Reprinted from THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Vol. 45, No. 6, 1544–1546, June 1969

Transillumination of the Larynx in Running Speech*

LEIGH LISKER,† ARTHUR S. ABRAMSON,‡ FRANKLIN S. COOPER, AND MALCOLM H. SCHVEY§

Haskins Laboratories, 305 East 43rd Street, New York, New York 10017

A fundamental distinction between speech sounds depends on whether the excitation is a noise source or a quasiperiodic one. The noisy or voiceless sounds are presumably produced with an opened and quiescent larynx, while for voiced sounds the larynx is closed down and in rapid oscillation. Direct evidence of this has come from motion pictures taken through the open mouth, a method obviously limited to particular sounds. Running speech can be studied by a transillumination technique. A fiber-optics bundle or a miniature incandescent bulb introduced into the laryngeal vestibule through the nose illuminates the glottis, while a photocell placed below the thyroid cartilage registers the variable light transmitted through the glottis and the tissues of the neck. The "glottograms" so obtained are compared with simultaneously recorded acoustic waveforms and intraoral air-pressure traces to determine how the voiced–voiceless distinction correlates with the closed-versus-open states of the larynx.

It has become increasingly clear to us that if we are to account in any simple way for all the acoustic features and physiological measures associated with the voiced–voiceless contrast in English, we must refer them back to events involving the larynx and to the timing of those events in relation to supraglottal gestures.[1] The work on transillumination of the larynx reported by Sonneson in 1960[2] suggested that this technique might be appropriate for obtaining data on the dynamics of laryngeal behavior during the production of those sounds, the stop and fricative consonants, for which voicing is distinctive. It would be more direct than the method of deriving inferences from acoustic analysis of the speech signal.

In brief, the technique we adopted involved directing a beam of light down onto the glottis and placing a photocell against the neck below the thyroid cartilage so as to pick up the light transmitted through the open glottis and the tissues of the neck. Careful positioning of light and photocell pickup permitted satisfactory light-level registration when the subject breathed quietly (glottis open); virtually zero light was transmitted when the breath was held (glottis closed). Variations in light intensity recorded during speech activity showed great consistency for repetitions of the same sentences. For the most part, those variations could reasonably be ascribed to variations in the area of glottal aperture. To be sure, there were occasional random fluctuations, even after the purely instrumental troubles were under control. We suppose that they reflect some sort of articulatory "noise," such as jostling of the light by the epiglottis or changes in the position of the larynx.[3]

In the first trials, the light source used was a high-intensity xenon arc whose output was conducted through a fiber-optics bundle. This was later replaced by a subminiature incandescent bulb introduced directly in the pharynx. In both cases, it was possible to achieve good placement of the light beam with minimal disturbance of articulation by going in through the nose. The transmitted light was picked up by a red-sensitive photomultiplier with a light cone. The output was stored on magnetic tape for later display

* This article is a revised version of a paper presented at the 71st Meeting of the Acoustical Society of America, Boston, 1–4 June 1966. See J. Acoust. Soc. Amer. 39, 1218 (A) (1966).

† Also located at University of Pennsylvania, Philadelphia, Pa. 19104.

‡ Also located at University of Connecticut, Storrs, Conn. 06268.

§ Also located at Columbia–Presbyterian Medical Center, New York, N. Y. 10032.

[1] L. Lisker and A. S. Abramson, "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements," Word 20, 384–422 (1964).

[2] B. Sonesson, "On the Anatomy and Vibratory Patterns of the Human Vocal Folds," Acta Otolaryngol. Suppl. 156, 1–80 (1960). See also A. Malécot and K. Peebles, "An Optical Device for Recording Glottal Adduction–Abduction during Normal Speech," Z. Phonetik, Sprachwissenschaft Kommunikationsforschung 18, 545–550 (1965).

[3] This uncertainty ought to be largely obviated by simultaneous direct viewing of the glottis through a coherent fiber-optics bundle. See M. Sawashima and H. Hirose, "New Laryngoscopic Technique by the Use of Fiber Optics," J. Acoust. Soc. Amer. 43, 168–169 (L) (1968).

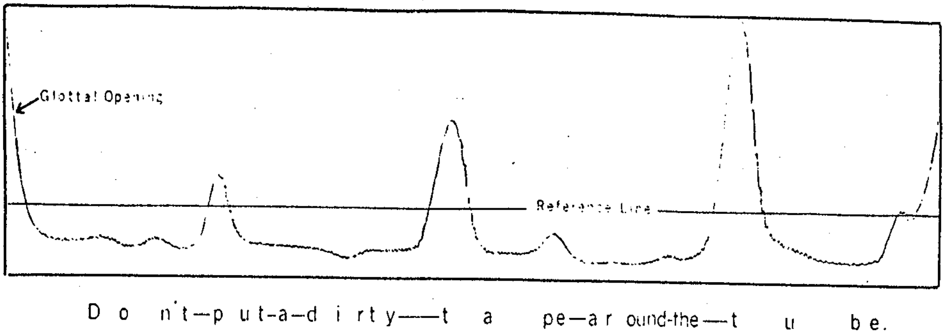

FIG. 1. Glottogram of an English sentence. Greater transillumination, thus larger opening of the glottis, is shown by upward movements of the trace.

in the form of glottograms such as the one shown in Fig. 1. The trace labeled "glottal opening" shows the variation in light registered by the photocell as the subject pronounced the sentence "Don't put a dirty tape around the tube." Gross opening of the glottis is indicated for the four voiceless stops in the sentence. They are /p/, in *put* and *tape*, and /t/, in *tape* and *tube*. The /t/ expected at the end of *don't* was apparently manifested as a shortening of the nasal consonant and, perhaps, a glottal stop. The intervocalic consonant at the end of *put*, which appears before the unstressed word *a*, is a voiced flap rather than a voiceless stop, as is also the medial consonant in the word *dirty*. The intervals of small aperture are seen as low-amplitude modulation at the frequency of the voice fundamental. We can also see how the glottis shifts from the open position of quiet breathing to a closed one just before the onset of speech, and how it returns to the open position after the cessation of speech activity. In addition to the glottogram, the display included oscillograms of the simultaneously recorded outputs of air and throat microphones and a pressure transducer placed so as to measure intraoral air pressure. Though these records, not shown in Fig. 1, have not yet been completely analyzed, they have been useful in enabling us to correlate features of the glottogram with other articulatory events. The preliminary data described here represent the first crude analysis of recordings from a single subject repeating a set of sentences selected to include all the English stop and fricative consonants in a variety of positions. In analyzing the glottograms and other records, we have begun by classifying these sounds simply on the basis of (1) whether there was an indication of gross opening of the glottis, and (2) whether there was any observable interruption of pulsing. In making this rough classification, we limited attention, for the time being, to the stops and fricatives in utterance-medial position between vowels and other characteristically voiced sounds. Moreover, we excluded from consideration items of ambiguous linguistic status, such as stops after /s/ and intervocalic flaps.[4]

Figure 2 gives the distributions of stop and fricative phonemes with respect to these two articulatory classes, considered separately. We found it advisable, in the case of the voiceless stops and the voiced fricatives, to distinguish between items on the basis of whether they precede stressed or unstressed syllabics. In Fig. 2, we see that, in the case of the stops, the voiced and voiceless categories are almost perfectly sorted on the basis of presence versus absence of either glottal opening or interruption of pulsing. For /ptk/, especially in unstressed position, there are several items whose glottograms fail to show any opening of the glottis, and others for which there is no apparent break in pulsing. For the fricatives, the picture is slightly more complicated. Here the voiceless set shows glottal opening and interruption of pulsing for virtually all items, independent of stress condition; but in the voiced set, too, there is a high percentage of items with open glottis and a few items with interruption of pulsing. In comparison with the stops, the fricatives appear to differ a bit more sharply with respect to pulsing interruption than glottal opening. That our records should show such a high incidence of glottal opening for the voiced fricatives as compared with the voiced stops is not unreasonable, if we suppose that a well-formed fricative requires audible turbulence, and that the airflow needed for

			Glottis		Pulsing	
			% Open	% Closed	% Broken	% Unbroken
STOPS	/bdg/ (N=56)		6	94	5	95
	/ptk/	Stressed (N=43)	96	4	94	6
		Unstressed (N=57)	84	16	82	18
FRICATIVES	/vðzž/	Stressed (N=21)	86	14	10	90
		Unstressed (N=47)	70	30	0	100
	/fθsš/ (N=72)		99	1	99	1

FIG. 2. Distribution of stops and fricatives with respect to glottal opening and continuity of pulsing.

[4] Identifying the stops after /s/ in such words as *spill*, *still*, *skill* with voiceless /ptk/ has been a linguistic convention, but these stops, once the /s/ frication has been removed, are identified as /bdg/ by speakers of English; see J. Lotz, A. S. Abramson, L. J. Gerstman, F. Ingemann, and W. J. Nemser, "The Perception of English Stops by Speakers of English, Spanish, Hungarian and Thai: a Tape-Cutting Experiment," Language and Speech 3, 71–77 (1960). As for intervocalic flaps, the medial consonant in common American English productions of words like *letter* is not readily identified as either /t/ or /d/; see A. Malécot and P. Lloyd, "The /t/: /d/ Distinction in American Alveolar Flaps," Lingua 19, 264–272 (1968).

this is most easily supplied when the glottis is at least partially open. In this connection, it is interesting to note that there are differences within the class of voiced fricatives: almost exclusively, the fricatives with higher levels of noise intensity, primarily /z/ and /ž/, show opening of the glottis as the normal accompaniment of the oral constriction.

If, instead of considering the two conditions of glottal aperture and pulsing continuity separately, we take their four possible combinations of conditions together, then we find the consonants distributed as shown in Fig. 3. It would seem that this represents some slight improvement in the resolution of voiced and voiceless categories. For example, all but two of 12 cases of /ptk/ with unbroken pulsing were produced with the glottis in an open position. For 98% of a total of 321 consonant tokens recorded, the situation can be summarized as follows: (1) The voiceless stops were produced with either opening of the glottis or interruption of pulsing, or both; (2) the voiced consonants, both stop and fricative, were produced without interruption of pulsing; and (3) the voiced fricatives were produced with either an open or a closed glottis; this distinction generally matched the difference between fricatives of high and low noise intensity. From our classification on the basis of glottal aperture and pulsing features, we should then put fricatives of the second kind together with the voiced stops.

By and large, the preliminary analysis of our transillumination records is encouraging. To be sure, there are certain questions raised by the data that can only be resolved after the observational technique has been developed much further. Thus, for the moment, we can do little more than make an educated guess as to just how the opening of the glottis is performed under each of the conditions of pulsing in running speech. In the near future, we expect to determine this for English and certain other languages of interest to us by means

GLOTTIS + PULSING

		OPEN		CLOSED	
		% Broken	% Unbroken	% Broken	% Unbroken
STOPS	/bdg/ (N=66)	3	3	2	92
	/ptk/ (N=106)	79	10	8	2
FRICATIVES	/vðzž/ (N=68)	3	72	0	25
	/fθsš/ (N=72)	97	1	1	0

FIG. 3. Combinations of conditions of glottal opening and pulsing for stops and fricatives.

of motion pictures of the glottis taken through a coherent fiber-optics bundle. We are encouraged in this expectation by the results of recent Japanese studies.[5] Nevertheless, much information is promised by our transillumination records as they stand, and awaits only the completion of a detailed quantitative study of the durations, magnitudes, and relative timing of the laryngeal gestures of opening and interruption of pulsing. But even on the basis of our first crude analysis, it appears that much of the complexity encountered in studying the voiced–voiceless difference among English stops and fricatives is obviated by referring to the relatively simple differences in laryngeal behavior.[6]

ACKNOWLEDGMENT

The research reported here was supported in part by grants from the National Institute for Child Health and Human Development and the National Institute for Dental Research.

[5] M. Sawashima, H. Hirose, S. Kiritani, and O. Fujimura, "Articulatory Movements of the Larynx," in Proc. Int. Congr. Acoust., 6th, Tokyo (1968), Paper B-1-1.

[6] The perceptual relevance of the acoustic consequences of variations in laryngeal timing has been demonstrated for English, Spanish, and Thai; see A. S. Abramson and L. Lisker, "Voice Onset Time in Stop Consonants: Acoustic Analysis and Synthesis," in Proc. Int. Cong. on Acoustics, 5th, Liege (1965), Vol. Ia, Paper A51; L. Lisker and A. S. Abramson, "The Voicing Dimension: Some Experiments in Comparative Phonetics," in Proc. Int. Cong. of Phonetic Sci., 6th (to be published).