24.941J / 6.543J / 9.587J / HST.727J The Lexicon and Its Features
Spring 2007

THE LEXICON AND ITS FEATURES
6.543J/HST.727J/9.587J/24.941J SPRING 2007

Professor Kenneth N. Stevens
*Lecture 1*   *02/08/07*

Readings:
1. Halle, M. and K.N. Stevens (1991) Knowledge of language and the sounds of speech.   In J. Sundberg, L. Nord, and R. Carlson (eds.), *Music, Language, Speech and Brain*, Basingstoke, Hampshire: Macmillan Press, 1-19.
2. Stevens, K.N. (1998) *Acoustic Phonetics*. Cambridge MA: MIT Press. Chapters 1 and 4.
3. Stevens, K.N. (2001) The properties of the vocal-tract walls help to shape several phonetic distinctions in language. In **Travaux du Cercle Linguistique de Copenhague,Vol. XXXI**, 285-297.
4. Keyser, S.J. and K.N. Stevens (2006) Enhancement and Overlap in the Speech Chain. **Language 82**, no. 1, 33-63.
5. Stevens, K.N. and S.J. Keyser (manuscript) Presented at Colloquium, Paris, France, July 2006.

Components of the speech production system
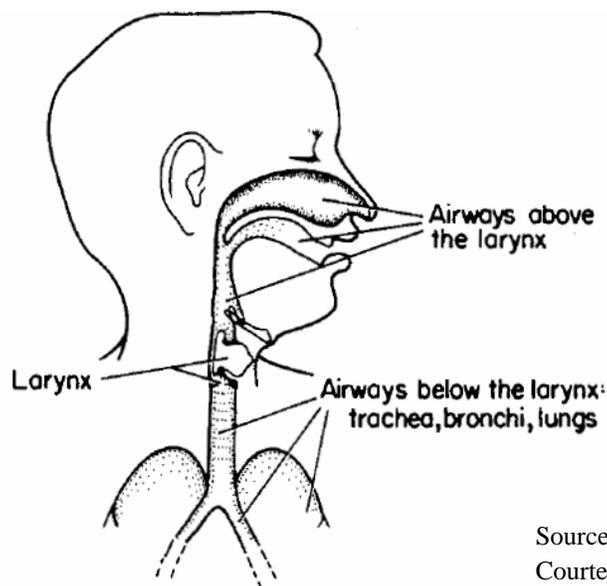
Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Figure 1.1 Schematic representation of the three components of the speech production system.

There are three major parts to the speech production system: airways below the glottis, the vocal tract proper from the glottis to the lips, and the nasal cavity. Usually the glottis forms a narrow opening so that the acoustic coupling between the subglottal and supraglottal system is small. However, the air from the lungs passes through the trachea and the glottis and there is a pressure drop across the glottis. Thus the glottis plays an important role in regulating airflow. The velopharyngeal port usually remains closed at all times except in the vicinity of a nasal consonant (in English).
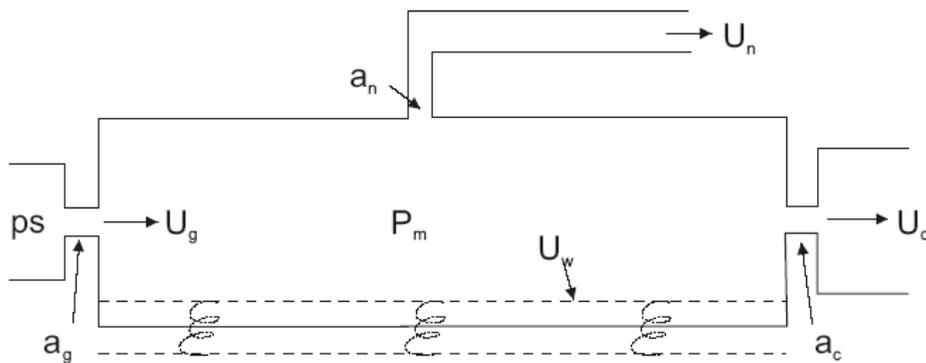
Fig. A

This picture schematizes the airflows and pressures in this three-component system, including a representation of a narrow constriction **ac** in the main vocal tract that might exist in the airway when a consonant is being produced. The dashed lines indicate that the vocal-tract walls can yield or move in response to changes in the intraoral pressure **Pm**. The subglottal pressure in **ps** and the opening to the nasal cavity is **an**. Ug, Un, Uw and Uc are airflows.

Our discussion of human production and perception of speech will cover four different areas:

(1) respiration, airflows, pressures, interaction with yielding walls

(2) generation of sources of sound, the filtering of these sources by the vocal tract and other airways, and the radiation of sound from the lips, the nose, and the vocal-tract walls

(3) the role of the human somatosensory characteristics in controlling speech production
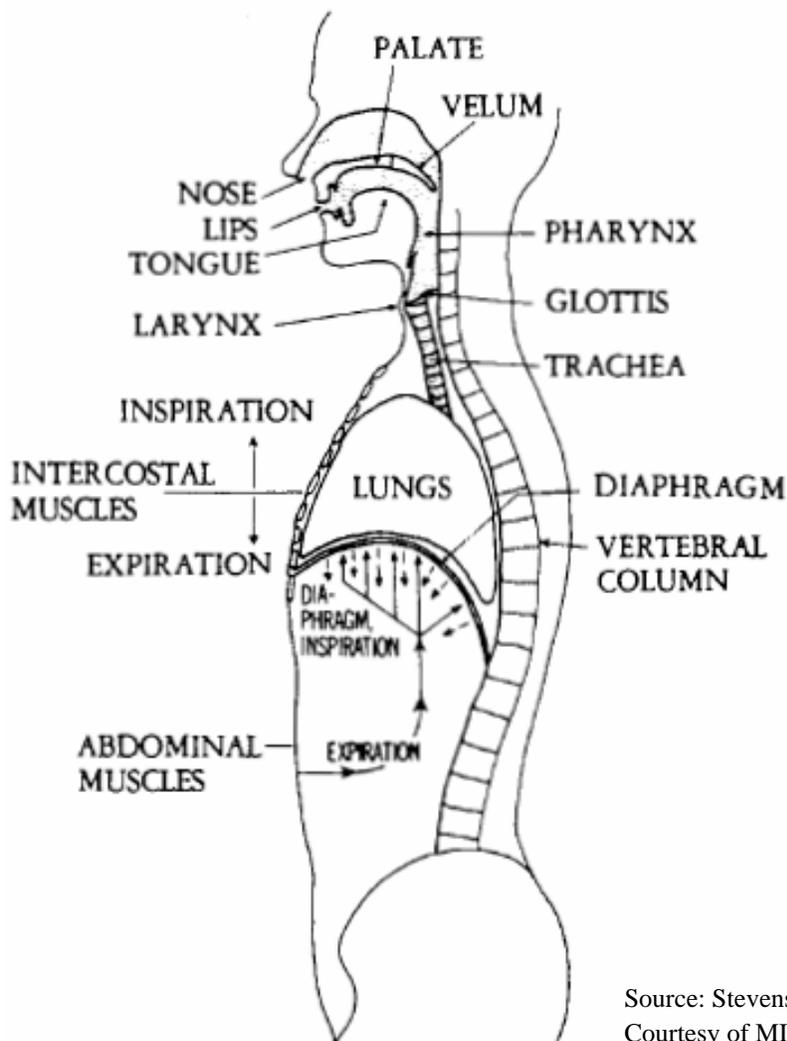
(4) the role of auditory response to sound

Airflow through a constriction
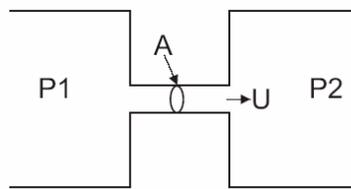
$$P2 - P1 \cong -\frac{\rho U^2}{2A^2}$$

Pressures P are in dynes/cm^2, sometimes in cm H$_2$O
1 cm H$_2$O $\cong$ 1000 dynes/cm^2
Volume velocities U are in cm^2/sec
Area A is in cm^2

It is approximately true for this type of constriction that there is no pressure recovery at the outlet.
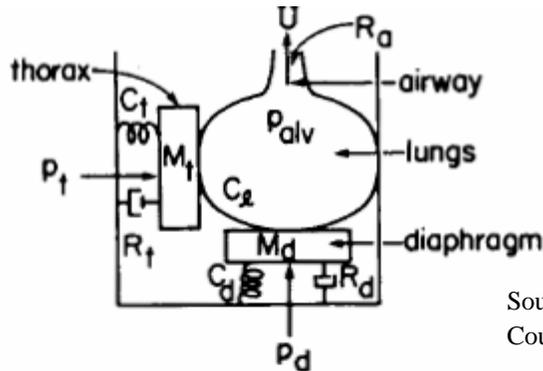
Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Figure 1.2 Schematic representation of the methods for controlling respiration. The principal inspiratory muscles are the external intercostals and the diaphragm. Expiratory muscles include the internal intercostals and the abdominal muscles. (From Pickett, 1980.)

This figure schematizes the respiratory system. When we take a breath in preparation for speaking, we expand the lungs by contracting intercostal muscles and by lowering the diaphragm. Following the inspiration the lung volume decreases in part due to passive return to a smaller volume, but with continued adjustment of the muscles of inspiration and inspiration. The net result is that the tracheal pressure remains roughly constant through much of the expiration, except for creating special prominences or inserting pauses. A typical subglottal pressure during normal speaking is in the range 6-8 cm H_2O.

A very simple schematization of the respiratory system during speech production, together with some rough values of components is shown below.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Figure 1.3 Schematization of the mechanoaerodynamic system for controlling respiratory pressure and flow. Two pressure sources are assumed: p_t for the thorax and p_d for the diaphragm. These forces give rise to an alveolar pressure p_{alv} and a flow U through the airways.

average values have been estimated (Van den Berg, 1960; Rothenberg, 1968). These estimates are as follows:

$C_t = 0.06 \, \text{cm}^5/\text{dyne}$

$M_t = 0.04 \, \text{gm}/\text{cm}^4$

$R_t = 4 \, \text{dyne-s}/\text{cm}^5$

$C_d = 0.09 \, \text{cm}^5/\text{dyne}$

$M_d = 0.2 \, \text{gm}/\text{cm}^4$

$R_d = 9 \, \text{dyne-s}/\text{cm}^5$

$C_\ell = 0.003 \, \text{cm}^5/\text{dyne}$

During phonation, the action of the subglottal pressure on the vocal folds can set these structures into vibration. A section through the larynx (Figure 1.4) shows the vocal folds. The subglottal pressure below the vocal folds (in the trachea) will cause them to vibrate. Pulses of air pass between the folds.
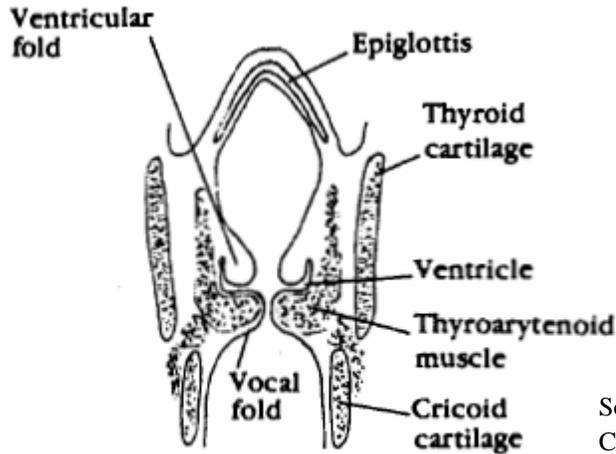
Figure 1.4 Coronal section through the larynx showing the vocal folds and the ventricular folds, together with the thyroarytenoid muscles and cartilages surrounding the folds. (From Dickson and Maue-Dickson, 1982.)

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

A typical waveform of the volume velocity through the glottis (the space between the vocal folds) is shown below in Figure 1.27. In this example, the airflow is zero for about one-half of a cycle and the airflow has a peak value of about 600 cm^3/s. The average value is about 150 cm^3/s.
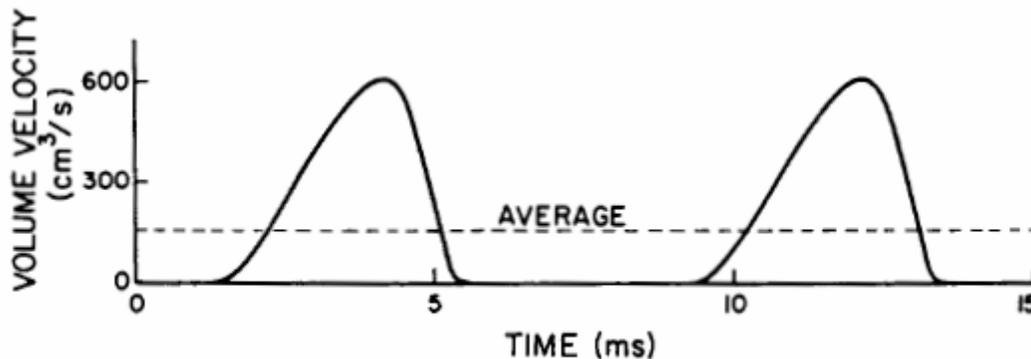

Figure 1.27 Sketch of the volume velocity of glottal airflow vs. time for normal adult male glottal vibration.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

During a typical respiratory cycle in speech, the lung volume might fluctuate by about 1000 cm3 (1 liter). Thus during phonation like that in Figure 1.27, a steady phonation could be maintained for about 7 seconds. During this kind of phonation for a vowel, there is just one narrow constriction in the airway --- the constriction at the glottis (although during phonation it is a time-varying phonation). Phonation of this kind, with just one major constriction, also occurs for some consonants, like nasal consonants, liquids (/l/ and /r/), and glides (/w/ and /j/). These are called sonorant consonants.

The flows, pressures, and the sound become more complicated when there is a constriction in the vocal tract above the glottis as well as a constriction at the glottis. This situation was represented

schematically in Figure A above. It is also shown as an equivalent circuit in Figure 1.36. In Figure 1.36, the additional constriction is at the lips.
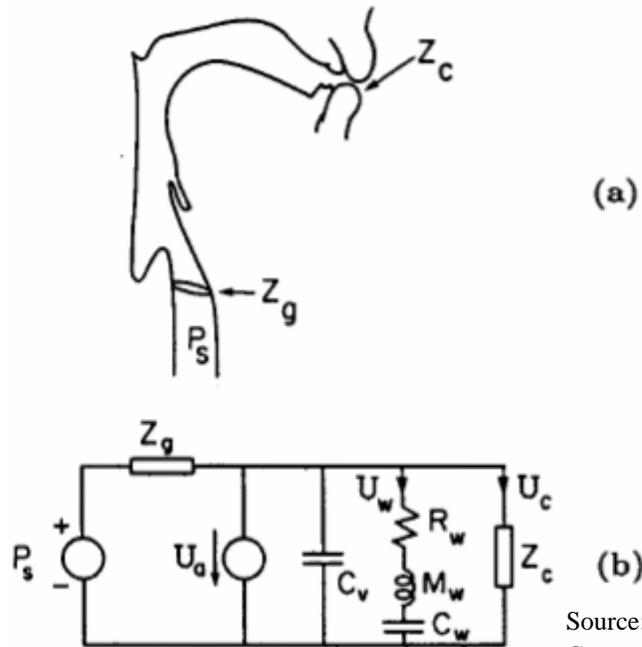
Figure 1.36 (a) Midsagittal shape of the vocal tract when there are constrictions at the glottis and at the lips. The acoustic impedances of these constrictions are Z_g and Z_c, respectively. (b) Equivalent circuit for the configuration in (a). See text.

Figure 1.36 shows the subglottal pressure Ps, the glottal constriction Zg, the supraglottal constriction Zc, and some other circuit elements representing the yielding walls of the vocal tract (Rw, Mw, Cw), the acoustic compliance Cv of the air between the two constrictions, and a "source" Ua that represents active expansion of the volume between the constrictions (or contraction if Ua is negative). All of these components can play a role in producing consonants that involve a pressure increase (or decrease?) in the oral cavity. These are the so-called obstruent consonants. For these consonants, the vocal folds no longer vibrate freely wit no supraglottal pressure, but either vibrate less strongly or not at all. Also for these consonants, there is a pressure drop across the supraglottal constriction, equal to the interaural pressure Pm. As a consequence there is a rapid flow of air through this constriction, and turbulence in this flow can generate a source of sound. This generation of noise (or sometimes a transient sound) is a kind of defining acoustic attribute of an obstruent consonant. Also when producing an obstruent consonant, a speaker senses the buildup of pressure in the mouth, and hence such a consonant has a somatosensory component.

[Sonorant]

With regard to the distinctive feature [sonorant], then, we have the following defining attributes: For the feature value [+sonorant]:
  (a) The glottal source can be generated with full amplitude without forming a downstream pressure increase.

(b) There is no noise generated by turbulent airflow at a supraglottal constriction.

(c) There is no excitation of somatosensory receptors due to a pressure increase.

For the feature value [-sonorant]:

(a) The glottal source is attenuated or obliterated due to a reduced transglottal pressure.

(b) Noise is generated due to turbulent airflow in the vicinity of a supraglottal constriction.

(c) Pressure is built up in the oral cavity behind the constriction; this pressure buildup produces a distinctive somatosensory response in the speaker's oral cavity.

Sound sources for speech

The defining attributes for the distinctive feature [sonorant] are concerned with sound <u>sources</u> that are produced in speech --- a noise source due to turbulence in the airflow near a constriction in the vocal tract, and a periodic source arising from vocal-fold vibration. We'll discuss these sources briefly before we move on to other features.

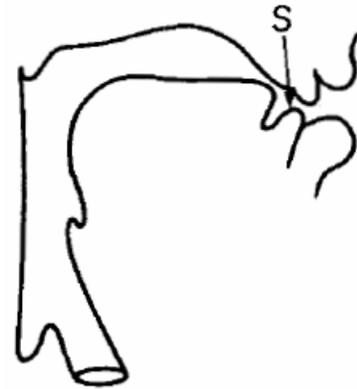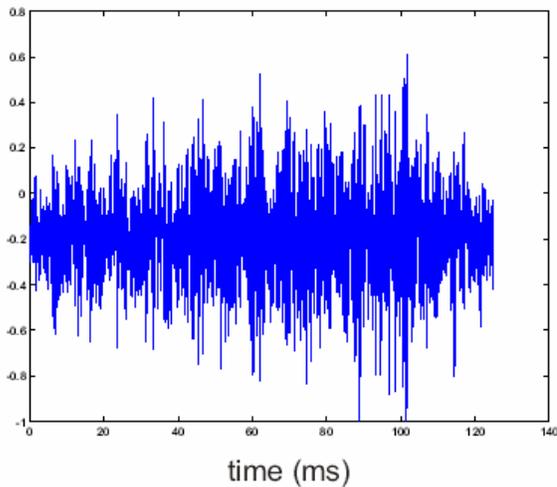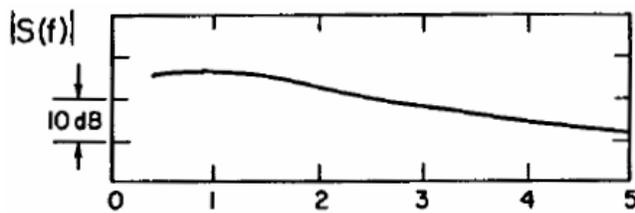A general model of speech sound production is given below.
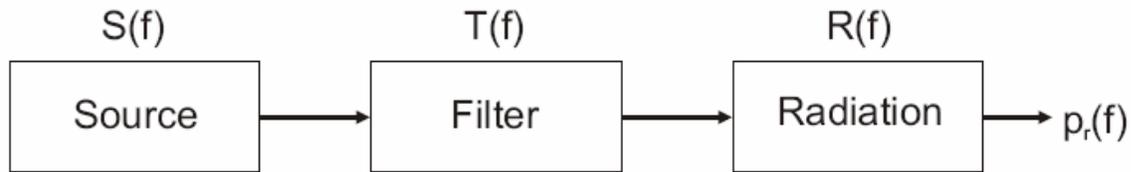
Glottal source:
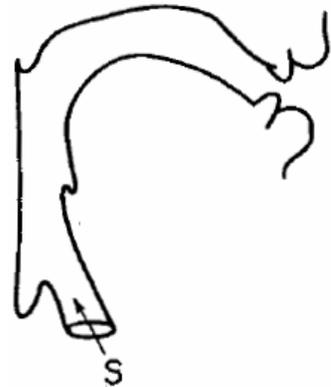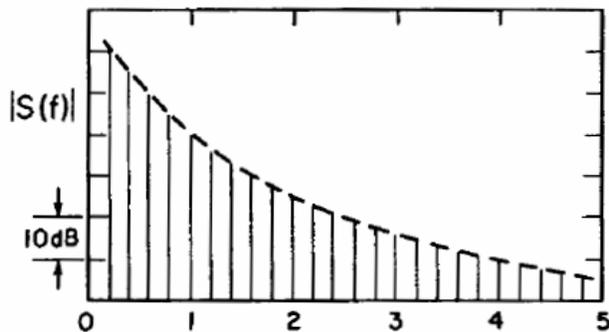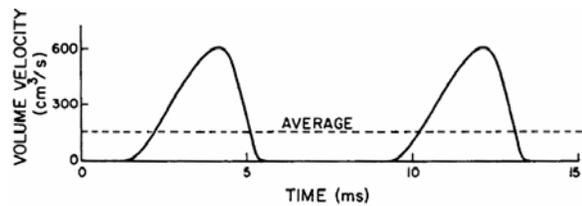
Source: Fig. 1.27 and Fig. 3.1 in Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Turbulence noise source:

Source: Fig. 3.1 in Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

Source: Stevens, K. N. *Acoustic Phonetics*. MIT Press, 1998.
Courtesy of MIT Press. Used with permission.

The filtering of these sources modifies the spectrum by introducing peaks in particular frequency ranges. One way of looking at it is that the sources excite the resonances or "natural frequencies" of the vocal tract. In the case of a noise source near a constriction, it will tend to excite the resonances of the part of the vocal tract anterior to the constriction. Usually, much of the noise is produced by a jet of air from the constriction impinging on an obstacle, such as the lower teeth.

There are a few things about the glottal source and the noise source that are relevant to the discussion of distinctive features.

The glottal source is produced by setting the vocal folds into vibration by applying a subglottal pressure. The vocal folds will only vibrate when certain conditions are met. If they are too far apart, they will not vibrate. If they are pushed too tightly together, the subglottal pressure will not be enough to start them vibrating. For an intermediate spacing, a glottal source with maximum amplitude will be generated. Also the stiffness of the vocal folds affects their ability to vibrate. When they are stiff, a greater transglottal pressure is needed to set them into vibration. When they are more slack, they will vibrate with a lower transglottal pressure. And there is a minimum transglottal pressure below which vibration cannot occur. When there is

vibration, the amplitude of the volume-velocity source is roughly proportional to the transglottal pressure to the power 3/2.

The amplitude of the turbulence noise source in the vicinity of a narrow constriction is roughly proportional to $\Delta P^{\frac{3}{2}}.A^{\frac{1}{2}}$, where ΔP is the pressure drop across the constriction and A is the cross-sectional area of the constriction. Turbulence noise can also be generated at the glottal constriction, and is called aspiration noise.

[Continuant] and [Strident]

A [-sonorant] or obstruent consonant is produced by creating a narrow constriction in the airway, usually in the oral cavity, and by building up pressure behind this constriction. The constriction can form a complete closure to generate a stop consonant, or can form a narrow opening to produce a fricative consonant. The contrast between these two types of obstruent consonants is captured by the distinctive feature. [continuant]: + for fricatives, - for stops.   In the case of a [+continuant] segment, there is continuous generation of noise during the constricted interval. For a [-continuant] segment, no sound is generated during the closure interval, and then a short burst of noise is produced at the time the closure is released. The pressure behind the constriction for these obstruent consonants creates forces on the surface of the articulator that is forming the constriction, and these forces influence the positioning of the articulator both during the fricative and in the time interval following the release.

Within the class of [+continuant] consonants, there are some consonants, produced by forming a constriction with the tongue blade, for which the amplitude of the noise spectrum at high frequencies (in the range of the third formant or higher) is greater than the amplitude of the spectrum in the adjacent vowel in the same frequency range. These consonants have been classified as [+strident]. For other fricative consonants the spectrum amplitude of the noise in this frequency range is much less.

The special characteristic of the production of [+strident] consonants can be illustrated with the models in Figure B.
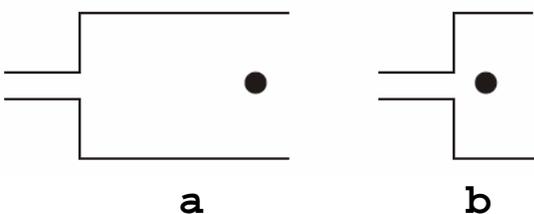
Schematic representations of the front cavity in a sound like /s/ (part a), and like /θ/ (part b).

a   **b**   **Figure B**

The model in Part **a** is intended to represent a sound like /s/ where there is constriction and an opening into a front cavity. The model in part **b** shows a shorter front cavity and a source closer to the constriction, as might exist in the fricative /θ/. Estimates of the sound pressure of the radiated sound show that the amplitude is much greater in **a** than in **b**. The difference is about 15 dB, assuming the same source amplitudes in the two cases. In **a** the source is located at a place where it is more effective in exciting the front cavity resonance. Thus it appears that the

[+strident] consonant achieves its high amplitude at high frequencies because of the relative position of the source.

[Stiff vocal folds]

We turn next to the voicing distinction for obstruents. All obstruent consonants in English can be either voiced or voiceless. One defining attribute of obstruent consonants is that they are produced with a raised intraoral pressure. This means that the transglottal pressure (the pressure drop across the glottis) is reduced relative to the subglottal pressure.

The various conditions are schematized in Figure C, where the phonation amplitude is plotted against vocal-fold stiffness for several values of intraoral pressure (assuming a fixed subglottal pressure of 8 cm H_2O). For an intraoral pressure of more than 6 cm H_2O, the transglottal pressure is 2 cm H_2O, and is below the threshold of phonation, and hence there will be no phonation. For smaller values of intraoral pressure (say 4 cm in the figure), phonation will occur over a range in which the vocal folds are slack, but for greater stiffness phonation will not occur. There is a value of stiffness where there is an abrupt change from phonation to no phonation. The name [stiff vocal folds] has been used to identify this feature: for [+stiff vocal folds] glottal vibration is inhibited, and for [-stiff vocal folds] glottal vibration is facilitated within the consonant region. Other articulator actions can also be implemented to inhibit vibration, such as spreading of the glottis. These matters relating to enhancing gestures for features will be discussed later.

The voicing feature [stiff vocal folds]
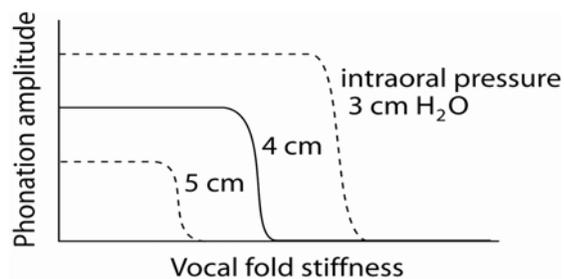
For a given intraoral pressure, increase in vocal-fold stiffness causes an abrupt termination of glottal vibration a particular threshold of transglottal pressure. Subglottal pressure assumed to be 8 cm H_2O

Figure C