Articulatory Phonology

1 What is a speech “gesture”?

2 How are gestures coordinated with each other?

3 How does this help explain phonological patterns?

Note: this lecture is an overview of work done primarily by Catherine Browman and Louis Goldstein (Haskins Labs & Yale University).
What is a speech gesture?

1. Coordinative structure - articulators (and the muscles that move them) are organized to produce vocal tract constrictions.

2. Dynamic system - using a spring-mass model to simulate the internal dynamics of articulator motion.

note: the system we are discussing here is an implemented model of phonology. i.e. it is tied to a speech synthesizer and produces speech. It is a research system which is not complete. It is intended to point the way toward a complete model by outlining an approach that, in principle, can be expanded upon as more knowledge about the articulators is acquired and better speech synthesis methods are developed.
Coordinative structures:

**Lip constriction** (two gestural specifications)
- LP - lip protrusion
- LA - lip aperture

**Tongue tip constriction**
- TTCL - tongue tip constriction location
- TTCD - tongue tip constriction degree

**Tongue body constriction**
- TBCL - tongue body constriction location
- TBCD - tongue body constriction degree

**Velo-pharyngeal port constriction**
- VEL - velic aperture

**Glottal constriction**
- GLO - glottal aperture
Dynamics. The motions of the articulators are calculated as a damped mass-spring model:

\[ m \ddot{x} + b \dot{x} + k(x - u) = 0 \]

- m = mass (set to 0 in this implementation)
- b = damping (set to critical damping so there is no oscillation)
- k = the spring stiffness
- x = the current location of the mass (\( \dot{x} = \text{velocity}, \ddot{x} = \text{acceleration} \))
- u = the new (target) location of the mass
2 How are gestures coordinated with each other?

- A “gestural score” coordinates the gestures.
- Relative timing specified in terms of phase relations among gestures.

```
V E L U M
T O N G U E   T I P
T O N G U E   B O D Y
L I P S
G L O T T I S

'p a n '

activation interval

specification of \( u \) and \( k \) for TDCL "narrow" and TBCD "phar"
```
The gestural score (with its specifications of dynamic system parameters) then produces smooth time-varying "movement curves" of the articulators.
3 How does this help explain phonological patterns?

The gestural approach to phonology is good at (offers important insight into):

- assimilation
- deletion
- resyllabification
- metathesis (maybe)

But not good at:

- morphophonemics
- dissimilation
- perceptually driven sound change
“perfect memory” spoken with an aspirated $[t^h]$
Gestural hiding can cause deletion.

“perfect memory” spoken with a deleted [t]
but the tongue blade gesture is still there!
Another case of “deletion” via gestural hiding.

“nabbed most” pronounced as [næbmoʊst], but the alveolar closure gesture is still there.
“seven plus seven” with an alveolar [n] between [v] and [b].
Gestural hiding can cause assimilation.

“seven plus seven” with an assimilated [m]. but the tongue blade gesture is still there!
Gestural blending can result in assimilation.

Movement of the tongue tip in /l/, /t/ and /ld/ in Castilian Spanish.

Place of articulation is backer for /l/ than for /t/ (and presumably also /d/).

Place of articulation in the /ld/ sequence is a blend of the two.

Gestural reorganization can lead to resyllabification.

“kiss ted” spoken more and more quickly - at the fastest rate the [tʰ] becomes [t] as if the syllable break moved [kɪs.tʰed] -> [kɪ.stɛd].

Phase of single glottal opening gesture in [st] is timed as for the cluster, losing the normal phasing for [tʰ].
Assimilation is different from gestural blending.

Consider the /s∫/ sequence in “tops shop”

Is it [∫], [∫∫], or a gesturally blended [s∫]?

Notice that the /∫s/ sequence in “fish soup” behaves a little differently – with less apparent assimilation of [∫] to [s].

EPG index: Speaker JR. Mean of four Type B/C fricatives.

Figure 11. EPG "alveolarity" index for the mean of four tokens classified spectrographically as either B or C (speaker JR).
Figure 12. Mean EPG “alveolarity” index for fricative events in Type D Cs#ʃ, and the two matched control contexts C#ʃ and Cs#C (speaker LA).
Figure 13. Mean EPG "alveolar" index for fricative events in Type D Cs#f, and the two matched control contexts C#f and Cs#C (speaker JR).
Some references:

