Phonetica
Editor: E. ZWIRNER, Münster in Westfalen
S. KARGER – BASEL/NEW YORK (Printed in Switzerland)
SEPARATUM

Phonetica *22:* 193–201 (1970)

Observing Laryngeal Adjustments during Running Speech by Use of a Fiberoptics System

M. SAWASHIMA, A. S. ABRAMSON, F. S. COOPER and L. LISKER

Research Institute of Logopedics and Phoniatrics, University of Tokyo, Hongo, Tokyo; Department of Linguistics, University of Connecticut, Storrs, Conn.; Haskins Laboratories, New Haven, Conn., and Department of Linguistics. University of Pennsylvania, Philadelphia, Pa.

Observing Laryngeal Adjustments during Running Speech by Use of a Fiberoptics System[1]

M. SAWASHIMA, A. S. ABRAMSON, F. S. COOPER and L. LISKER

Research Institute of Logopedics and Phoniatrics, University of Tokyo, Hongo, Tokyo; Department of Linguistics, University of Connecticut, Storrs, Conn.; Haskins Laboratories, New Haven, Conn., and Department of Linguistics, University of Pennsylvania, Philadelphia, Pa.

During the past decade, several techniques have been developed by which changes in glottal aperture during running speech may be observed indirectly. The techniques involve transillumination [SONESSON, 1960], measurement of electrical impedance across the glottis [FABRE, 1957], ultrasonic monitoring [ASANO, 1968; MINIFIE, KELSEY and HIXON, 1968] and X-ray motion pictures [HOLLIEN, 1965]. In earlier work at Haskins Laboratories the transillumination technique was used [LISKER, ABRAMSON, COOPER and SCHVEY, 1969]. The transillumination data provided valuable information on laryngeal maneuvers during speech. The data, however, raised some questions which might only be solved by visual examination of the glottis. We then began to use a flexible fiberoptics bundle to examine the glottis visually and photographically. The technique was developed at the University of Tokyo [SAWASHIMA and HIROSE, 1968].

The fiberoptics bundle consists of two groups of glass fibers, the image guide and the light guide. A 16 mm cinecamera was attached to the eye piece of the image guide. The fiber bundle, with an outside diameter of about 5 mm, was inserted through the nose and positioned in the hypopharynx at the level of the tip of the epiglottis to get a

[1] This article is an expanded version of papers presented orally at three meetings held in November and December, 1969, viz: 78th Meeting of the Acoustical Society of America, San Diego, Motion pictures of the vocal folds in speech; 45th Annual Convention of the American Speech and Hearing Association, Chicago, Observation of laryngeal adjustments during running speech by use of a fiberoptics system; 1969 Annual Meeting of the Modern Language Association of America, Denver, Viewing the vocal folds in running speech.

good view of the glottis. Surface anesthesia was necessary for the nasal cavity and the epipharynx.

Three American English talkers were used as subjects. Short sentences containing both voiced and voiceless consonants were read aloud. The cinecamera was driven by a synchronous motor at a frame rate of 60/sec. Speech signals were recorded on a tape recorder together with synchronization time marks. The present paper is a report of findings obtained so far for one subject, the second author.

Figure 1 shows selected frames of the motion picture for the sentence *Rub Billy's head with this towel.* Each frame is correlated with the proper point in the sound spectrogram. A narrow-band trace is displayed above the wide-band pattern to show the voicing during speech. The symbols at the bottom are those of a broad phonetic transcription of the utterance; they are aligned approximately with appropriate spans of the spectrogram.

The left-most frame shows the larynx in respiratory position before the utterance. We see the anterior commissure of the vocal folds at the upper left corner of the frame, and the wide open glottis and vocal processes of the arytenoid cartilages towards the lower right corner.

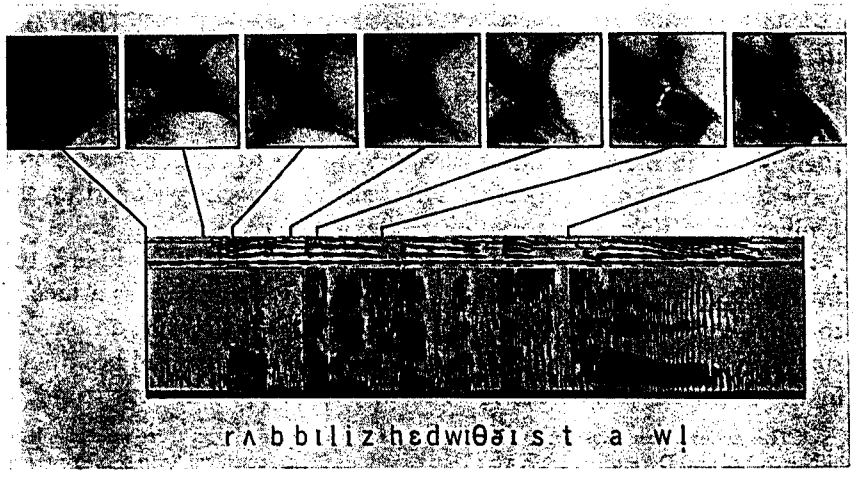

Fig. 1. Rub Billy's head with this towel. From top to bottom: Selected views of glottis in running speech. First few harmonics of narrow-band spectrogram. Wide-band spectrogram. Broad phonetic transcription.

The next frame shows the situation immediately before voice onset. The larynx is in phonatory position, the arytenoids are closed, while a narrow opening is seen along the membranous portion of the glottis. The next frame shows almost the same position of the larynx, in which the blurred edges of the vocal folds indicate vibratory motion.

The 4th frame is for the labial closure of the [b] of *Rub Billy's.* Here also, the larynx is in phonatory position with vibrating vocal folds. Its appearance is almost the same as in the next frame for the following vowel. The 6th frame is for the transition from [z] to [h] of *Billy's head.* The glottis is open with separated arytenoids. The sharp image of the vocal fold edges indicates the cessation of vibration. The last frame shows the glottis just before the release of [t] for *towel.* The opening of the glottis is as large as, or a little larger than[2], the opening during the transition from [z] to [h].

Figure 2 shows selected frames for voiced consonants in the sentence *His zoology book is available in paperback.* For the left-most two frames, we see the glottis for the vowel [ɪ] and fricative [z] respectively. Both of them show closed arytenoids with vocal fold vibration. The third frame is for the affricate [ǰ] of *zoology.* Here the arytenoids are

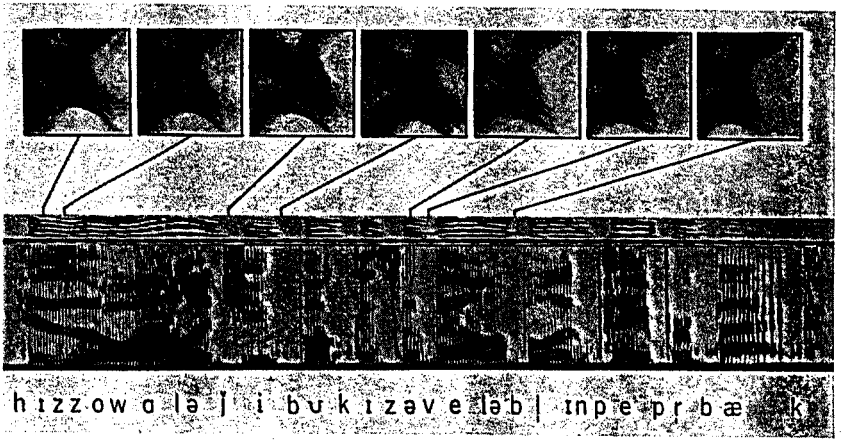

Fig. 2. His zoology book is available in paperback.

[2] As a result of articulatory jostling of the fiberoptics bundle and movements of the larynx, great precision of measurement of the size of the glottal aperture is impossible with the present technique.

separated slightly and the glottis is open. This position is quite similar to the one in the next frame, which is for the [b]-closure of *book*. In both cases voicing dies away during the closure period.

The 5th and 6th frames show the glottis for the vowel [ə] and fricative [v] of *available* respectively. No difference is observed between the vowel and [v]. The final frame is for the [b]-closure of *available*. The arytenoids for this [b] are in closed position, almost the same as for the vowel and [v], while voicing ceases during the closure period, so that a narrow opening of the glottis is visible.

Figure 3 shows two different openings of the glottis for the same phoneme, the voiceless stop /p/; the first example has an aspirated release, while the second is unaspirated. The left frame is the glottis for the latter part of the long [p]-closure in the sentence *Keep peace on the campus*. The glottis is wide open with separated arytenoids. A considerable length of voicing lag with turbulent noise after the release

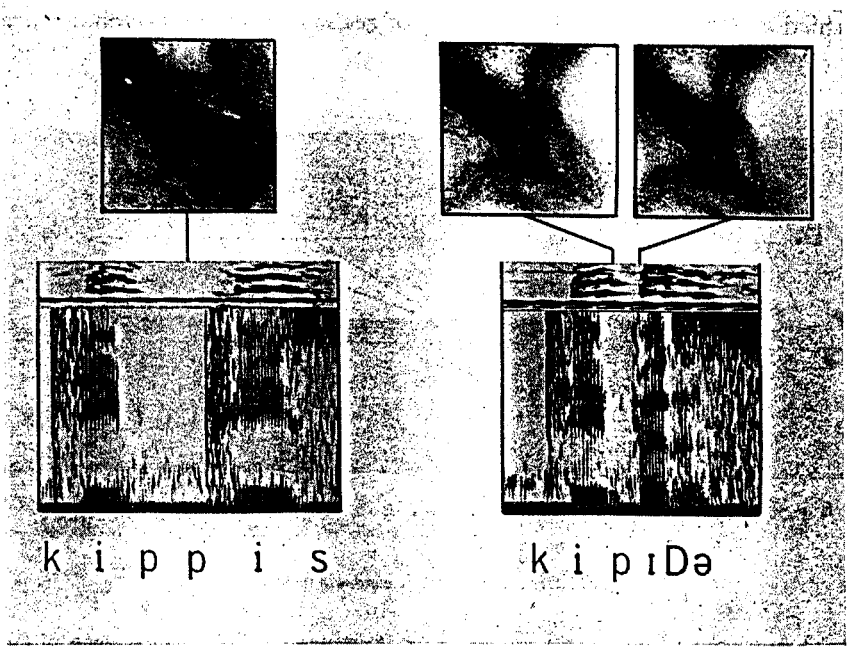

Fig. 3. Keep peace... and *Keep it a...*

is shown in the spectrogram[3]. The next frame is for the [p]-closure in the sentence *Keep it a secret*[4]. The opening of the glottis is very small and the position of the arytenoids is almost the same as in the next frame, which is for the following vowel.

In figure 4, we see three different openings of the glottis for the same phoneme /k/. The left-most frame shows the glottis for the [k]-closure of *campus* in the sentence *Keep peace on the campus*. The arytenoids are wide apart with a large glottal opening. The sound spectrogram shows a considerable length of voicing lag with turbulent noise after release appropriate to the auditory impression of aspiration[5]. The second frame is for the [k]-closure of *book is* in the sentence *His zoology book is available in paperback*. In this case, the arytenoids are slightly separated and the glottal opening is smaller than that in the left frame. The spectrogram of this unaspirated stop shows less voicing lag. The

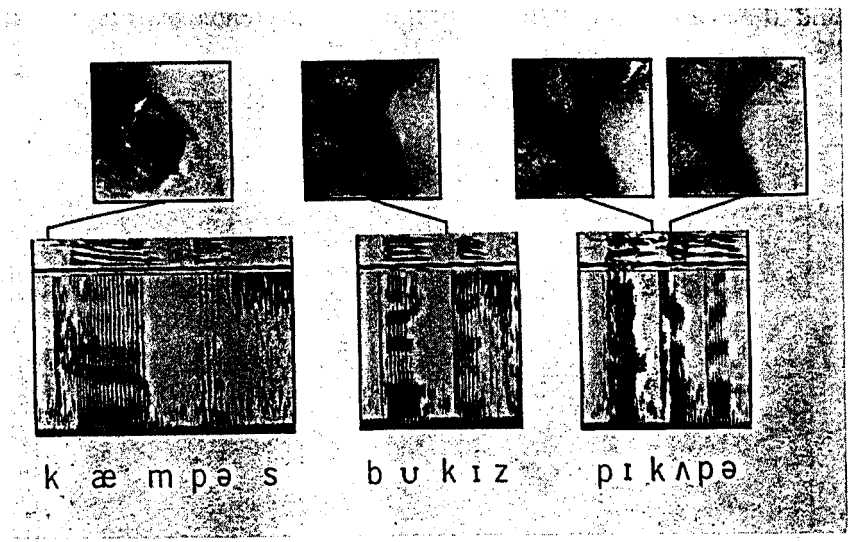

Fig. 4. ...campus; ...book is...; pick up a...

[3] For extensive data on voice timing relative to stop-release as a category differentiator in English and a number of other languages, see LISKER and ABRAMSON [1964, 1967].

[4] The apical stop in *it* is morphophonemically /t/. The [D] in the transcription in figure 3 indicates that this consonant was in fact either [d] or a voiced flap, commonly found in American English in intervocalic position before an unstressed vowel.

[5] This apparent correlation of aspiration with large glottal aperture is in agreement with recent X-ray data on Korean [KIM, 1970].

next frame is for [k] of *pick up* in the sentence *Pick up a pack of pipe tobacco*. For this [k] the arytenoids remain closed very much as in the phonatory position found for the following vowel, shown in the last frame. The glottal opening for this unaspirated [k] is very small and limited to the membranous portion of the glottis.

Figure 5 shows the glottis for voiceless /s/ in different phonological environments. The left frame is for word-initial stressed [s] of *secret* in the sentence *Keep it a secret*. The right frame is for the word-final unstressed [s] of *peace* in the sentence *Keep peace on the campus*. Both of them show a considerable width of glottal opening with separation of the arytenoids. The opening for the stressed [s] seems to be a little larger than that for the unstressed one; however, the difference is not so pronounced as for the voiceless stop.

At this point in our study, we can come to several tentative conclusions: (1) Variations in glottal opening occur in running speech, and these are effected by controlling the arytenoid cartilages. (2)

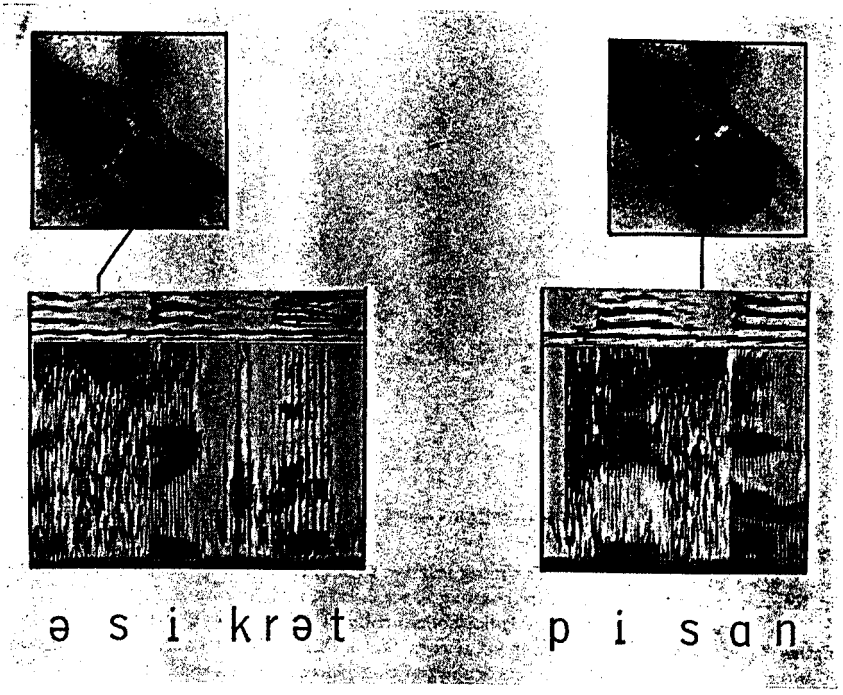

Fig. 5. . . . a secret and *. . . peace on. . .*

Speech sounds with a predominant non-transient noise are produced with opening of the arytenoids. (3) Speech sounds having predominant voicing are produced without opening of the arytenoids. (4) Other speech sounds are produced with a variety of glottal openings, ranging from those for which there is clearly a separation of the arytenoids to those which are not distinguishable from phonatory position. Such sounds are the voiceless unaspirated stops, certain of the voiced fricatives, and those varieties of English /b, d, g/ in which voicing is interrupted.

In a sequel to the present progress report, we expect to present an extensive analysis of the film footage for all three speakers used in the experiment. We are now examining appropriate frame-sequences for stop and fricative consonants, for the following features: (1) Opening and closing movements of the arytenoid cartilages. (2) Interruption and resumption of vocal-fold vibration. (3) Maximum width of glottal aperture. (4) Width of glottal aperture at the time of oral release of the consonant.

The perceptual relevance of variations in voice timing, at least for utterance-initial stops, has been demonstrated [Lisker and Abramson, 1970; Abramson and Lisker, 1970], but the question remains as to whether these timing differences are under the direct control of temporal instructions to the glottis or, as Kim [1970] would have it, instructions controlling the size of the glottal aperture. If we understand a recently proposed model correctly [Slis, 1970], it is not direct management of the larynx that is relevant but rather some kind of high-level control effected by means of certain muscles of the pharynx[6]. We hope that our continued research into laryngeal adjustments will help clarify these matters.

Acknowledgement

The research reported here was supported mainly by the National Institute of Dental Research, Grant DE-01774.

[6] See Slis [1970] for references to earlier versions of the model proposed by him and his coll. Their data and arguments do not convince us that the main mechanism underlying voicing distinctions is not in the control of the intrinsic muscles of the larynx. Of course, present knowledge of speech physiology does not enable us to deny that some non-laryngeal muscles may be involved at least marginally. For example, it is possible that the pharynx is sometimes actively expanded to maintain a transglottal pressure imbalance during voiced consonants.

Summary

Earlier findings on the acoustic features of voicing distinctions in consonants and underlying laryngeal adjustments have raised certain questions that can be answered only by direct observation of the glottis during running speech. In the present study this was done for American English without interfering with articulation by inserting a coherent fiberoptics bundle through the nose into the pharynx and coupling its outer end to a cinecamera set to run at 60 frames a second. From data on a single speaker, it appears that certain classes of sounds may be distinguished by whether or not the arytenoid cartilages move apart. Some consonant classes, particularly voiceless unaspirated stops before unstressed vowels, show some variability in this respect.

Zusammenfassung

Beobachtungen zu laryngalen Vorgängen in fortlaufender Rede mit Hilfe eines fiberoptischen Systems

Frühere Befunde zu den akustischen Merkmalen der Stimmhaftigkeitsunterscheidungen in Konsonanten in Abhängigkeit von laryngalen Vorgängen haben Fragen aufgeworfen, die nur durch direkte Beobachtung der Glottis in fortlaufender Rede beantwortet werden können. In der vorliegenden Untersuchung wurde dies für das amerikanische Englisch durchgeführt, wobei durch Einführen eines kohärenten fiberoptischen Bündels durch die Nase in den Rachenraum und durch Verbinden des äußeren Endes mit einer 60 Bilder in der Sekunde aufnehmenden Filmkamera die Artikulation nicht beeinträchtigt wurde. Aus dem Datenmaterial eines einzigen Sprechers läßt sich tentativ folgern, daß bestimmte Lautklassen dadurch unterschieden werden, ob die Stellknorpel sich öffnen oder nicht. Einige Konsonantenklassen, besonders stimmlose unaspirierte Verschlußlaute vor unbetonten Vokalen, zeigen eine gewisse Veränderlichkeit in dieser Hinsicht.

Résumé

Observation à l'aide d'un guide optique du travail laryngal dans le discours suivi

Des constatations antérieures relatives aux caractéristiques acoustiques de la sonorité des consonnes et à leur relation avec le travail laryngal ont soulevé des problèmes qui ne peuvent être résolus que par l'observation directe de la glotte pendant le discours suivi. Pour la présente étude cela a été réalisé sur l'anglo-américain: grâce à l'introduction par le nez d'un guide optique de lumière cohérente glissé dans le pharynx et relié à une caméra réalisant 60 images à la seconde, l'articulation n'a pas été gênée. Les faits observés sur un seul locuteur permattraient de conclure que certaines catégories de sons se distinguent les unes des autres par l'ouverture ou la non-ouverture du cartilage aryténoïde. Quelques catégories de consonnes, en particulier les occlusives non aspirées suivies d'une voyelle inaccentuée, montrent à cet égard une certaine variabilité.

References

Abramson, A. S. and Lisker, L.: Discriminability along the voicing continuum: Cross-language tests; Proc. 6th int. congr. phon. sci., Prague 1967, pp. 569–573 (Academia, Prague 1970).

Asano, H.: Application of the ultrasonic pulse-method to the larynx (in Japanese). J. Otolaryng. Jap. *71:* 895–916 (1968).

Fabre, P.: Un procédé électrique percutané d'inscription de l'accolement glottique au cours de la phonation: glottographie de haute fréquence; premiers résultats. Bull. Acad. nat. Méd. *3–4:* 66–69 (1957).

Hollien, H.: Stroboscopic laminagraphy of the vocal folds. Proc. 5th int. congr. phon. sci., Münster 1964, pp. 362–364 (Karger, Basel/New York 1965).

Kim, C.-W.: A theory of aspiration. Phonetica *21:* 107–116 (1970).

Lisker, L. and Abramson, A. S.: A cross-language study of voicing in initial stops: Acoustical measurements. Word *20:* 384–422 (1964).

Lisker, L. and Abramson, A. S.: Some effects of context on voice onset time in English stops. Speech *10:* 1–28 (1967).

Lisker, L. and Abramson, A. S.: The voicing dimension: Some experiments in comparative phonetics. Proc. 6th int. congr. phon. sci., Prague 1967, pp. 563–567 (Academia, Prague 1970).

Lisker, L.; Abramson, A. S.; Cooper, F. S., and Schvey, M. H.: Transillumination of the larynx in running speech. J. acoust. soc. Amer. *45:* 1544–1546 (1969).

Minifie, F. D.; Kelsey, C. A., and Hixon, T. J.: Measurement of vocal fold motion using an ultrasonic Doppler velocity monitor. J. acoust. soc. Amer. *43:* 1165–1169 (1968).

Sawashima, M. and Hirose, H.: New laryngoscopic technique by the use of fiber optics. J. acoust. Soc. Amer. *43:* 168–169 (1968).

Sli, I. H.: Articulatory measurements on voiced, voiceless and nasal consonants. Phonetica *21:* 193–210 (1970).

Sonesson, B.: On the anatomy and vibratory patterns of the human vocal folds. Acta otolaryngol., suppl. *156* (1960).

Authors' addresses: M. Sawashima, Research Institute of Logopedics and Phoniatrics, University of Tokyo, *Hongo, Tokyo* (Japan); A. S. Abramson, Dept. of Linguistics, University of Connecticut, *Storrs, CT 06268*; F. S. Cooper, Haskins Laboratories, 270 Crown Street, *New Haven, CT 06510*; L. Lisker, Dept. of Linguistics, University of Pennsylvania, *Philadelphia, PA 19104* (USA)

Laryngeal timing in Karen obstruents. Producing speech: Con-temporary issues: For Katherine Safford Harris, ed. by Fredericka Bell-Berti and Law-rence Raphael, 155â€“65. Woodbury, NY: American Institute of Physics Press.Â 1970. Observing laryngeal adjustments during running speech by use of a fiberoptics system. Phonetica 22.193â€“201. DOI: 10.1159/000259320. During normal deglutiition, a bolus of food triggers the swallowing reflex as it passes in region behind tongue -Larynx elevates, epiglottis drops down to cover aditus, aryepiglottic folds tense by action of aryepiglottic muscle and VF are abducted. Abdominal fixation. Process of impounding air in thorax to stabilize the torso and provide muscles with a structure upon which to pull -Plays important role in childbirth, defecation and vomiting.Â Property of a material that causes it to return to its original shape after being displaced (restoring force).