THE ESSENTIALS OF SPEECH MECHANICS
SUMMARY

Gary S. Tong (2006)

CONTENTS

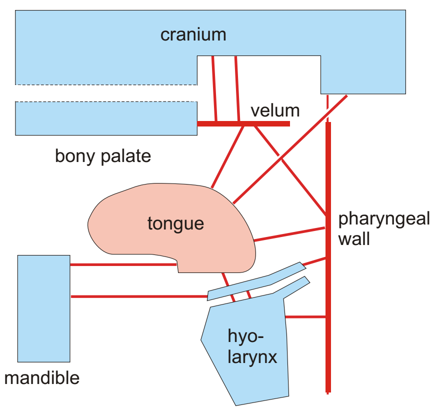
A. METHOD for analyzing the speech mechanism

1. Start with the root behavior of vocalization: **respiration**. Observe what functional modification of the respiratory framework generates the speech framework. Going from the simple to the complex, continue to reconstruct step-by-step the series of gradual modifications that lead to the structuring of speech.

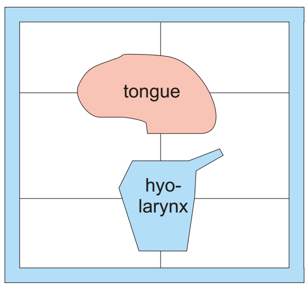
2. Eliminate all **higher** level functions and **variables** of speech that are not fundamentals common to all languages. This is possible through authentic reproduction by a single person (who is both subject and observer) of the articulation of all major (or any) language(s), and judging which behaviors are characteristics of specific languages, and which belong to speech in general.

3. Observe and analyze speech behavior in precise **anatomic** terms. It is essential to clearly perceive the behavior of individual muscles and muscle groups in speech.

4. Evaluate data and and apply **rigorous** testing.

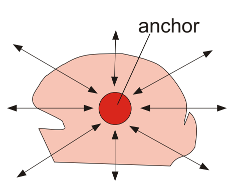


Floating units fig. a.1



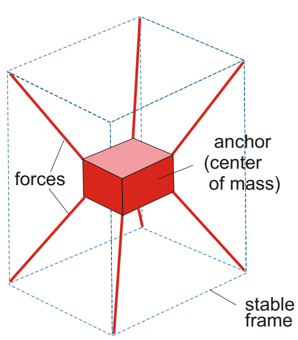
Floating units model

fig. a.1a



Convergent forces in the tongue

fig. a.2



Model of anchor frame

fig. a.3

B. THE ESSENTIALS OF SPEECH MECHANICS

This outline describes the essential structure of speech, built as a hierarchical matrix of forces and their intersecting nodes/anchors the behaviors of which are governed by **glottoregulation** and **metaperistalsis**. These processes are described in the following pages.

ARTICULATION

A. Anchor farmeworks

a.1 Sling attached structures — "Floaters"

The interconnected organs of articulation, the tongue, velum, pharyngeal constrictors and hyo-larynx are relatively free "floating", structural units. They are sling-attached and are **moveable** in various directions by externally imposed muscles. These muscles come into direct contact and coactivity with each other without any intermediary bone, (excepting the hyoid bone which also floats). *See figs.a.1 and a.1.a.*

a .2. Nodes/anchors

Active **external muscles** inserting in such floating muscular organs spontaneously generate an **antagonist** stable **base** within the organ against which they can exert force. Within the organ, where the forces, or more generally their resultants, converge a region of tension develops and serves as a gathering point, or a **node** for these forces. *See fig a.2.*

In the sling attached structures such a node takes the place of bone, or cartilage, the normal **anchoring** material of muscles. The size and location of the anchors varies with the identity and energy level of the incoming forces. The anchor may be called an analog of the center of mass of a system. (*See Appendix: center of mass*)

The main agent of articulation is the tongue and the forces, or their resultants, of its extrinsic and intrinsic muscles converge at one specific region of the tongue mass. Such a region is a lingual **node** or lingual **anchor,** through which the incoming forces are unified to work as a single mechanical **framework.** The tongue performs several functions and has an appropriate anchor frame serving each.

a. 3. An anchor lies at the dynamic center of its framework of forces. (Cf. F.1. Envelopes, below.) Except at sufficiently low energy states of its muscular frame an anchor remains masked by superimposed frames. fig. a.3 (See g.8 *Superimposition*, and g.10 *Masking* .

a. 4. The speech framework

Nodal structures exist in each of the "floating" organs of speech and their monadic interaction forms a composite **matrix** or **unit framework** analogous to a geometry of linear forces interactively "hinged" at various nodes. See fig. a.4.

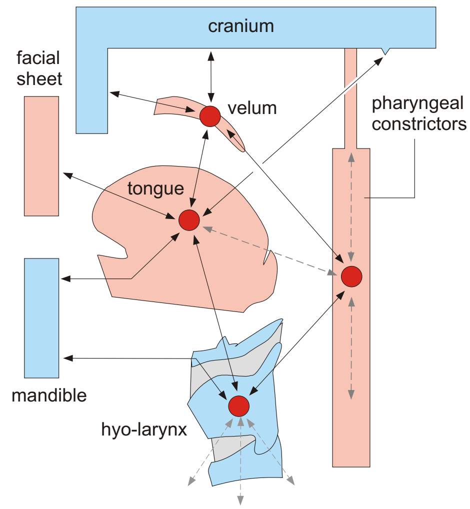
Reducing **complex** speech production to a unitary mechanical matricial device of nodes and convergent forces gives us an effective tool for analysis. It can be expected that, like any system, the "floating" speech organs, without a cohesive, integrated structuring would be highly inefficient in their actions and interactions.

The functional integrity of the separate subunits is generated by a physiological regulatory mechanism centered around the larynx, providing a **constant** against which the many variables in speech can operate, one which directs or regulates the pattern of their cooperative behavior. (This constant, **glottoregulation,** briefly described below, is discussed in section *Fundamentals of Speech*.)

Alternate articulation occurs whenever within the global monadic mechanism any of the various anchors or forces initiates coactive behavior by the others. The symmetries of this framework underlie the agonist-antagonist interactions of speech activity.

Due to such **monadic** unity any subframe of sound production will uniquely reflect the behavior of the entire larger frame structure. The map of forces in the framing of a specific language will differ significantly from those of others in measurable ways. Variation in gesticulation is evidently characteristic of different languages, but we can readily shown that this is not cultural but physio-mechanically formed behavior. See figure a.5.

The framework mechanism allows us to systematically distinguish and analyze the components from the whole framework. The speech system has three major sub-components: a) articulation, b) phonation and c) speech respiration.



The lingual anchor matrix of speech

This simplified diagram shows the lingual anchor as the central agent of the speech framework and does not include the mandible, the oral sphincter, respiratory organs, etc., as these may be considered secondary articulators. Other outlying musculature includes the infralaryngeals. For simplicity the hyoid bone and the larynx are here condidered a single mechanical unit.

fig. a.4

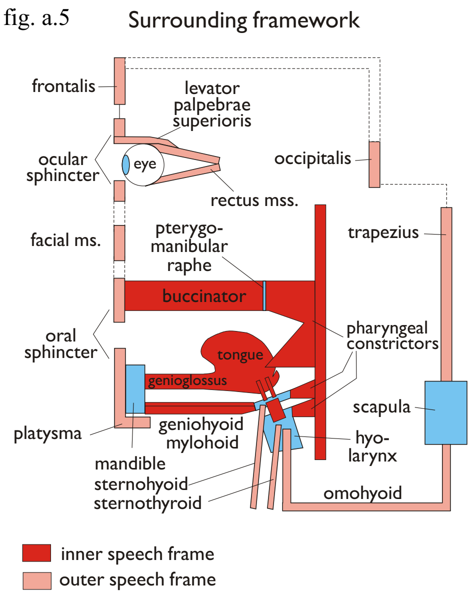
Note: The mechanics described here are easiest to follow and are immediately understood through first hand experience in proprioception. To do so it is essential to set the expertimetnal framework by reducing to a tonic level all direct and indirect parts of speech production, or simply put, to sufficiently relax these. At that time anchors located and isolated and their behaviors can be analyzed. Several of the many available methods for finding and perceiving lingual anchors are offered in the Appendix, p.000

B. Behaviors

b.1. Glottoregulation

Glottoregulation is an agonist-antagonist **balancing** of forces surrounding the **larynx** that maintains the appropriate optimal glottal state during all actions of the upper visceral system. Since the larynx is to a measurable degree highly sensitive to external pressure, and its distortion by **external** forces creates immediate glottal stricture, there is an **equalization** behavior built into the speech mechanism which constantly the glottal state by balancing compensation for external distortions. Every instance of speech production contains a glottoregulative factor. (This topic, including data demonstrating this process is covered in section *Foundations of Speech*).

fig. a.5



The **mandible** is directly tied to lingual anchors through the genioglossus muscle, etc. The **facial** musculature is connected to the pharyngeal constrictors through the oral sphincter and buccinator. The (external) eye muscles, aside from neural paths, are united with the facial sheet through the levator palpebrae superioris.

Note: Accessibilty to measurements:

The lingual frame is the **central** agent of speech but its muscular behavior is difficult to instrumentally measure. However, **externally** to the lingual frame there lie, monadically united with it, coactive subframes including those of the lips, mandible, face, eyes, arms and hands that are observable and accessible to measurements. Any action of the tongue has a unique corresponding counterpart in the musculature of the face, hands, arms, eyes, etc. Cf. the built-in coaction of sound production and head movement in animals and especially gesticulation in humans. **(***Details in Chapter 000*).

Figure a.5 shows the main connections between the inner and outer frameworks of speech. The velar apparatus is not included for simplicity.

Since facial behavior, gesticulation, voice pitch, etc. distinctly varies among the major language families, languages, and to a lesser extent among individual speakers, data can be collected from such externally accessible regions to help reflect the behavior of more interior, inaccessible regions of the frame.

b.2. Metaperistalsis

The second basic mechanical function of speech is a highly developed form of peristalsis. The most **primitive** behavior of the pharyngo-visceral tract, which later evolved into the upper respiratory-masticatory, or upper visceral (UV) tract was **peristalsis.**

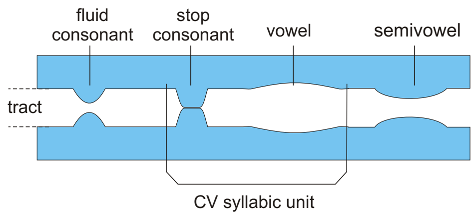
As evolution is a process that modifies, rather than creates ex nihilo, it may be inferred that the mechanics of the modern UV are complex developments of peristalsis. In fact, this appears justified because the interpretation of speech in such terms easily and consistently explains its behaviors. Syllabification is one instance. The diagrams of the wave of phonation of the word "tomato" as closed-open sequences and of the sound produced by a goat on slides by MacNeilage (source) are diagrams of peristaltic waves.

To differentiate the more complex articulatory and masticatory mechanics of open-close sequences from simple one-way passages of bolus, this behavior involved can be termed **metaperistalsis**. **(**For additional details, see the section *Phoneme Production*, p. 000, (not included in this outline). **Peristaltic** action has been **recognized** as the mechanical function of suction and of bolus transport in **suckling.** Cf. Woolridge, 1986; Buckley, et. al 2006, and others listed on p. 22.

The earliest vertebrate ancestral feeding tube was **linear**. Metaperistalsis evolved as a complex reorganization with bifurcations, detours and overlaps of the original linear tract when a) the nasal and the aural tracts emerged from the oral and pharyngeal line and were repositioned, b) when the jaw and tongue and allied musculature appeared, c) when the hypobranchial (under-gill) muscles moved to cover the face, head, neck and shoulders and were physically connected to the branchial (gill-derived) structures of the tract such as the hyolarynx. This has produced complex overlapping, simultaneously occuring patterns of peristaltic expansions and constrictions. For instance, while a vowel is expansive, its simultaneously occuring phonation further down the tract is constrictive.

fig. b.1

Peristaltic patterns in articulation



In this diagram dashed lines indicate the median cross-section of the tract. Vowel and consonantal expansions and compressions of tract segments are shown for various syllables types.

1. Any syllable, whether vocalic or consonantal (V, VC, CV) is built from peristaltic units, from a variety of patterns composed of partial phonating closure, full tract closure and open tract units. These actions are partial or complete peristaltic three-element movements.

2. The production of phonemes is also a modified peristaltic derivative. Full stops are complete closures, while other phonemes, vowels and consonants (voiced, voiceless, affricates, fluents, etc.) are modified epansions expansions and closures of the tubular tract.

In view of the discovery of the **genetic toolbox**, which supports the notion that nature builds on existing resources**,** it is not surprising that peristalsis can be shown to be the fundamental behavior later modified into a more complex system of respiratory, masticatory and sound producing mechanics that we can call **metaperistalsis**.

Peristalsis occurs when tubular visceral tract undergoes a segmentally moving wave sequence of alternating contraction and expansion of tract cross section. Such behavior has been recognized in suckling and deglutition, but metaperistaltic functions can be successfully interpreted to generate all actions of the upper respiratory-visceral tract and so underlie speech production as well.

The upper visceral tract contains muscles of somatic developmental origin as well, however these unite with the true visceral structure to generate behaviors that are mechanically still visceral.

The ordered sequences of **metaperistalsis** can be developed into more complex behaviors by **modifications** such as combining peristaltic segmental units or elements in variant groupings and in directional reversal. (Note that whereas primary power in respiration generates inspiration, in speech the primary power is assigned to expiration. Cf. also food input vs. Regurgitation).

An example of visceral segmental functioning in **speech** is in syllabification. *See fig. b.1.* Phrasing, on the other hand, reflects metaperistaltic segmentation in **respiration**.

As sound production is a coaction of articulation and respiration (the latter in the form of pulmonary pressure, phonation and speech respiration), speech mechanics can be seen as a combination of two different sources of peristaltic action. The **integration** of respiratory and masticatory behaviors in a complexly constructed alternation of contractions and expansions in particular sequences underlies the **functions** of the upper visceral system, the main ones being respiration, feeding (suckling/ mastication/deglutition) and speech. In respiration one breathful of air is a metaperistaltic segment. "Breathfuls" can be added by segments or glided into continously without obstruction. A **yawn** glides through all segments with maximal tract expansion. (For details see Chapter on Respiration, Chapter 000.)

b.2.1 Support for the *frame and content* theory of MacNeilage and Davis

Metaperistalsis clarifies both the **frame** and associated **content unity** and its **serial/parallel** forces components proposed by MacNeilage and Davis. Peristalsis has two force components: longitudinal and transverse. In the mechanism of sound production these manifest as the **serial** and **parallel** aspects of speech. The unit CV frame-plus-content is, on the other hand, interpretable as the module consisting of two consecutive peristaltic segments, in which the content is **parallel** or **transverse** and the sequence of C and V is **serial** or **longitudinal.** A description of the basis for the constancy and detemination of the frame content is not covered here as it involves additional combined functions involving the jaw, tongue and the cyclicities of food ingestion, in suckling (cf. Woolridge, 1986) and in mastication (cf. Hiiemae and Palmer, 2003). For details see Ontogeny of phonemes, Chapter 000.

C. Hierarchical organization

c. 1. Hierarchical series of anchors

Speech production, being an ordered system, is **hierarchically** organized. In terms of anchors this means that there is for a given function a single t**op level** node which controls, or interacts with a number of lower level nodes. This is a two-way relationship; a higher anchor is to an extent affected by subordinate anchors when these are present. (The source and mechanics of hierarchical sequence ordering is not discussed here.) See fig. c.1.

Hierarchical rank is a function of the amount of **energy** input applied to a particular framework. The mechanical behavior of the speech mechanism is such that it is "pre-wired" to initiate and complete action sequences when energy is applied. When going into the speech mode, as energy input increases, so does tension, cross section of oral tract, and the degree of approximation of articulative posture. (Cf. increase in size of oral tract cross sections in relation to energy level of respiration.)

The "pre-wired", **built-in** a nature of anchor sequences is the basis of language learning. Without such device the acquisition of human speech or for that matter, that of certain animal vocalization would be excessively complicated. Similar built-in automatization is evident in suckling, eating, walking, etc.

When the speech framework is present, with increasing general tension applied to the framework, starting at the lowest energy level, the framework passes through successive hierarchical levels. Normally this passage through levels can be rapid, and apparently simultaneous, but its is observable when tension is gradually increased, either experimentally or in moments of naturally slowed speech.

A natural hierarchical sequence is the dynamically optimal path of transition between anchor frames. It is therefore present in all efficient movements, e.g., as in the locomotor behavior of limbs. Hierarchical sequences, i.e., dynamic paths can be terminated and switched to other paths through glide or pulsed transformations. (See anchor transformations, g.11).

fig. c.1