# INTERACTION BETWEEN ACOUSTIC SOURCES AND VOCAL-TRACT CONFIGURATIONS FOR CONSONANTS

KENNETH N. STEVENS

Research Laboratory of Electronics and
Dept. of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge MA 02139 USA

## Abstract

When a narrow constriction is formed in the vocal tract, there are substantial interactions between the vocal-tract configurations and the sources of sound in the vocal tract. In the case of liquids and glides, the interaction causes the amplitude of the glottal source to decrease somewhat relative to its amplitude for vowels. When the constriction is narrow enough to cause frication noise, there are stron interactions between vocal-tract configuration and both the noise source and the glottal source. A theoretical analysis of some aspects of this interaction is presented, together with some illustrative data.

## 1. Introduction

It is usually assumed that the acoustic source due to vocal-fold vibration for vowels is largely independent of the vocal-tract configuration. That is, the shape and frequency of the pulses of volume velocity passing through the glottis are not significantly influenced by the vocal-tract configuration, and the acoustic properties imposed on the sound by the vocal tract are largely independent of the glottal source. Recent analytic and experimental work has indicated that there are some interactive effects, between source and vocal tract [1, 2]. The interactive effects on the source become strongest when the vocal tract is constricted and also when the fundamental frequency is in the vicinity of the first formant frequency (F1). The effect is to modify somewhat the waveform of the airflow pulses emerging from the glottis, whereas the effect on the pattern of mechanical vibration is relatively small.

The basic mechanism of the interaction is that, during the open phase of the glottal cycle, the airflow is determined not only by the impedance of the glottal opening, which is largely resistive, but also by the impedance of the sub- and supraglottal airways, which is largely reactive in the frequency range of glottal vibration. This reactive impedance, which is determined by the acoustic mass of the airways, produces a skewing of the volume-velocity waveform relative to the waveform of glottal area versus time. There is also a small effect of the vocal-tract configuration on the frequency of glottal vibration. This is thought to be due to mechanical coupling between displacement of the tongue root and rotation of the thyroid cartilage, leading to a change in vocal-fold tension [3].

The approximation of independence between acoustic sources and vocal-tract shapes is far from being valid when the airways between the glottis and the vocal-tract output contain a narrow constriction, as is the case for many consonants. When there is vocal-fold vibration during the interval in which the airways are constricted, the shape and frequency of the pulses of glottal airflow can be strongly dependent on the supraglottal configuration, including the size of the constriction, the impedance of the vocal-tract walls, and the way the walls are displaced during the constricted interval. Under some circumstances, the airflow through the constriction is sufficient to generate a source of turbulence noise, and the characteristics of this noise are clearly dependent on the shape of the vocal tract. The aim of this paper is to explore the nature of the interaction between acoustic sources and vocal-tract shapes for these more constricted, consonantal configurations.

## 2. Glottal source for liquids and glides

The articulatory configuration for liquids and glides is generally characterized by a constriction or narrowing at some point along the vocal tract such that the cross-sectional area is smaller than that for vowels. Measurements have been made to determine the effect of this type of constricted configuration on the glottal source [4, 5].

In one type of measurement, the spectrum of the sound was compared at two points within each of a number of utterances containing liquids and glides: (1) when the vocal tract was constricted during the liquid or glide and (2) when the vocal tract was relatively unconstricted in a vowel adjacent to the consonant. Measurements of the amplitude of the fundamental component were made in these regions, and appropriate corrections were made for the influence of the first-formant frequency on the amplitude of this component in the glottal spectrum. A change in the amplitude of this component in general indicates a change in the amplitude of the glottal pulses. The results for six speakers showed that when the vocal tract was maximally constricted for a liquid or glide, there was a reduction of about 3 dB in the amplitude of the glottal pulses, on the average. This relatively modest change can presumably be ascribed, at least in part, to an acoustic interaction between the glottal airflow and the increased impedance of the airway for the constricted configuration.

In addition to the acoustic measurements, an electroglottograph was used to obtain estimates of glottal opening and closing times for a number of glottal cycles in the same liquids and glides and in the adjacent vowels. The aim was to determine whether there was evidence for an influence of the vocal-tract constriction for the consonant on the mechanical aspects of vocal-fold vibration. The measurements showed that there was an average increase of about 10 percent in the glottal open time within the vibratory cycle during the constricted interval for liquids and glides relative to that for a vowel.

Models of vocal-fold vibration predict effects in the same direction, i.e., a decreased amplitude of the glottal pulses and an increased duration of the glottal opening, but current models are not sufficiently refined to predict accurately the magnitude of the effects. In any event, it seems clear that when the vocal tract is constricted to form a liquid or a glide, there is a modest influence on the glottal volume-velocity source. This influence is in a direction that tends to reduce the overall amplitude of the sound during the consonant, and thus can be considered to enhance the contrast between the syllabic peaks for vowels and the intensity minima for nonsyllabic segments.

## 3. Turbulence noise source for fricatives

The generation of turbulence noise in the vocal tract for a fricative consonant is accomplished by forming a constriction in some region along the supraglottal airways and directing the airstream against an obstacle or surface in the vocal tract. In a sense, then, the properties of the source are a direct result of interaction of the vocal-tract shape and the airflow in the vocal tract.

A typical configuration of the vocal tract for a fricative consonant is shown in Fig. 1a. The factors influencing the properties of the source of turbulence noise can be explored by examining the properties of the sound that is generated when air flows through a mechanical model like that in Fig. 1b. The rapid airflow through the downstream constriction generates turbulence noise, and if an obstacle like that shown in the figure is present in the airstream, this sound source is concentrated in the vicinity of the obstacle. The constriction at the left of the model represents the glottal opening. A number of measurements have been made of the sound generated by airflow in a model of the type shown in Fig. 1b [6].

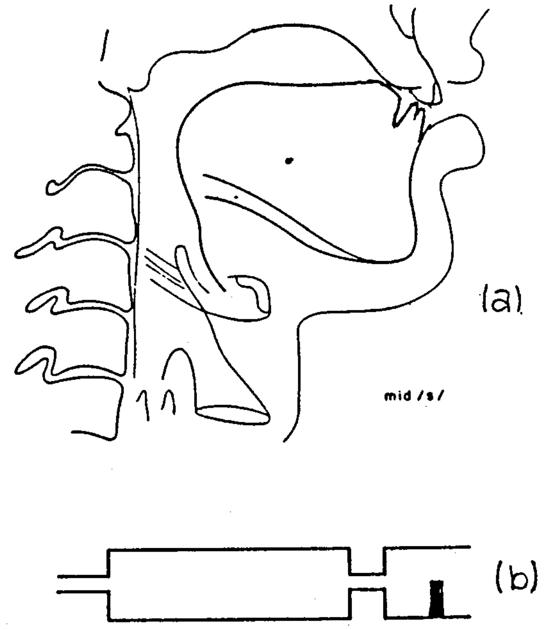

**Fig. 1(a)** Midsagittal configuration of vocal tract for a fricative consonant. **(b)** Mechanical model of fricative noise generation with an obstacle downstream from the constriction.

Figure 2 shows the spectrum of the sound pressure measured some distance from the model with and without the obstacle downstream from the constriction. The difference in level is as great as 30 dB in some frequency regions. Turbulence noise in the vocal tract is never entirely free of some obstacle or surface, but clearly large differences in source strength are obtained for different configurations of obstacles in the airstream. This fact is evidently exploited in the selection of articulatory gestures for producing feature combinations and contrasts for use in language.

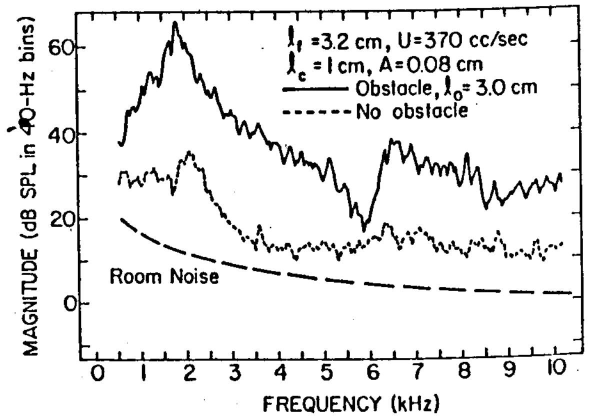

**Fig. 2** The solid curve shows the spectrum of the sound radiated from a model like that in Fig. 1(b). Front-cavity length and distance from constriction to obstacle are both about 3 cm. The dotted curve is the spectrum for the same configuration, but with the obstacle removed. (From Shadle, 1985).

The results of a series of experiments with models of the type shown in Fig. 1b, together with theoretical studies of turbulence noise generation, have led to procedures for calculating the amplitude and spectrum of the sound-pressure source $p_s$ for different values of airflow and constriction size. The relation that has been developed is

$$p_s = K_1 U^3 A^{-\frac{5}{2}} = K_1 \left(\frac{2\Delta P}{\rho}\right)^{\frac{3}{2}} A^{\frac{1}{2}} \qquad (1)$$

where $U$ = volume velocity through the constriction, $\Delta P$ = pressure drop across the constriction, $\rho$ = density of air, $A$ = cross-sectional area of the constriction, and $K_1$ is a constant that depends on the configuration of obstacles and surfaces against which the airflow impinges.

Equation (1), together with standard equations relating pressure and airflow in constricted tubes, can be used to calculate the levels of the noise sources at the glottal constriction and at the supraglottal constriction in the model of Fig. 1b as the cross-sectional area of the supraglottal constriction is manipulated. Figure 3a shows the results of this type of calculation when the cross-sectional area of the glottis is fixed at 0.2 cm². The amplitude of the source at the downstream constriction has a broad maximum for constriction sizes in the range 0.05 to 0.15 cm². Over this range, the amplitude of the turbulence noise source is relatively insensitive to the size of the supraglottal constriction. This result suggests that, in producing a fricative consonant, a speaker is not required to adjust the size of the constriction precisely in order to obtain a fixed maximum amplitude for the noise source. The figure also shows that as the supraglottal constriction becomes larger, the amplitude of the noise generated at that constriction decreases, and the noise source at the glottis becomes dominant. This glottal source in the model corresponds to aspiration noise in the vocal tract.

One other type of interaction has been ignored in estimating the source levels in Fig. 3a. Since the walls of the vocal tract are not rigid, an increase in pressure behind the constriction will cause the walls of the vocal tract to displace outwards in response to the pressure. This displacement of the walls can have an influence on the size of the constriction. In the configuration of Fig. 1a, for example, the heightened pressure behind the alveolar constriction exerts a force that causes a downward displacement of the tongue blade that is sufficient to cause an increase in the size of the constriction. If the constriction is adjusted to have a "resting" cross-sectional area in the absence of an applied subglottal pressure, then the cross-sectional area when the pressure is applied will become larger. It is possible to calculate approximately this displacement, since some data are available on which to make estimates of the mechanical compliance of the vocal-tract walls [7]. Based on these calculations, we have recomputed the curves in Fig. 3a, except now we have plotted on the abscissa the resting area of the constriction, rather than the actual area after the subglottal pressure is applied. The new curves are given in Fig. 3b. We again observe a maximum in the amplitude of the noise generated near the downstream constriction, except that now the maximum is much broader. That is, the level of the noise is even more stable in response to perturbations in the area of the constriction. This example, as well as others not discussed here, illustrate that properties of the vocal-tract walls can play a significant role in determining the characteristics of sound sources in the vocal tract.

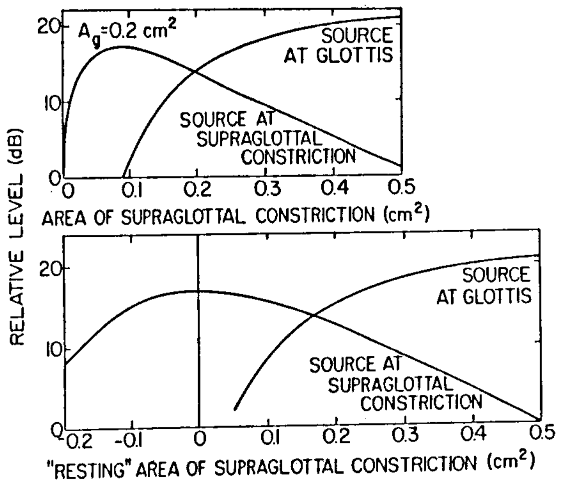

**Fig. 3(a)** Calculated levels of the turbulence noise sources at the supraglottal and glottal constrictions as a function of the area of the supraglottal constriction. The area of the glottal constriction is fixed at 0.2 cm². **(b)** Same as (a), except that the size of the constriction is modified by the presence of the intraoral pressure. The abscissa is the area that the supraglottal constriction would assume before application of the subglottal pressure.

## 4. Source characteristics for voiced fricatives

A voiced fricative consonant is produced with a supraglottal configuration similar to that for a voiceless fricative, described above. The difference, however, is that the vocal folds continue to vibrate over at least part of the time interval when the vocal tract is in the constricted configuration. If noise is to be generated at the supraglottal constriction and if glottal vibration is to continue, then the pressure in the space between the glottis and the constriction must be intermediate between the subglottal pressure and atmospheric pressure. In order to achieve this condition, the speaker must maintain a rather careful balance between the configuration of the glottis and the cross-sectional area of the supraglottal constriction.

If the average pressure in the vocal-tract volume between the glottis and the supraglottal constriction is $P_m$ and the subglottal pressure is $P_{sub}$, then the transglottal pressure is $P_{trans} = P_{sub} - P_m$, and the pressure across the constriction is $P_m$. This pressure drop $P_m$ is the value of $\Delta P$ that is used to esimate the amplitude of the turbulence noise source at the constriction, in equation (1). When the vocal folds are in a configuration appropriate for voicing, experimental observations and theoretical analysis have shown that the amplitude of the volume-velocity pulses at the glottis is roughly proportional to $P_{trans}^{\frac{3}{2}}$, i.e., proportional to $(P_{sub} - P_m)^{\frac{3}{2}}$ in this case [8]. However, when the transglottal pressure becomes less than a particular threshold value, usually considered to be about 2-3 cm $H_2O$, vocal-fold vibration is no longer maintained.

In Fig. 4 we estimate the level of the turbulence noise source at the constriction and the level of the glottal source as a function of the cross-sectional area of the supraglottal constriction. The average area of the glottis is assumed to be fixed (at 0.13 $cm^2$ in this case). The noise level is plotted relative to the maximum level that would be obtained for a voiceless fricative consonant produced with a glottal opening as indicated in Fig. 3. The amplitude of the glottal source is plotted relative to the amplitude that is obtained with the same glottal configuration but with no supraglottal constriction. When the constriction size decreases to about 0.07 $cm^2$, the transglottal pressure drops to about 2 cm $H_2O$, and vocal-fold vibration can be assumed to cease for smaller constrictions.

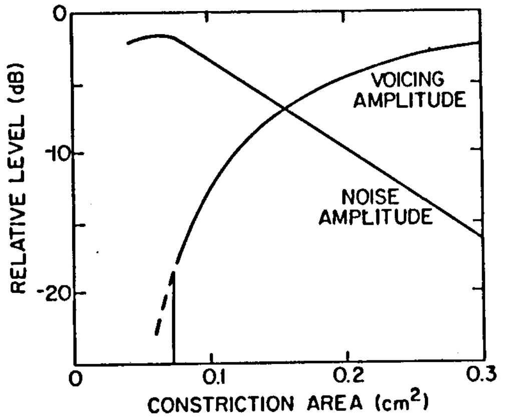

**Fig. 4** Calculated amplitudes of noise source and glottal source for a model of a voiced fricative consonant, as a function of the cross-sectional area of the supraglottal constriction. The average glottal area is assumed to be fixed at about 0.13 $cm^2$. Calculated voicing amplitude is given relative to the amplitude when the vocal tract is unconstricted, as it might be for a vowel. Noise amplitude is given relative to the amplitude for a model of a typical voiceless fricative. The vertical line at about 0.07 $cm^2$ indicates the constriction area below which vocal-fold vibration cannot be sustained.

Figure 4 indicates that there is some instability in adjusting the supraglottal constriction size. If strong voicing is to be maintained (say around a constriction size of 0.2 $cm^2$), then the amplitude of the noise source becomes small, whereas a strong noise source can only be achieved at the expense of weakened or even cessation of voicing. It often happens that voicing is maintained only over some part of the interval when the vocal tract is constricted for a voiced fricative.

Examples of measurements of the voicing amplitude and the noise amplitude are shown in Fig. 5 for the intervocalic voiceless and voiced fricatives [s] and [z]. In the case of the voiceless fricative, the amplitude of the noise remains rather stable throughout the interval, and the glottal source turns off and on rather abruptly. There is a brief reduction in noise amplitude just before voicing onset, as the constriction is released before the vocal folds have adducted to a position appropriate for voicing. For the voiced fricative, there is considerable variation in both noise amplitude and glottal source amplitude. The noise amplitude tends to be 5-10 dB less than that for the voiceless cognate, over much of the constricted interval.

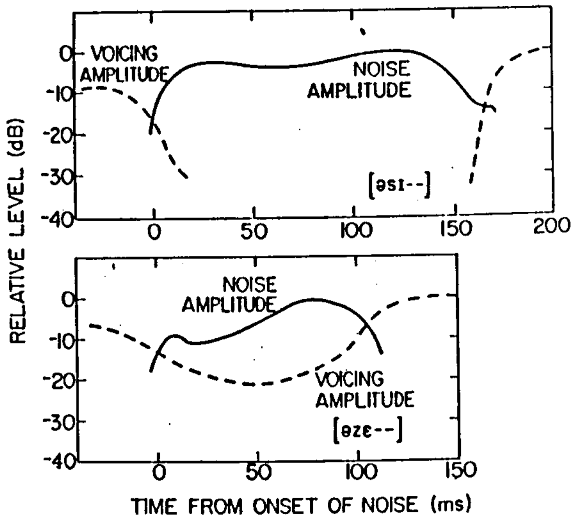

**Fig. 5** Measurements of noise amplitude (peak spectrum amplitude above 3.5 kHz) and voicing amplitude (peak spectrum amplitude at $F1$) as a function of time for an intervocalic voiceless fricative [s] (top panel) and voiced fricative [z] (bottom panel). The voicing curves are plotted with 0 dB as the maximum level, and the noise amplitude curves are plotted with the reference of 0 dB being the peak level for the voiceless cognate.

The reduced amplitude of the glottal source for the voiced fricatives [v] and [ð] has been observed for a variety of utterances by Bickley and Stevens [5]. They report an average reduction in amplitude (relative to that for a vowel) of about 10 dB for these fricatives.

## 5. Concluding remarks

When a constriction in the vocal tract becomes narrower than the minimum cross-sectional area for a vowel, there is a substantial increase in the interaction between the vocal-tract shape and the characteristics of the source. For liquids and glides, the contriction is not sufficiently narrow to produce turbulence noise, but the impedance of the constriction is sufficient to cause a decreased amplitude of the glottal source. When the constriction is made narrower, turbulence noise is generated in the vicinity of the constriction. For a fixed glottal opening, the amplitude of this noise is reasonably stable over a substantial range of supraglottal constriction sizes. Simultaneous generation of both glottal vibration and turbulence noise at a supraglottal constriction requires a rather careful adjustment of the supraglottal and glottal constrictions. It is to be expected, therefore, that one or other of the sources may not continue throughout the entire consonantal interval for a voiced fricative.

This analysis has indicated that the interaction between acoustic sources and vocal-tract shapes tends to be much greater for constricted or consonantal configurations than it is for more open vocalic configurations. We still have much to learn about the characteristics of the glottal and turbulence noise sources for different types of consonants, particularly for obstruent consonants for which there is an increased pressure in the intraoral space.

## 6. Acknowledgements

This work was supported in part by a grant from the National Institute of Neurological and Communicative Disorders and Stroke and the National Science Foundation.

## 7. References

1. M. Rothenberg. *Acoustic interaction between the glottal source and the vocal tract.* In **Vocal Fold Physiology**, K.N. Stevens and M. Hirano, eds. Tokyo: University of Tokyo Press, 305-318 (1981).
2. G. Fant. *Preliminaries to analysis of the human voice source.* **Quarterly Progress and Status Report** (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), STL-QPSR 4, 1-27 (1982).
3. K. Honda. *Relationship between pitch control and vowel articulation.* In **Vocal Fold Physiology: Contemporary Research and Clinical Issues**, D.M. Bless and J.H. Abbs, eds. San Diego: College-Hill, 286-297 (1983).
4. C.A. Bickley and K.N. Stevens. *Effects of a vocal-tract constriction on the glottal source: Experimental and modelling studies.* **J. Phonetics, 14**, 373-382 (1986).
5. C.A. Bickley and K.N. Stevens. *Effects of a vocal-tract constriction on the glottal source: Data from voiced consonants.* In **Laryngeal Function in Phonation and Respiration**, T. Baer, C. Sasaki, and K.S. Harris, eds. Boston: College-Hill, 239-253 (1987).
6. C. Shadle. **The acoustics of fricative consonants.** Technical Report 506, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge MA (1985).
7. K. Ishizaka, J.C. French, and J.L. Flanagan. *Direct determination of vocal tract wall impedance.* **IEEE Trans. on Acoustics Speech and Signal Processing, ASSP-23**, 370-373 (1975).
8. A. Bouhuys, J. Mead, D.F. Proctor, and K.N. Stevens. *Pressure-flow events during singing.* In **Sound Production in Man, Annals of the New York Academy of Sciences, 155**, 165-176 (1968).