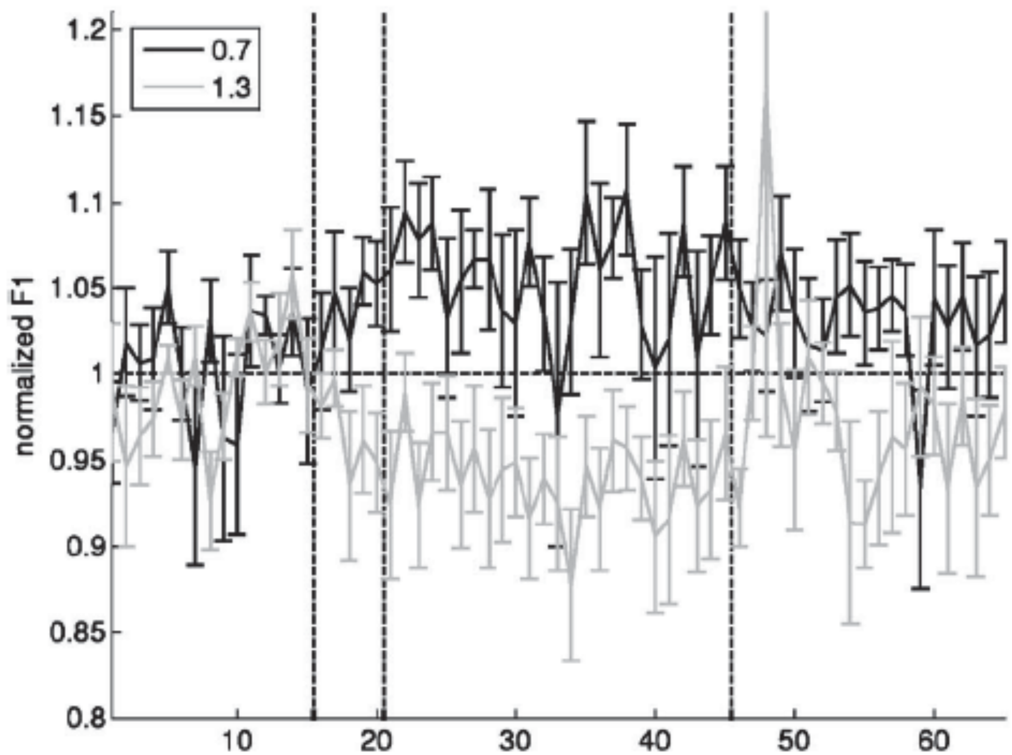
Figure removed due to copyright restrictions.
Source: Figure 2, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.



Results - [ɛ] words without feedback

- Reduced compensation.
- Compensation is affected by ongoing feedback, but some compensation persists in its absence.

Figure removed due to copyright restrictions.
Source: Figure 2, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.



Reproduced from Villacorta, Virgilio M., Joseph S. Perkell, and Frank H. Guenther. "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception." The Journal of the Acoustical Society of America 122, no. 4 (2007): 2306-2319. https:/doi.org/10.1121/1.2773966, with the permission of the Acoustical Society of America.

Results - other vowels, no feedback

- Formants measured as proportions of baseline averages.
- Compensation generalizes to other vowels.
- Suggests the forward model is modified by adjusting overall articulatory to acoustic mapping, not individual vowel targets.
- Less compensation with [i].
  - Stronger orosensory feedback due to tongue-teeth/palate contact?
  - Perhaps compensated for modified F1 of [ε] by proportionately increasing opening, relative to vowel height, as in hyperarticulation. Least effect on highest vowels
  - Not much scope for lowering F1 in [i].

Figure removed due to copyright restrictions.
Source: Figure 5, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.

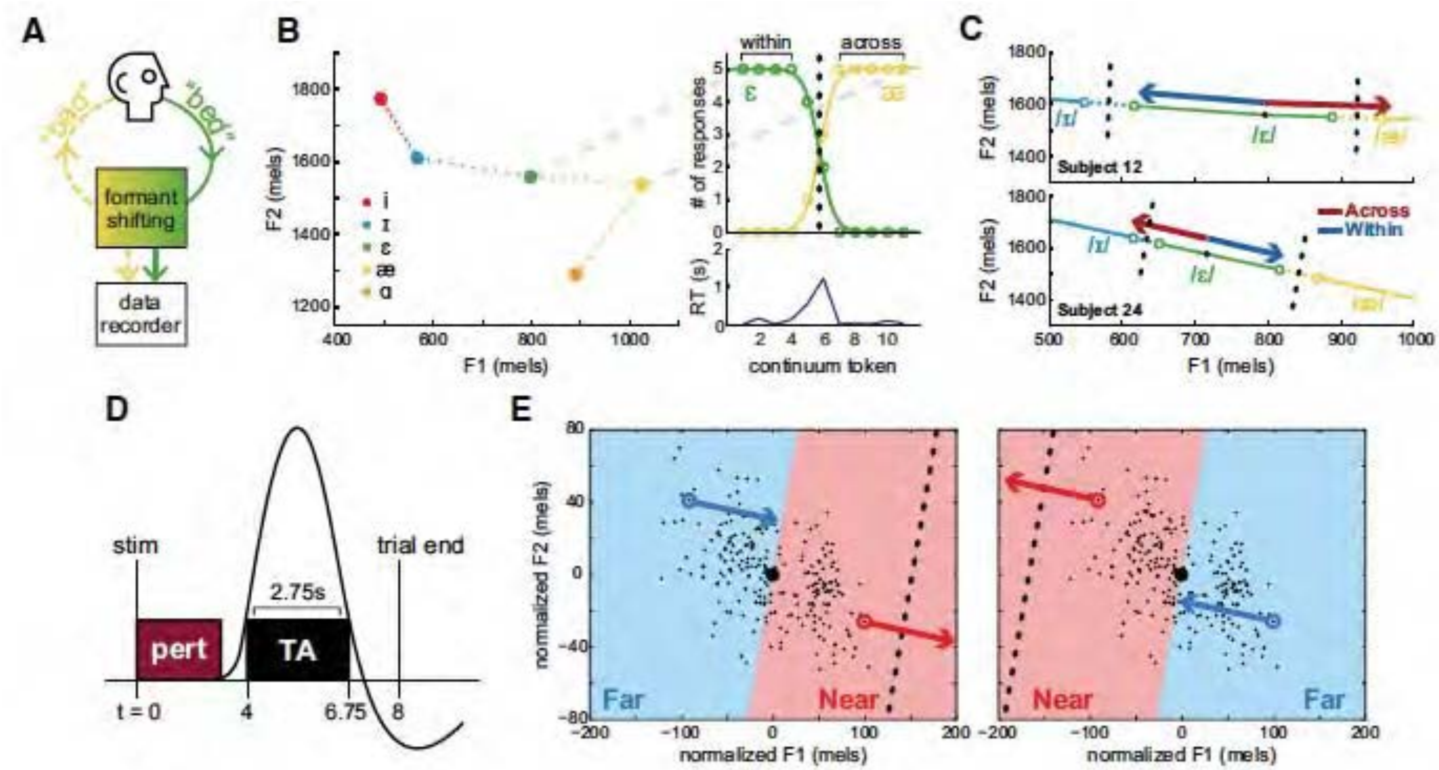
Results - other vowels, no feedback

- No compensation for shifting F1 of [ʌ] up.
  - But there is no [ʌ] word - it seems they mean [ʊ] *put*.
  - Due to one group of speakers (1.3*, male) anti-compensating.
  - [ʊ] is quite variable across dialects.
- F2 shifts in [ʌ]/[ʊ] in response to either F1 shift.
- F2 shifts up in [ɑ] in response to F1 lowering.
  - Why?

Figure removed due to copyright restrictions.
Source: Figure 5, Villacorta, Perkell & Guenther (2007) "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception" The Journal of the Acoustical Society of America.

Modified Auditory Feedback – What determines the magnitude of compensation?

- Magnitude of compensation also depends on whether the formant shift would bring the vowel close to a category boundary (Niziolek & Guenther 2013)
- Greater compensation for shifts of tokens that were near to the relevant category boundary, although all compensation was partial.



Source: Niziolek, Caroline A., and Frank H. Guenther. "Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations." Journal of Neuroscience 33, no. 29 (2013): 12090-12098.

Auditory Targets

- Compensation for modified auditory feedback argues for the existence of auditory/perceptual targets in speech production.
- Compensation appears to involve modification of a forward model of articulatory-to-acoustic mapping rather than adjustment of the articulatory specifications of individual vowels.

Limits on the role of auditory targets?

- Browman & Goldstein (1990) report that the phrase ‘perfect memory’ can be produced with a tongue-tip gesture for the /t/ of ‘perfect’, but ‘careful listening reveals no evidence of the /t/, and no /t/ release can be seen in the waveform.
  - replicated by Tiede et al (2001)

separate words phrase

Figure removed due to copyright restrictions.
Source: Figure 13, Browman, Catherine P., and Louis Goldstein. "Tiers in articulatory phonology, with some implications for casual speech." Papers in laboratory phonology I: Between the grammar and physics of speech (1990): 341-376.

Limits on the role of auditory targets?

- If the targets for [t] are auditory, why would speakers produce an articulatory gesture that has no audible consequences?
- Perkell (2012): frequent words have stored motor plans, which are executed in running speech without calculating the consequences of coordination between words.
- Somatosensory goals for [t]?
- Is the [t] really inaudible – e.g. formant transitions into [kt]?