



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

An artificial voicing waveform for laryngectomees

Andersen, Jørgen Bach; Langvad, Bjarne; Møller, Henrik; Rold, Ove

Published in:
Electroacoustic Analysis and Enhancement of Alaryngeal Speech

Publication date:
1980

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Andersen, J. B., Langvad, B., Møller, H., & Rold, O. (1980). An artificial voicing waveform for laryngectomees. In A. Sekey (Ed.), *Electroacoustic Analysis and Enhancement of Alaryngeal Speech* (pp. 198-205). Charles C Thomas Publisher.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

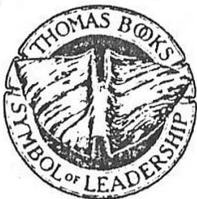
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

ELECTROACOUSTIC ANALYSIS AND ENHANCEMENT OF ALARYNGEAL SPEECH

Edited by

ANDREW SEKEY, Ph.D.

*Department of Electrical and Computer Engineering
Computer Systems Laboratory
University of California
Santa Barbara, California*



CHARLES C THOMAS • PUBLISHER
Springfield • Illinois • U.S.A.

Published and Distributed Throughout the World by
CHARLES C THOMAS • PUBLISHER
2600 South First Street
Springfield, Illinois, 62703 U.S.A.

This book is protected by copyright. No part of it
may be reproduced in any manner without written
permission from the publisher.

© 1982 by **CHARLES C THOMAS • PUBLISHER**
ISBN 0-398-04547-X

Library of Congress Catalog Card Number: 81-9082

*With THOMAS BOOKS careful attention is given to all details of
manufacturing and design. It is the Publisher's desire to present books that are
satisfactory as to their physical qualities and artistic possibilities and appro-
priate for their particular use. THOMAS BOOKS will be true to those laws of
quality that assure a good name and good will.*

Library of Congress Cataloging in Publication Data

Main entry under title:

Electroacoustic analysis and enhancement of alaryngeal speech.

Based upon an interdisciplinary conference held on June 26-27, 1980 at the
University of California, Santa Barbara.

Bibliography: p.

Includes index.

1. Speech, Alaryngeal—Congresses. 2. Electro-acoustics—Congresses. 3. Ar-
tificial larynx—Congresses. 4. Laryngoplasty—Congresses. I. Sekey, Andrew.

RF540.E36

617'.53301

81-9082

ISBN 0-398-04547-X

AACR2

Printed in the United States of America
PS-R-1

AN ARTIFICIAL VOICING WAVEFORM FOR LARYNGECTOMEES

J. BACH ANDERSEN, B. LANGVAD, H. MØLLER, AND O. ROLD

ABSTRACT

The well-known mechanical artificial larynx is placed externally against the lower throat and excites, through vibrations, the vocal-tract resonances. Not all laryngectomees are able to use this aid and for patients in intensive care it may be impossible. A compact speech aid consisting of a portable sound generator and a thin plastic tube has been developed. A periodic signal, stored in a memory, excites a small loudspeaker, which again excites the tube. The tube is placed in the back of the mouth. The device is in principle similar to earlier mechanical devices. The new idea involves the possibility of optimizing the waveform for maximum intelligibility, partly compensating for the nonoptimized position of the source. For maximum power output, tube resonances are utilized. In developing the waveform, speech tests and spectrograph analysis have been used.

BACKGROUND

THE MOST-USED SPEECH aid for persons without a larynx is the mechanical vibrator, which is externally pressed against the neck tissues. There are several situations where this device is difficult to use. One of these is in the case of older people, where the transmission through the flesh impairs

Mr. Niels Pedersen is gratefully acknowledged for the practical construction of the prototype, the phonetics laboratory of Aarhus University for the loan of their spectrograph, and the ENT department of Aalborg Hospital South for their support. The work was supported by Aalborg City's Foundation for Medical Research.

an efficient coupling to the oral cavities. Another is the situation in the hospital, where patients in intensive care have temporarily lost the use of their larynx and are unable to use a mechanical device. In the latter situation there is an urgent need for communication, which requires a simple apparatus to be used without extensive training.

The device to be described rests on the simple idea of putting an external sound source into the mouth to replace the lost larynx or, for patients in intensive care, the blocked larynx. Since the remainder of the person's articulatory system is intact, he just "speaks" as usual as if he had a larynx. Earlier versions of this device exist, e.g. in the form of a pipe with a small mechanical vibrator in the pipe head. This new version relies on modern electronic circuitry, which makes it relatively easy to create various spectral distributions and various pitch heights.

In the following, the design of the apparatus is described in detail, and some results and performance tests are discussed.

DESCRIPTION OF APPARATUS

A photo of a prototype version is shown in Figure 10-1 and in a block diagram in Figure 10-2. The system generates a periodic

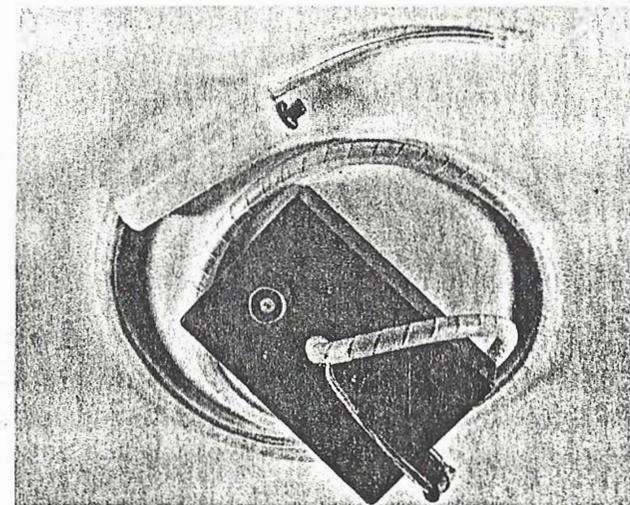


Figure 10-1. Photograph of prototype voice generator.

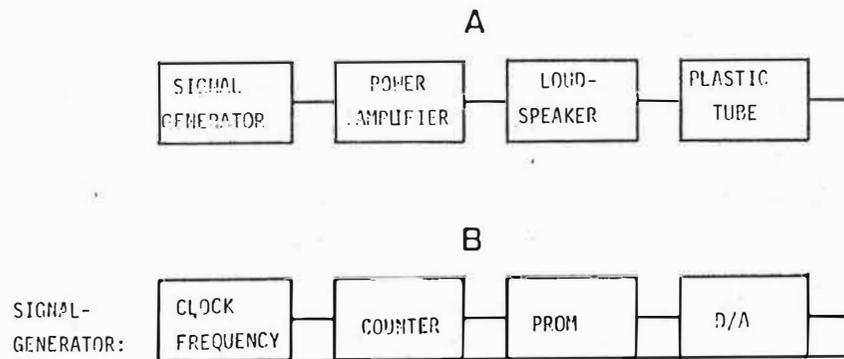


Figure 10-2. (A) Block diagram of the system; and (B) signal-generating system.

signal with a time period corresponding to a pitch of 120 Hz for men and 240 Hz for women. The time sequence within a period is stored in a PROM as indicated; the time variation of the signal will be discussed later. After a digital-to-analogue conversion, a power amplifier excites a small loudspeaker, which again excites a 1.5 m long plastic tube. A periodic sound signal with a prescribed power spectrum is emitted from the end of the tube, which is put into the mouth. The diameter of the tube is only about 5 mm, and it creates only minor articulation problems to have the tube in the mouth. A small switch in the hand turns the power on or off, but the pitch and level of the signal are constant.

THE ACOUSTICAL SYSTEM

Two design criteria are important for the acoustical system, by which we mean the loudspeaker and tube. The system must be power efficient, since we are using a battery-driven energy source,

and it must be able to cover the necessary frequency range. The system and its electrical analogue are shown in Figure 10-3, where the tube is modelled by a transmission line. It is important that the loudspeaker be shielded, since a direct "unmodulated" signal from the box will have a deteriorating effect on the intelligibility of speech. One of the problems with the mechanical vibrator is that the articulated sound must compete with the direct signal from the vibrator.

Analysis of the system shows that the tube resonances can be used advantageously since we have a periodic source. The lowest frequency (120 Hz for men) corresponds to a tube length of 1.50 m for half a wavelength. The details of the analysis are not given here; however, the conclusions derived from data analysis are that the membrane area should be small in order to give a high-resonance frequency. Furthermore, the $B \times l$ product of the loudspeaker should be high. Figure 10-3C shows the transfer function of the system from the voltage input to received sound pressure 10 cm from the tube ending.

Although the impedance of the load changes when the tube is in the mouth, analysis and experience show that the mismatch from the generator to the load is so large that we can assume that the flow at the tube ending is effectively independent of acoustical loading, i.e. independent of the position in the mouth.

THE ELECTRICAL SYSTEM

The electrical system gives the input to the loudspeaker as previously described, the critical point being the shape of the signal. If we know the required spectrum at the tube ending, it is a simple matter to find the electrical signal as a function of time. The problem is, of course, that the complete system includes the oral cavities of the patient, taking into account that the source is displaced relative to the natural position of the larynx.

The first attempt to find the optimum spectrum is based on the philosophy that the vowels should be correct on average, since they are the main components depending on the spectrum of the vocal cords.

The spectrum has been found by a substitution method, where

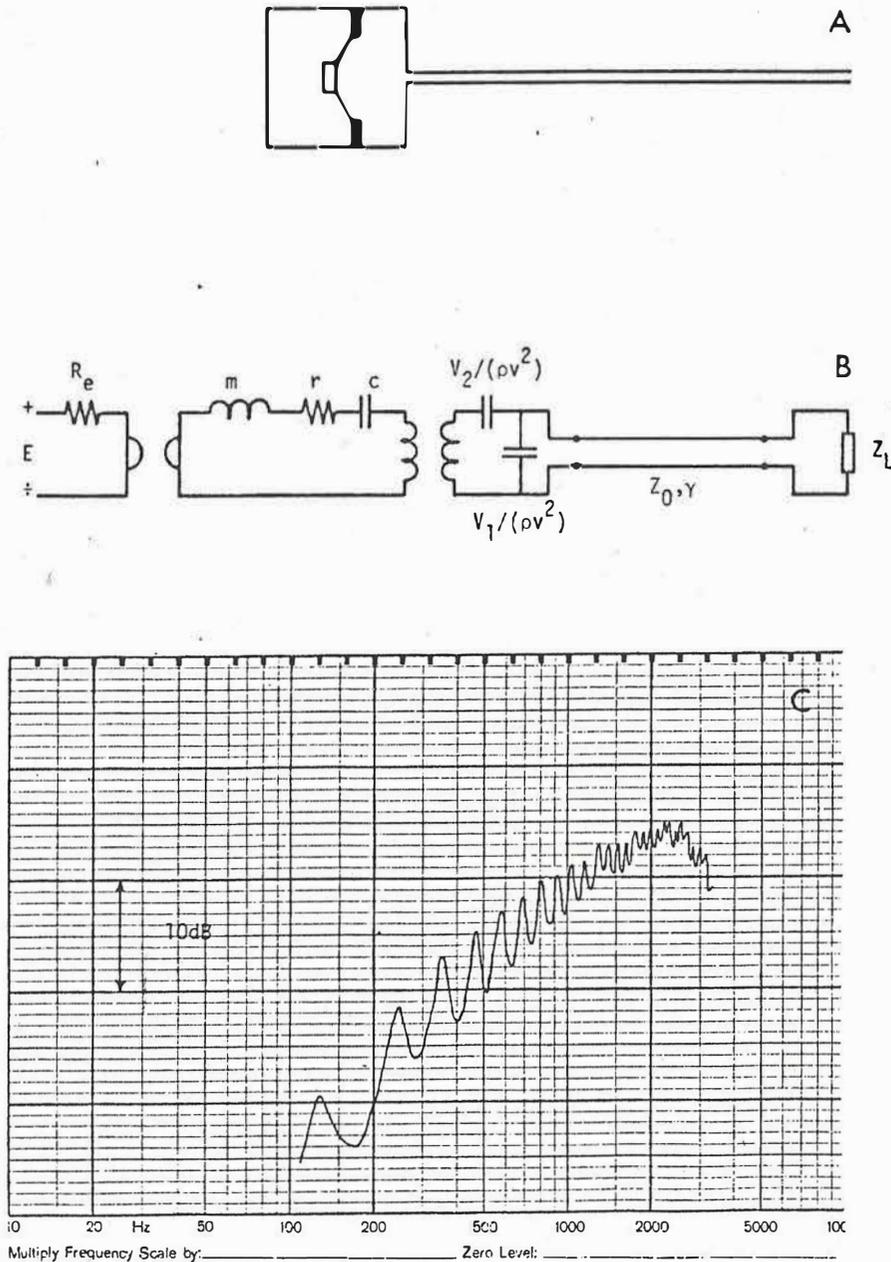


Figure 10-3. (A) Acoustical system consisting of enclosed loudspeaker and tube; (B) electrical equivalent circuit of (A); and (C) sound pressure measured 10 cm in front of tube ending for constant voltage input.

a normal speaker pronounces a vowel with his natural voice and afterwards makes an effort to do the same with the tube in the mouth and a known electrical-input spectrum. The difference spectrum was averaged over all vowels and over four speakers; the power spectrum of the input signal could then be determined.

Utilizing the fact that the ear is insensitive to phase, the phases of the various harmonics may be shuffled around to give a minimum peak-to-peak signal, which is significant for an efficient use of the battery voltage. The resulting signal shape in the time domain was found by using an optimization technique; an example for the male voice is shown in Figure 10-4. In this system, 256 different points of the period are stored with an eight-bit accuracy in the memory.

PERFORMANCE

In its present form the speech aid is not fully satisfactory. It has been used successfully by some volunteers, but it seems to require a little training. Especially for patients in bed it has been difficult in the sense that the intelligibility was low, perhaps because the original spectrum was developed with the aid of healthy subjects who simulated a closed glottis.

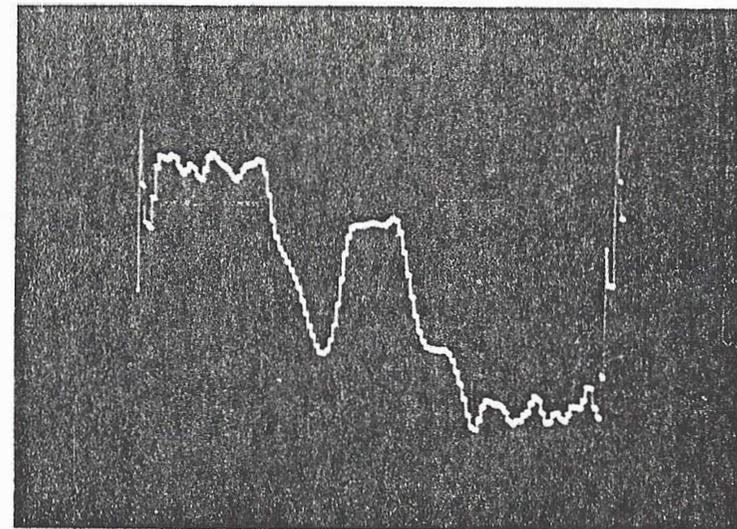


Figure 10-4. One period of electrical input to loudspeaker.

In Figure 10-5, spectrograms are shown of a Danish sentence spoken by one of the authors with (a) a natural voice, (b) a speech aid and closed glottis, and (c) a speech aid, open glottis, and open to the nasal cavity. It is clear that the situation in (c) is considerably worse than in (b).

Comparing Figure 10-5a and 10-5b, it is noted that smooth formant transitions in natural speech are only partly reconstructed. One explanation for this may be that the position of the sound inside the cavity instead of at one end introduces zeros in the transfer function at certain frequencies. These zeros may not be fully compensated with the chosen signal due to the narrow bandwidth of the zeros. One possible remedy for this may be a multielement source.

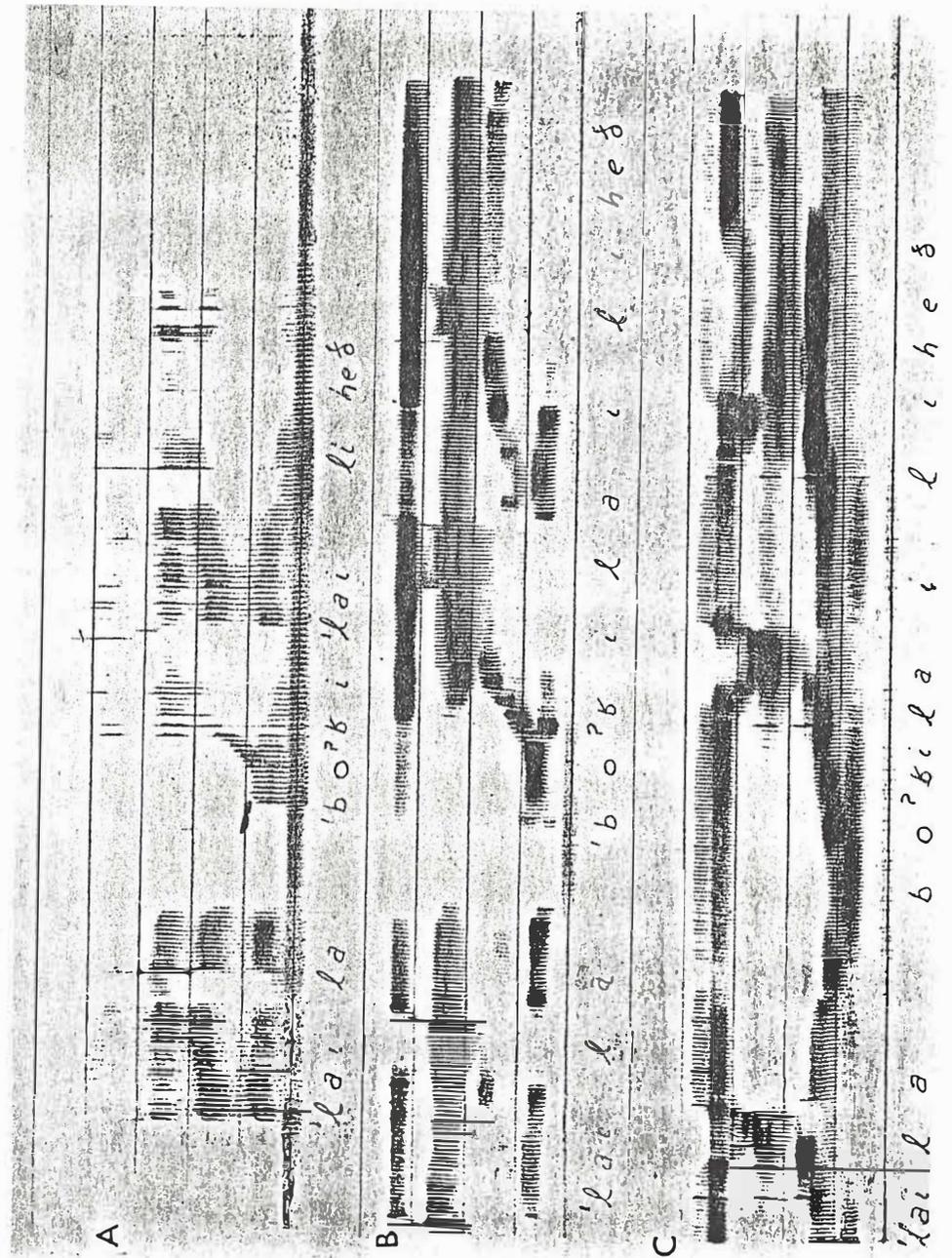


Figure 10-5. The Danish sentence "Laila bor i lejlighed" spoken by one of the authors: (A) with natural voice; (B) with speech aid and closed glottis; and (C) with speech aid and glottis open to nasal cavity.