THE BOOK WAS DRENCHED

170383

UNIVERSAL
LIBRARY

OU_170383

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 536.4 / P55V Accession No. 2615

Author Potter R. K and others

Title Visible Speech

This book should be returned on or before the date last marked below.

VISIBLE SPEECH

THE BELL TELEPHONE LABORATORIES SERIES

Theory of Vibrating Systems and Sound, by Irving B. Crandall

Contemporary Physics, by Karl K. Darrow
(Second Edition)

Speech and Hearing, by Harvey Fletcher

Probability and Its Engineering Uses, by Thornton C. Fry

Elementary Differential Equations, by Thornton C. Fry
(Second Edition)

Transmission Circuits for Telephonic Communication
by K. S. Johnson

Transmission Networks and Wave Filters, by T. E. Shea

Economic Control of Quality of Manufactured Product
by W. A. Shewhart

The Application of Electromechanical Impedance Elements in Transducers and Wave Filters, by Warren P. Mason

Rhombic Antenna Design, by A. E. Harper

Poisson's Exponential Binomial Limit, by E. C. Molina

Electromagnetic Waves, by S. A. Schelkunoff

Fundamental Theory of Servomechanisms, by LeRoy A. MacColl

Network Analysis and Feedback Amplifier Design
by Hendrik W. Bode

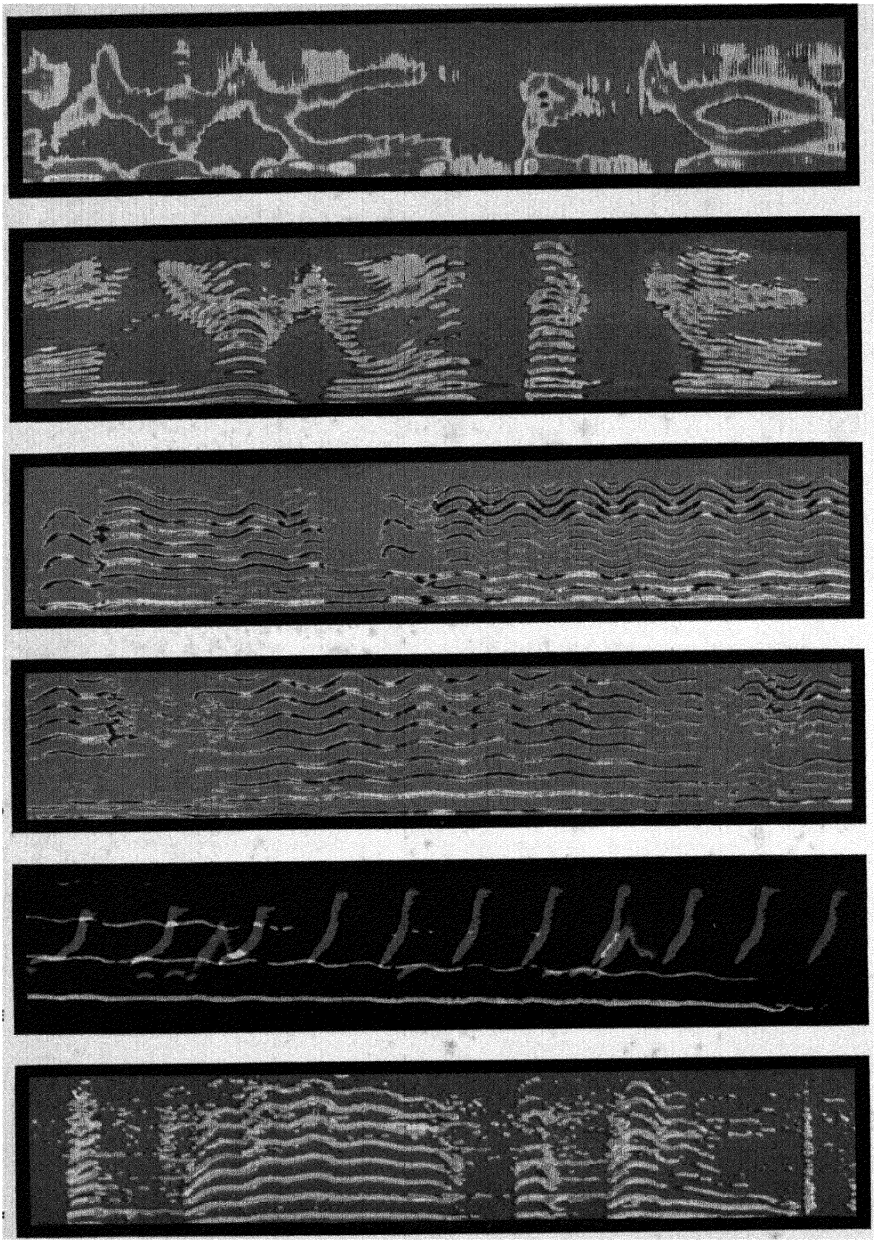
Quartz Crystals for Electrical Circuits, by Raymond A. Heising

Capacitors – Their Use in Electronic Circuits
by Manfred Brotherton

Fourier Integrals for Practical Applications
by George A. Campbell and Ronald M. Foster

Visible Speech
by Ralph K. Potter, George A. Kopp, and Harriet C. Green

PUBLISHED BY D. VAN NOSTRAND COMPANY, INC.



Sound spectrograms in which color indicates the relative energy in different parts of the frequency scale. Each spectrogram represents about two seconds of time, and a frequency range of 70 to 3,500 cycles per second. All except the first are high-resolution patterns. A—The words "unusual picture," male voice. B—The same words, male voice, in high resolution, showing more overtone detail. C—Two sobs by Caruso at the end of "Vesti la giubba" from *I Pagliacci*. (RCA Victor record 6001-A). D—The words "La victoire" by Louise Homer in "Je viens célébrer la victoire" from *Samson and Delilah*. (RCA Victor record 5003-B). E—Three superposed bird songs: cardinal, red; whippoorwill, blue; screech owl, green. (Cornell University Laboratory of Ornithology recording). F—The words "The play's the thing" from *Hamlet*, spoken by John Barrymore. (RCA Victor record 6827-A).

VISIBLE SPEECH

RALPH K. POTTER
DIRECTOR OF TRANSMISSION RESEARCH
BELL TELEPHONE LABORATORIES

GEORGE A. KOPP
FORMER MEMBER OF TECHNICAL STAFF
BELL TELEPHONE LABORATORIES

HARRIET C. GREEN
FORMER MEMBER OF TECHNICAL STAFF
BELL TELEPHONE LABORATORIES



NEW YORK
D. VAN NOSTRAND COMPANY, INC.
1947

RALPH K. POTTER, Director of Transmission Research at the Bell Telephone Laboratories, initiated the "Visible Speech" project, and has closely supervised every phase of its development.

GEORGE A. KOPP, formerly a member of the technical staff of the Bell Telephone Laboratories, is now Associate Professor of Speech in the College of Literature, Arts, and Sciences at the University of Michigan.

HARRIET C. GREEN, formerly a member of the technical staff of the Bell Telephone Laboratories, is now Assistant Professor of Special Education at the Michigan State Normal College.

Copyright 1947 by D. Van Nostrand Company, Inc.

All rights reserved. This book, or any parts thereof, may not be reproduced in any form without written permission from the authors and the publishers.

Published simultaneously in Canada by McClelland & Stewart, Ltd.

Printed in the United States of America

First Edition

VISIBLE SPEECH

Preface

This book deals with a new form of visible speech obtained by the recording of an analysis of speech which is somewhat similar to the analysis performed by the ear. The visible patterns so produced have proved to be readable and hence may offer a means for visual hearing. In a sense, these patterns are an automatically written phonetic alphabet.

Originally much of the material presented here was prepared for a textbook to be used in an experimental training program in the reading of visible speech. Accordingly, it is written in non-technical language. During the preparation of the book, it seemed desirable to enlarge its scope to include material of special interest to the various groups that will ultimately evaluate this development. Such groups include the deaf, those concerned with teaching the deaf, those concerned with teaching languages, and those interested in phonetics, in the nature and improvement of speech, and in vocal music. To broaden the scope, information has been included on the progress to date, the present problems, and the possible applications to specialized fields concerned with speech and vocal music.

It is hoped that the experimental state of the development will be recognized, and that this initial presentation will elicit a critical evaluation of present progress as well as suggestions that will help guide future work.

In the following pages, we have attempted to emphasize the experimental nature of the development, and to avoid dogmatic or arbitrary statements concerning present evidence and information. The tentative nature of our conclusions, principles, and concepts cannot be overstressed. An enormous number of unanswered and even unformulated questions exists in the field of casual relationships, linking the visible patterns, the physical factors, and the physiological functions. While each of the types of modulation and the combined effects of articulatory and phonatory changes may be described in terms of total modifications in the visible patterns, and may be associated in a general way with the underlying physiology, kinesiology, and physics, it is not possible at this time to assess the specific contributions of the various parts of the speech mechanism to the configuration as a whole. For instance, what is the effect of varying tension in the vocal tract upon the position, shape, and appearance of the resonance areas of the vowel; or, what is the specific contribution of a particular part of the vocal mechanism, as it functions in combination with the total mechanism, to the aural and visible sound patterns? Difficulties in exact measurement, complexities inherent in both physiologic and acoustic interrelationships, and limitations in instrumentation, make definitive assumptions or conclusions concerning visible speech and the physiology, physics, and psychology of speech dependent upon future

research. Such problems, in addition to those connected with the evaluation of changes or modifications in translation techniques and the resultant variations in pattern form, offer ample basis for long term programs of research to be carried on in many different fields of interest.

Many members of the Bell Telephone Laboratories have participated enthusiastically in the development described in the following pages. Since the development period included the war, there were frequent personnel changes. In fact, very few individuals connected with the project at the start have continued with it to the present time. Consequently, as much as we should like to do so, it is impractical to acknowledge all contributions in detail. As a compromise, there are listed below the names of those in the project group during some part of the period when the experimental training studies were being brought to a conclusion and while material was being collected and organized for this volume.

Experimental Training Class — Mr. Edgar Bloom, Jr., Miss Rosemary T. Kennedy, Miss Selma Graber, Mrs. Evelyn B. Leyland, Mrs. Leanore L. Levine, Miss Anna J. O'Connor, Miss Frances Aaron, Miss Mary F. Barberi, Mrs. Florence W. Davis, Mrs. Doris F. Gibson, Mrs. Rose M. Rogers and Miss Florence P. Raedler.

Sound Spectrograph Development — Mr. Walter Koenig, Jr., Mr. A. E. Ruppel, Mr. H. J. Christopher, Mr. C. W. Vadersen, Mr. R. L. Wallace, Jr.

Translator Development — Mr. H. W. Dudley, Mr. R. R. Riesz, Mr. L. Schott, Mr. O. O. Gruenz, Jr.

Pattern Evaluation Studies — Mr. N. R. French, Mr. S. S. A. Watkins, Dr. G. E. Peterson.

Publication Problems — By arrangement with Mr. R. K. Honaman, Director of Publication for the Bell Telephone Laboratories, Mr. K. M. Collins designed the book and handled the many problems associated with publication and printing.

Supervision — Throughout the early stages, Mr. C. H. G. Gray was responsible for immediate supervision of the development. During the summer of 1945 this responsibility was transferred to Dr. J. C. Steinberg who in this capacity has cooperated with the authors in completing the plans for this book.

In conclusion the authors wish particularly to acknowledge the inspiring encouragement that has been given this project from the beginning by Dr. O. E. Buckley, President of the Bell Telephone Laboratories.

RALPH K. POTTER
GEORGE A. KOPP
HARRIET C. GREEN

Contents

Foreword by Dr. Oliver E. Buckley . . . . . *page* xv

Part One
Development and Basic Principles of Visible Speech

CHAPTER 1 - *Introduction* . . . . . *page* 3

CHAPTER 2 - *The Sound Spectrograph* . . . . . *page* 8

CHAPTER 3 *The Direct Translator* . . . . . *page* 16

CHAPTER 4 - *The Training Program* . . . . . *page* 22

CHAPTER 5 - *The Basic Principles of Visible Speech — The New ABC's* *page* 28

Basic Speech Processes *page* 30
The Six Sound-Pattern Groups *page* 33
Sounds Influence Each Other *page* 38
The Hub - The Key to How Sounds Change Each Other *page* 39
Essential Characteristics of the Hub Principle *page* 43
How Stressed and Unstressed Sounds Look *page* 51
The New ABC's . . *page* 53
The ABC's of Visible Speech . . *page* 54

Part Two
Lesson Units — Visible Patterns of American Speech

CHAPTER 6 — *Suggestions for Using the Lesson Chapters* . . . *page* 59

CHAPTER 7 — *The Patterns and the Hubs of the Vowels and Diphthongs* . *page* 65

Each Front Vowel Has a Distinct Pattern . *page* 66
Each Back Vowel Has a Distinct Pattern . . *page* 69
Each Mid-Vowel Has a Distinct Pattern . . . *page* 72
Each Diphthong Has a Distinct Pattern . . *page* 74

CHAPTER 8 — *Reading the Stop Sounds* . . . . . *page* 79
Unit 1 — Reading the **p** (pay) and **b** (be) . *page* 81
Unit 2 — Reading the **t** (to) and **d** (day) . . *page* 87
Unit 3 — Reading the **k** (key) and **g** (go) . *page* 94
Unit 4 — Comparative Review of the Stop Sounds . . . . *page* 101
Unit 5 — Reading Combinations of the Stop Sounds . . . . *page* 104

CHAPTER 9 — *Reading the Fricative Sounds* . . . . *page* 110
Unit 1 Reading the **h** (he) and **ɦ** (ahead) . *page* 111
Unit 2 — Reading the **f** (for) and **v** (vote) . *page* 118
Unit 3 — Reading the **θ** (thin) and **ð** (then) . *page* 126
Unit 4 — Reading the **s** (see) and **z** (zoo) . *page* 134
Unit 5 — Reading the **ʃ** (she) and **ʒ** (azure) . *page* 142
Unit 6 — Comparative Review of the Fricative Sounds . . *page* 149
Unit 7 — Reading Combinations of the Fricatives and Stops . . . *page* 153

CHAPTER 10 — *Reading the Vowel-Like Sounds — The Nasal Sounds* . *page* 166
Unit 1 — Reading the **m** (me) . . . *page* 167
Unit 2 — Reading the **n** (no) . . . *page* 174
Unit 3 — Reading the **ŋ** (sing) . . *page* 183
Unit 4 — Comparative Review of the Nasal Sounds . . *page* 188
Unit 5 — Reading Combinations of the Nasals, Stops, and Fricatives *page* 190

CHAPTER 11 — *Reading the Vowel-Like Sounds — The Glides* . . *page* 202
Unit 1 — Reading the **w** (we) . *page* 203
Unit 2 — Reading the **j** (you) *page* 211
Unit 3 — Reading the **r** (read) *page* 218
Unit 4 — Reading the **l** (let) . . *page* 231
Unit 5 — Comparative Review of the Glide Sounds . *page* 244
Unit 6 — Reading Combinations of Glides and Other Sounds *page* 246
Unit 7 — Reading Combinations of Glides, Stops, and Fricatives *page* 251

CHAPTER 12 — *Reading the Vowels* . . . . . *page* 270
Unit 1 — Reading the Front Vowels . . *page* 272
Unit 2 — Reading the Back Vowels . . *page* 275
Unit 3 — Reading the Central Vowels . *page* 278
Unit 4 — Reading the Diphthongs . . *page* 280
Unit 5 — Summary Diagrams of the Vowels *page* 281
Unit 6 — Profiles of the Vowels . . . *page* 282

Part Three
Applications of Visible Speech

CHAPTER 13 — *Deaf Interests* . . . . . *page* 287
The Problem . . . . *page* 289
Patterns for Speech Education *page* 294
Specific Plans . . . *page* 297
Conclusions . . *page* 301

CHAPTER 14 — *Phonetic Interests* . . . . . *page* 302
Pattern Possibilities . . *page* 302
Phonetic Research Possibilities . *page* 307
Miscellaneous Illustrations . . *page* 315
Need for Mutual Understanding . *page* 324

CHAPTER 15 — *Speech Correction Interests* . . . . *page* 325
Spectrograms Illustrating Certain Speech Defects . . . *page* 329

CHAPTER 16 — *Foreign Language Interests* . . *page* 343

CHAPTER 17 — *Vocal Music Interests* . . . . . *page* 376

CHAPTER 18 — *Generalities and Speculations* . . . . *page* 410

Appendix I — Reference Listing of Phonetic Symbols . . . *page* 425

Appendix II — Summary of the Profiles and Hub Areas . . . *page* 427

Bibliography of Technical Publications on Visible Speech . . . *page* 430

Index . . . . . *page* 431

Foreword

Improvement of the telephone and extension of its service are principal aims of research in Bell Telephone Laboratories. While requirements of the average telephone user with normal hearing are employed as a guide in this work, the needs of others are not overlooked. Several years ago the laryngaphone made the telephone available to many who had lost the ability to speak and restored to them the social contacts that only vocal communication can provide. Likewise, improvements in hearing aids brought welcome relief to those whose hearing was impaired, and the incorporation of an amplifier in the telephone has made the latter instrument fully usable to those who have only the bare minimum of oral sensitivity required to distinguish speech sounds. But there remains, in the totally deaf, a large group whom the telephone cannot at present serve directly.

It was with the hope of making the telephone available to the totally deaf that the development of a mechanism for portraying speech sounds visibly was undertaken. The problem we set out to solve still awaits complete solution, but as so often happens in research, more significant results have been achieved than were at first imagined. There has evolved a simple method of depicting speech sounds that makes feasible their quick recognition and interpretation by eye. For the first time it has become possible to represent sound graphically so that pictures defining it fully can be printed in books. But much more importantly, a mechanism has been produced which enables deaf students to see their own voices which they cannot hear, and to make corrections under a teacher's guidance.

In a device so new, applied to problems so old, it is easy to overrate practical possibilities and to underrate the difficulty of their achievement. Only time and continued research will bring forth a true evaluation of the results already achieved. But the ever-growing enthusiasm of those who have worked with these pictures of sound, and the benefit already achieved by one of our own number, gives hope that, in this line of research, a significant contribution has been made toward helping those to whom the inventor of the telephone devoted so much of his own life's effort.

In his ingenuity, enthusiasm and perseverance in leading this research, Ralph Potter, the senior author of this book, has exemplified the heritage of a great tradition which stems from Alexander Graham Bell and today inspires those who carry on the work he started. Joining with Mr. Potter as co-

authors are two who, with the benefit of knowledge and experience in teaching the deaf to speak, have explored the application of this new art of depicting speech sounds and contributed much to its organization.

Nothing is so important for the progress of a new method or art as a comprehensive compilation of existing knowledge that gives a basis of fact and terminology on which others can build and in the light of which they can fruitfully explore. That purpose this book is designed to serve.

OLIVER E. BUCKLEY
President, Bell Telephone Laboratories

VISIBLE SPEECH

PART ONE

DEVELOPMENT AND BASIC PRINCIPLES OF VISIBLE SPEECH

Chapter 1

Introduction

In 1867, Melville Bell, the father of Alexander Graham Bell the inventor of the telephone, published a book called "VISIBLE SPEECH: THE SCIENCE OF UNIVERSAL ALPHABETICS." Although the present volume also is entitled "VISIBLE SPEECH," the meaning attached to the term is substantially different. There is, however, an interesting relation between the two that deserves mention.

Early handwritten symbols

Melville Bell called his symbols "visible speech" because they specified the pronunciation of word sounds so accurately, that speech thus symbolized could be repeated by anyone else familiar with the method. This was true even though the words were in an unfamiliar tongue or modified by an unusual accent. Alexander Graham Bell, then a young man, assisted his father in the demonstration of these symbols. With Alexander off-stage and out of hearing range, Melville Bell would ask members of the audience, preferably from distant parts of the country or world, to come up and say a few words of their own choosing for him to write down. After making such records, he would call in his son who would read aloud the written sounds. After each reading the original speaker would pronounce the words so that the audience could hear that the pronunciation had been accurately imitated. The words "*visible speech*," written in Melville Bell's symbols, are shown in Fig. 1.



Fig. 1 — The words "Visible Speech" in the handwritten symbols developed by Melville Bell in 1867.

Upon their disclosure these handwritten symbols were hailed as a major contribution to phonetics and the education of the deaf, but the most important effect upon the world at large may not have been evident at that time. Alexander Bell, like his father, was deeply interested in the problems of deafness, and planned to devote his life work to this field. To him, the visible speech symbols suggested a further accomplishment, that of actually making speech visible in a form that could be read. Although his subsequent search for a specific device to accomplish the translation of speech into a readable form was unsuccessful, it led to investigations of electrical devices and ultimately to his discovery of the telephone.

Many others have searched for a visible speech translator, and numerous ways for making records of speech sounds have been devised in the past, but these records have

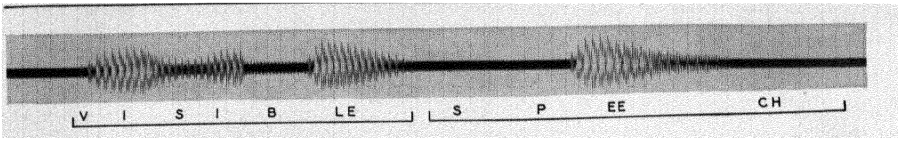


Fig. 2 — The words "Visible Speech" in the form recorded by the oscillograph.

not proved to be readable. For example, both oscillograph traces such as are illustrated by Fig. 2, and sound tracks on moving picture films, are unreadable to the eye. At the same time all the information is present in such records and it is possible to obtain the original sound by means of a suitable reproducing device.

New voice-written symbols

To enable the eye to obtain the information in a speech sound, a translator must provide, at least in an approximate fashion, a translation similar to that made by the ear. It should spread out the dimensions of speech so that they are visible to the eye as they are audible to the ear. Such a translation of sounds into readable patterns has been realized through the development of special translating devices.* One instrument for this purpose has been named the "sound spectrograph." This is essentially a device for making visible records of the frequency, intensity, and time analysis of short samples of speech. The analysis is so made that the variations of the vocal resonances with

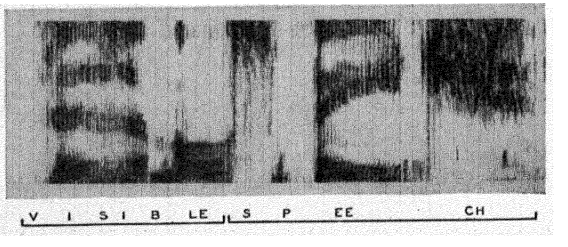


Fig. 3 — The words "Visible Speech" as recorded by the sound spectrograph.

time are portrayed conspicuously. It is these variations which appear to be of considerable importance to readability. A sound spectrographic record for the words "*Visible Speech*" is shown in Fig. 3. The pattern is a new form of visible speech, a system of natural phonetic symbols translated from speech itself, a type of record that would seem to fulfill Alexander Graham Bell's early ambition.

A spectograph development project was started early in 1941. American participation in the war emphasized the military applications of, and needs for, a visual translation

* Originally described in the magazine "Science" in the issue of Nov. 9, 1945.

of sound, and during the war period major interest was centered upon these military requirements. It was not until military needs had been met that opportunity was offered for a resumption of the development of visible speech toward the goal of visual hearing. Because of the military interest, publication of information concerning the nature or progress of the project was delayed until the end of the war. In order to present adequately the information accumulated during the course of the developmental work, it seemed that a book of the present scope was required.

Evidence of readability

Evidence that the sound spectrograph patterns are readable appears in a comparison of patterns for different words and different speakers. Different words have a different appearance, an essential requirement if they are to be told apart. But the same words spoken by different individuals have a similar appearance, also an essential requirement if the symbols are to be of practical use. That the patterns of different words have a different appearance is evident to some extent in Fig. 3, and it will be more apparent from an examination of the illustrations which follow. That the patterns of the same words by different speakers have a similar appearance is shown in Fig. 4. This figure presents six sound spectrograms of the word "*speech*" as pronounced by six speakers with a wide range of voice quality. There is an evident similarity in the appearance of the patterns which is probably as distinct as it is in the sound of these same words.

It was not until the autumn of 1943 that there was opportunity to explore the visual hearing possibilities of sound spectrograph patterns in a systematic way. At that time, an experimental training group comprising several girls with normal hearing was organized, and a training program was started under the supervision of the authors. The progress of the group was slow, as might be expected since the subject was then new to the teachers, and the teaching methods were correspondingly uncertain. In addition, the only instrument then available for obtaining speech patterns was the sound spectrograph which required minutes to record a short sample of speech on paper. Nevertheless, before the end of the first training year it was demonstrated that both stationary and slowly moving patterns could be read by the entire group. The training experience is described in more detail in Chapter 4.

Late in 1944, the first direct translator, described in Chapter 3, was made available for class use. This direct translator produced patterns like those shown in the spectrogram of Fig. 3. The patterns appeared upon a screen similar to the type used in television receivers, except that this screen was moving continuously. As the words were spoken, patterns were formed at the right edge of the screen and moved slowly towards the left. The several words visible on the screen could be read from left to right as one reads the words on this page.

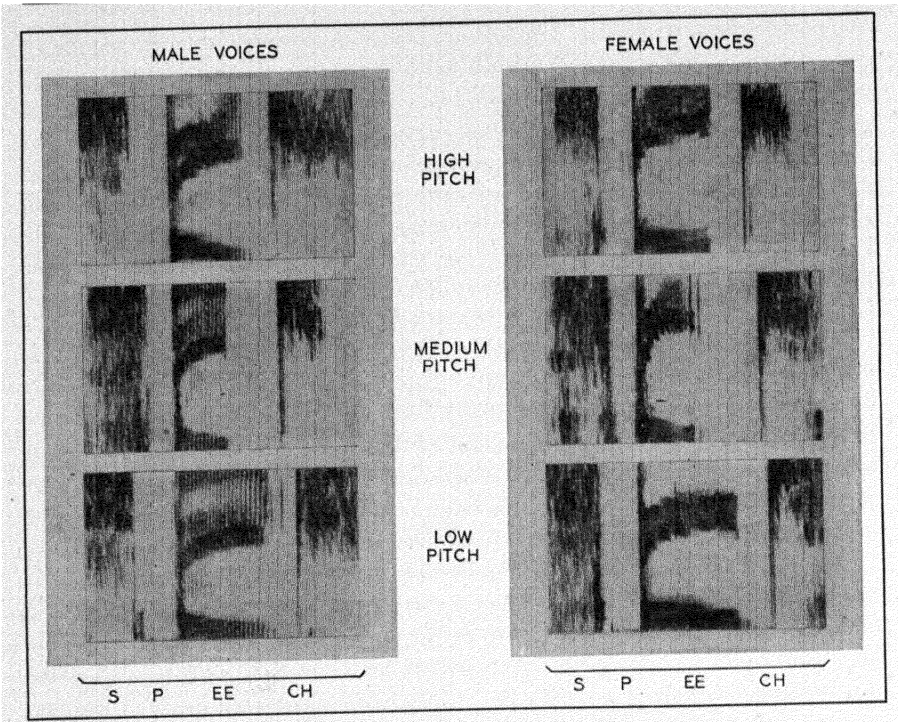


Fig. 4 — Patterns of the word "speech" by six different speakers with a wide range of voice quality.

Although the results of this first experimental training have been encouraging, they do not provide conclusive answers to questions concerning the possible use of visual hearing by the severely deafened. There is still some uncertainty both as to whether the new ABC's may be learned easily enough to be practicable, and as to when the present experimental translators can be developed into a form suitable for general use. Time will be required to determine the answers to both of these questions.

Need for development

Even though the ultimate usefulness of visible speech for the deaf has not been determined and practical aids are not yet available, there are a number of reasons that make publication of present information desirable. It is clear that these natural phonetic symbols will assist those interested in the analysis and written discussion of

sound. In order to assess the practicability of the instruments, patterns, and training procedures already developed for visual hearing, it is essential that they be put to the test of actual use with the severely deafened and the congenitally deaf. While it is possible to anticipate many requirements of these groups, there are numerous others that will arise in the course of practical and direct use. The contribution of the deafened, and of those concerned with the problems of the deaf, is basic to the future development of instruments, of training methods, and of applications of visible speech. Therefore, it was decided to publish the available information concerning visible speech in order to elicit from all interested groups a critical evaluation of present progress, and suggestions for future development.

To describe visible speech adequately requires something akin to a textbook; accordingly, a large part of this volume is just that — a textbook that should enable anyone to learn as much about the subject as is possible through the medium of words and pictures. The text material is preceded by a description of the sound spectrograph; a discussion of the development of direct translators; an account of the progress and experience with the experimental training program; definitions of the special terms that have been introduced to facilitate the study of the patterns; and an explanation of the phonetic principles basic to the reading of visible speech. The text itself consists of lesson units selected to develop the knowledge and skills essential both to an understanding of visible speech and to the acquisition of a basic reading vocabulary. The lesson units are arranged in progressive order of difficulty and are written with both the adult hearing and the severely deafened reader in mind. In the final chapters, practical applications of visible speech to various specialized fields are discussed, and a number of examples are included.

For those who are interested in more detail concerning certain parts of the development, a bibliography of the technical papers so far published is included at the end of this book.

Chapter 2

The Sound Spectrograph

While involved technical details are not required to accomplish the purposes of this book, some technical explanation of principles will assist those who wish to understand the association between audible effects of sound and the visible effects as portrayed by the translation process.

The simplest of sounds is the so-called pure tone. It is produced by a simple sine-wave form of oscillation such as that shown by the oscillographic trace in Fig. 1. An example of such an oscillation is the sound wave produced by a whistle when blown with constant frequency and intensity. In the trace, intensity is proportional to the square of the vertical swing, and frequency is proportional to the number of oscillations in a given interval along the horizontal axis or time dimension. The sensations of loudness and pitch of the tone are related, respectively, to the intensity and frequency of the wave.

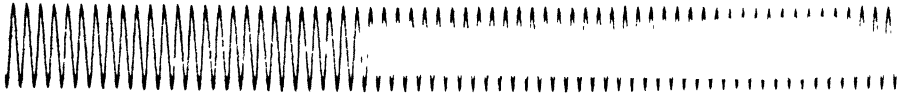


Fig. 1 — Oscillographic representation of a simple oscillation such as a steady tone.

A simple tone such as is represented in Fig. 1 is uninteresting in itself; in fact, it may be distinctly annoying as, for example, when a constant squeal is produced on the radio by interference. But, these simple tones are the threads that combine to make up the beautiful patterns of sound that delight us in music, or the hodgepodge patterns that irritate us in noise. They are the substance of sound.

When a number of simple sine waves are combined, for example, by blowing many whistles simultaneously, a complex wave of the type shown in Fig. 2 may be produced.



Fig. 2 — Oscillographic picture of a complex wave made up of many simple oscillations.

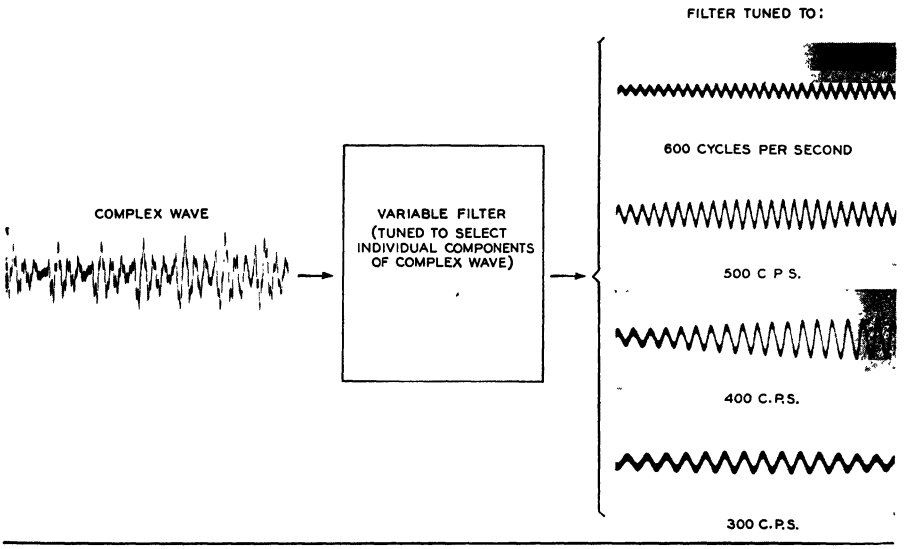


Fig. 3 — Illustrating the selection of different individual components in a complex wave by changing the tuning of an analyzing filter.

The simple waves that combine to produce a complex wave tend to lose their identity as far as the eye is concerned. An expert on wave analysis might make a rough estimate of the frequency components present in a complex wave, but he could do this only for the less complicated cases. In order to unravel a very complex wave and to recover the simple oscillations, it is necessary to use an analyzer of some kind. One form of analyzer is a filter, made up of tuned electrical circuits, that will pass oscillations in a particular frequency range. The common radio receiver is an example. It is an adjustable filter that is tuned to pass the frequencies of one radio station at a time, and to suppress stations having higher and lower frequencies. Similarly, apparatus containing tunable circuits may be designed to pass oscillations in one frequency range of the voice, and to suppress higher and lower frequency ranges.

Separation of wave components

A diagrammatic illustration of the action of such a filter is given in Fig. 3. At the left is shown an oscillogram of a complex wave. To select the various individual components that make up this wave, the variable filter is tuned successively across the audible frequency range in much the same way that a radio receiver might be tuned

across the radio frequency range. At the right are shown the simple waves selected by the filter for the settings indicated.

Speech waves, of course, have many more than four components. In addition, the components slowly change with time as the sounds are formed in succession. Hence, the wave shapes are not steady and unchanging as in Fig. 2, and the results of an analysis such as is shown in Fig. 3, will change slowly with time. When a wave form repeats itself over and over, the simple waves that comprise it are harmonically related. In other words, their frequencies are integral multiples of a lowest or fundamental frequency. The voiced sounds of speech approximate harmonic waves during the steady state portions of the slow changes previously noted. In other types of complex waves, such as the hiss of escaping steam, the oscillations may occur in an entirely random manner. These waves also are made up of simple oscillations, but their frequencies are inharmonically related. Whispered and unvoiced sounds of speech are produced by waves of this type.

With a single variable filter, a complex wave can be analyzed in only one frequency band at a time just as a radio receives only one station at a time. If it were essential to receive simultaneously all of the frequency bands occupied by a complex wave, it would be necessary to use an extensive array of filters each adjusted to a slightly different frequency position. Within limits, such an elaborate machine could analyze any complex wave as fast as it was received. It is believed that the ear functions in a somewhat similar manner. The indications are that within the cochlea of the inner ear there is an array of filtering elements that man could not duplicate, at present, with

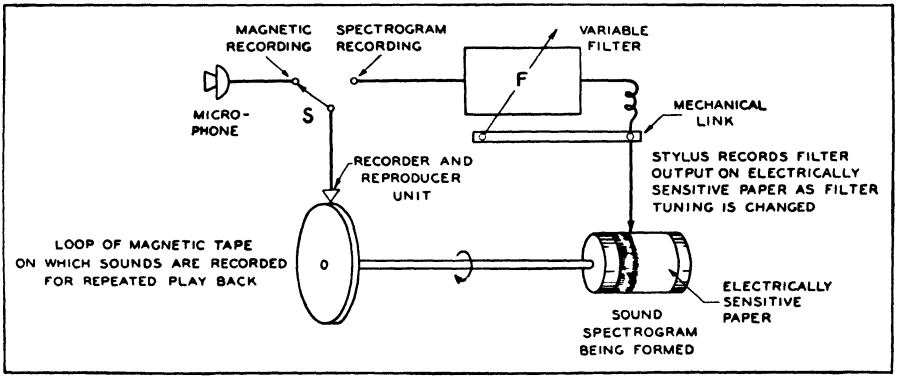


Fig. 4 — Schematic diagram of the sound spectrograph illustrating how the analysis of a repeated sample of speech is recorded line by line to form a complete pattern.

less than a small roomful of apparatus. Successive elements are tuned to progressively different frequency bands. Apparently, when simple oscillations occurring in a complex wave coincide with the oscillations to which an element is tuned, the element initiates nerve impulses, and the results of this analysis are carried over neural circuits to the brain.

Analysis by repetition

To duplicate the inner ear accurately would be a sizable undertaking with present day circuit techniques, but if it were permissible to repeat the complex wave over and over again, only one adjustable filter would be required. The output for each adjustment could be recorded progressively so that, after a sufficient number of filter readjustments and repetitions of the complex wave at the filter input, all the simple components would be obtained. This is the principle of the so-called "sample type" of sound spectrograph. In effect, it is an instrument that analyzes, one band at a time, the simple oscillations of a complex wave, and records the intensity variations in each band side by side in an orderly fashion upon a sheet of paper. The result is a visible pattern of the sound in its three fundamental dimensions — frequency, intensity, and time.

A schematic diagram of the arrangement used to accomplish this type of analysis and recording is shown in Fig. 4. With the switch "S" thrown left, one may talk into the microphone and record on a loop of magnetic tape a short sample of speech such as the words, "*Speech we may see.*" Then the switch "S" is thrown right, and the magnetic tape record is reproduced over and over again. The repeated speech sample goes to the input of the variable filter "F." Adjusted first to some starting frequency such as 50 cycles per second, its tuning would be shifted say 15 cycles at a time for each repetition of the words. The filter output is connected to a stylus resting upon electrically sensitive paper wrapped around a drum as shown at the right. The simple oscillations separated out from the complex wave are recorded side by side as both the filter tuning and stylus position shift together over the frequency range. The darkness of the trace increases as the intensity of the oscillations increases. Hence, the record is one of varying shade rather than an oscillating line such as is recorded in the familiar oscillogram.

When the filter "F" has a band width of approximately 45 cycles per second, a pattern such as is illustrated by Fig. 5, on page 12, is obtained. The parts of the pattern are rather fully identified. These patterns contain details of speech sounds that are of interest in studies of speech defects, in speech rehabilitation, and in phonetic research.

When the filter "F" has a band width of about 300 cycles, as shown in Fig. 6, much of the finer detail is not portrayed. However, there remains in broad outline, a picture of the intensity-frequency-time distribution, where frequency and time are shown

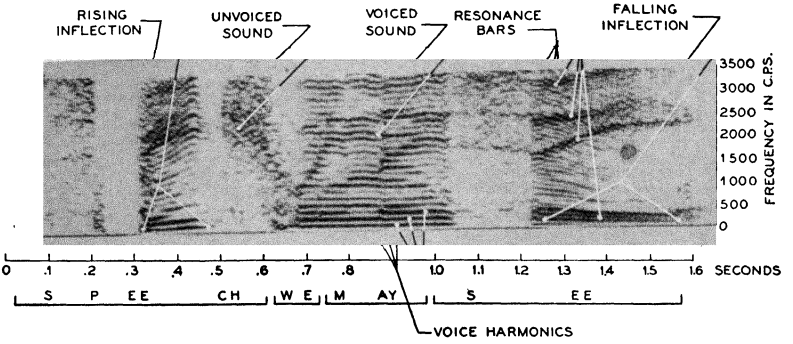


Fig. 5 — Sound spectrogram of the words "Speech we may see" using a narrow-band analyzing filter (45 cycles) to portray harmonic structure.

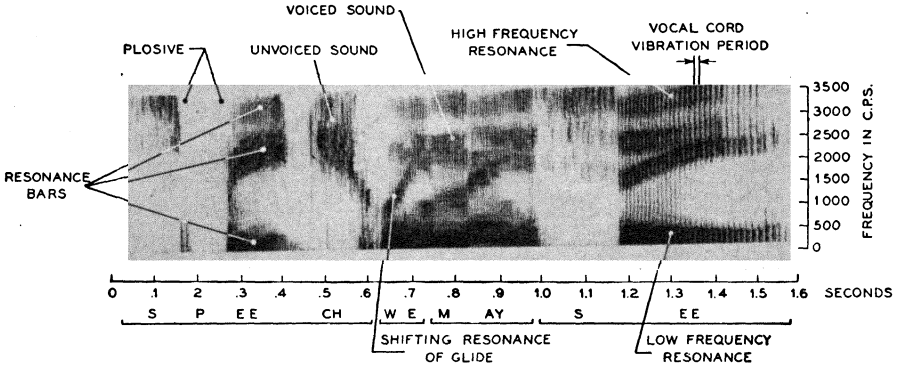


Fig. 6 — Sound spectrogram of the words "Speech we may see" using a wide-band analyzing filter (300 cycles) to emphasize vocal resonances.

vertically and horizontally, respectively, and intensity is shown by the shade of darkness. Patterns of this type are the ones of special interest in visual hearing.

When the words, "*Speech we may see,*" were whispered, the records shown in Fig. 7 and in Fig. 8 were obtained. Although the appearance of these two spectrograms is different, Fig. 7 for the narrow-band filter contains less information important to the recognition of the word sounds than does Fig. 8 for the much wider-band filter. The latter shows more clearly how the speech energy is distributed. It is interesting to note that, neglecting the detailed form of vertical striations, Fig. 8 shows a general energy distribution somewhat similar to that which appears in the patterns of Fig. 6.

Intelligibility tests made with whispered and voiced speech reveal that when the former is loud enough to be heard easily, it may be understood about as well as the latter, although voicing makes speech more pleasant to listen to and helps to identify the speaker.

A natural inference appears to be that the movements of the articulators, the tongue, lips, jaw, etc., play a prominent role in producing the intelligible characteristics of speech sounds. An important conclusion is that the effects of these movements may be portrayed adequately for purposes of reading, by the wide-band filters. This bears upon the design of translators, as will be discussed later.

Experimental equipment

In Fig. 9 is shown a model of the sound spectrograph that has been used for obtaining most of the patterns shown in this book. The recorder unit contains the loop of magnetic tape, the driving motor, the recording drum, and the stylus. The two latter

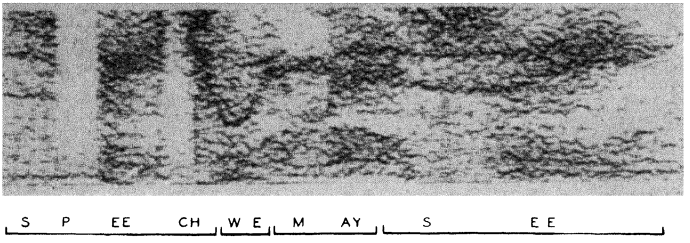


Fig. 7 — Sound spectrogram of the words "Speech we may see" spoken in a whisper, using a narrow-band filter (45 cycles).

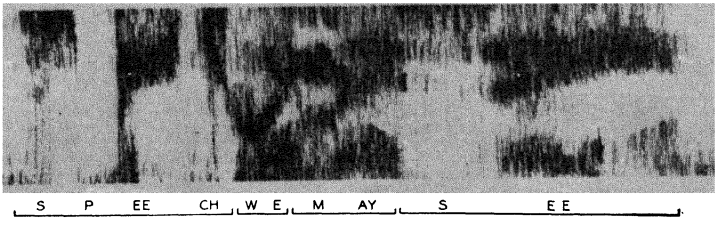


Fig. 8 — Sound spectrogram of the words "Speech we may see" spoken in a whisper, using a wide-band filter (300 cycles).

elements are mounted on the cover of the unit. The analyzer unit contains the variable filter and associated circuits. The equipment on the lower shelf consists of a power supply unit, a loud speaker, and a pre-amplifier and filter for the microphone. Samples of speech having duration time less than 2.4 seconds may be recorded on the magnetic tape. The time required for the analysis is about 2.5 minutes. As normally used in visible speech studies, the spectrograph provides an analysis covering the frequency range from 70 to 3500 cycles in 200 steps. Experimental investigations in which the frequency range was extended to 11,000 cycles indicate that the detail essential to the legibility of the patterns is largely within the 3500-cycle band. Although the different speech sounds are ordinarily spoken with different levels, the level variations are greatly reduced in the analyzing process. This is done in order to portray adequately the detail of each speech sound, and also to overcome the effects of the limited contrast provided by the recording paper.

Several other experimental types of sound spectrographs have been developed for different purposes. In one type, provision was made for the recording of much longer sound spectrograms. In another type, different colors were used instead of different shades of darkness to portray intensity. In still another type, patterns having logarithmic frequency scales were studied. In an effort to duplicate the resolving properties of the ear, the band widths of the analyzing filters employed in this type were made to depend upon frequency. Although the advantages in the visible speech application of these types of visual presentation over that obtained with the present patterns are not outstanding, the whole subject of presenting visual detail that corresponds closely with auditory detail needs further study. This includes not only the effects of filter band width on the portrayal of essential details, but also the effects of the limited range of visual contrast available in recording papers.

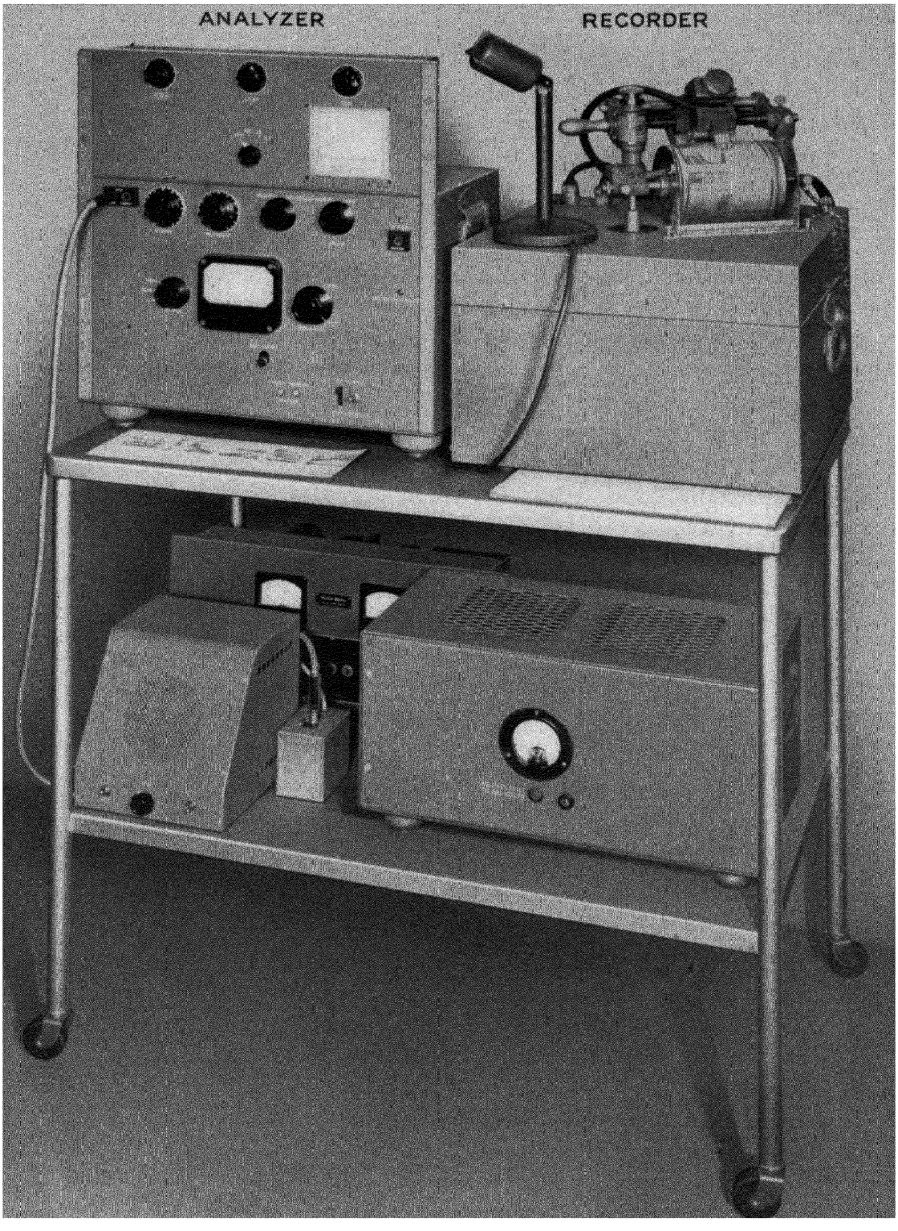


Fig. 9 — Sound spectrograph used in obtaining most of the patterns shown in this book.

Chapter 3

The Direct Translator

The method of translation used in the sound spectrograph, as described in the previous chapter, is not suitable for use in visual hearing conversations. A different method is needed in order to enable the severely deafened to carry on conversations directly, or over the telephone, or to permit them to see radio speech. Such a translation has to be direct or instantaneous; and therefore, the instruments that have been developed for this purpose are called direct translators.

Two general problems are involved in the direct translation from audible to visible speech. First, how to obtain from speech the information necessary to construct the visible patterns. Second, how to use this information in forming the visible display. There are many possible solutions to both problems, and the experimental equipment to be described represents only two of the numerous methods that have been considered. The availability of materials and techniques influenced considerably the choice of the methods of translation that have been used.

Original model

In the original model, a series of twelve fixed filters was employed to perform the analysis rather than one variable filter as used in the sound spectrograph. Each filter passed a 300-cycle band, thus providing for a speech band width of 3600 cycles. (This arrangement of filters does not appear to be the optimum, and translators employing a larger number of filters with overlapping bands and different band widths are being investigated.) The output of each filter was used to form a trace of light on a moving screen of phosphorescent material similar to that used in television receivers. The brightness of the trace was related to the intensity of the speech components within the band passed by the filter. The traces for the various filters were placed side by side along the vertical or frequency dimension. As successive instants of speech were analyzed, the resultant traces remained visible on the moving screen to form a pattern that had a time dimension. In this manner, instantaneous patterns were formed having the dimensions of frequency, intensity, and time. The process is illustrated schematically in Fig. 1. When the words "*one, two, three*" are spoken into the microphone the separate filter outputs are recorded, each in proper frequency relation, as indicated by the schematic pattern in the figure.

In Fig. 2 is shown the first direct translator that was made available to the experimental training group. Needless to say, it is far too large and complicated for other than experimental use. The patterns appear on a screen in the small dark rectangle

near the middle of the vertical panel, and are produced by a new form of cathode-ray tube. This tube consists of a large glass sphere with necks extending from the top and bottom. A band on the inner surface of the sphere is coated with phosphorescent material similar to that used for television tube screens. The entire tube is rotated so that, as viewed from the front, the surface of the sphere moves from right to left. Although the tube rotates continuously, the electron beam that rapidly traces out the patterns under the control of the individual filters remains focused toward the right edge of the viewing aperture. It is held in this position by fixed magnetic field coils within which the sphere revolves. The beam sweeps vertically at a rate of about 500 times per second, scanning in succession the outputs of the twelve filters. Generally, pattern heights of one inch and screen speeds of four inches per second were employed. This adjustment provided suitable patterns for a speaking rate of about 90 to 120 words per minute, and permitted some 3 to 4 words to be visible on the screen at one time.

This cathode-ray translator was used by the visible speech class during most of their training with direct translation. The patterns that were obtained enabled the members of the class to read satisfactorily at the slower speaking rates.

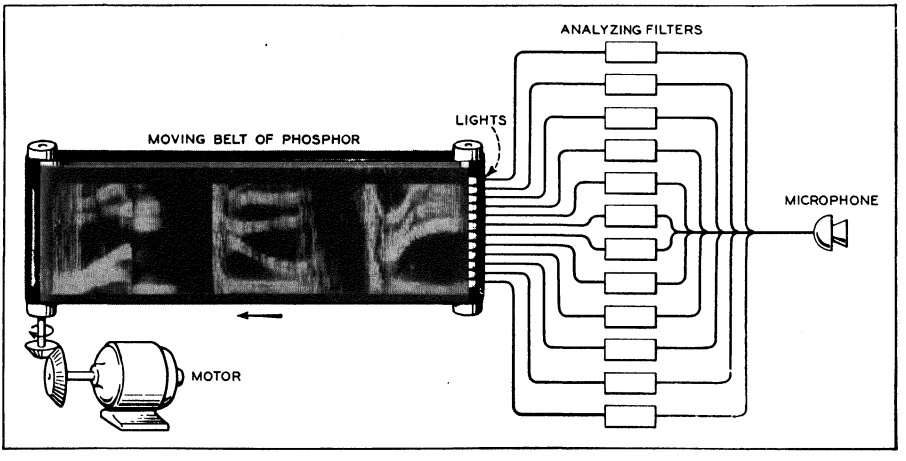


Fig. 1 — Schematic diagram showing how the direct translators convert speech sounds to visible speech patterns.

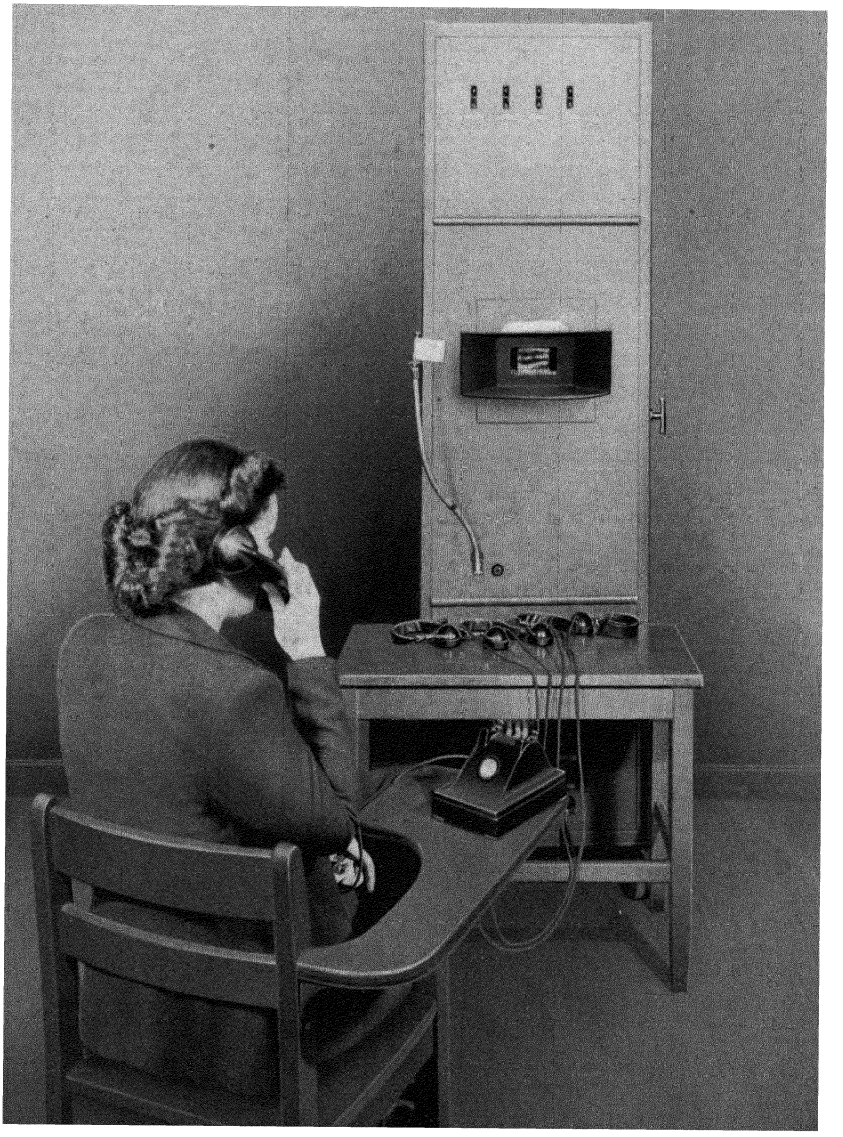


Fig. 2 — First experimental model of a direct translator in which the visible speech patterns appear on the moving screen of a new type cathode-ray tube.

Other experimental units

In Fig. 3 is shown the second experimental model of a direct translator, obviously very much smaller than the cathode-ray model. In this instrument, an attempt was made to produce portable equipment for individual use in the training program. The translator may be operated from a standard power outlet. It weighs about 20 pounds, and is roughly the size of a portable typewriter.

The small translator uses a different method in forming visible speech patterns. Connected to the output of each filter is a tiny so-called "grain-of-wheat" filament lamp. Twelve of these miniature lamps are arranged side by side. They trace the speech patterns in light, upon a rotating cylinder that is covered with a layer of phosphorescent material. The patterns are viewed in the tilted mirror at the top rear of the unit as shown in Fig. 3.



Fig. 3 — Second experimental model of a direct translator in which the visible speech patterns are traced by tiny lights on a moving cylindrical surface of phosphorescent plastic.



Fig. 4 — Experimental model of a large screen translator in which visible speech patterns are traced by small glow lamps on a moving belt of phosphorescent plastic.

Fig. 4 shows a large model of this type that was constructed for demonstration use. In this larger translator, argon-nitrogen glow lamps of about ½-inch diameter replace the grain-of-wheat lamps. The lamps form traces on a continuously moving belt of phosphorescent material. The visible part of the belt is some 10 inches high and 50 inches wide, and is viewed with the aid of a tilted mirror.

Both of the translators employing lamps are still in the early experimental stage and have been used only to a limited extent in the training program.

Pattern improvement

While the translator patterns contain the information necessary for reading words, they do not show the pitch of the voice. For speech training, particularly in the case of the severely deafened, it is essential that the translator also show the pitch used

in speaking. Attempts are being made to display pitch by varying the position of a moving trace and by varying the brightness of a line of light. A satisfactory solution is not yet available, but the indications are that a practical solution is possible.

In designing direct translators there is a problem of evaluating the readability of the patterns obtained. Is this filter arrangement better than that? Is this form of pattern better or worse than that? What is the effect of using phosphorescent materials with different decay and color characteristics? Which type of microphone is better for use with a given unit? What is the advantage of using a wider frequency range? There is no end to the questions that constantly arise in connection with the design and use of the translation equipment.

In addition to subjective judgments, it was very desirable to provide an objective method of pattern evaluation. To meet this need, a type of discrimination test was devised wherein an observer is asked whether one pattern is like either a second or a third in a group of three patterns. The test method as applied to visible speech operates as follows: Words, the patterns of which were likely to be most easily confused, are grouped into a "similarity series." Pairs of words are then selected from these series and arranged in test groups such as "sat, fat, fat" or "do, do, two." These are displayed on the translator under test, and an observer is asked to check whether the middle word is most similar to the first or last of the group. If the basis of selection were pure guesswork, about 50 per cent of the answers would be correct. An index called the "discrimination index" is given a value of zero for the above condition and a value of 100 when all identifications are correct. By making tests with an adequate range of sounds, a fairly dependable comparison of the readability of different displays is possible.

Much remains to be done on direct translator design both with regard to the detail and form in which the information should be portrayed, and the technical means of accomplishing the portrayal. Considerable progress has been made on the requirements for obtaining readable patterns for the person with normal hearing, but experience with deaf readers and methods of teaching them has been rather limited. One of the purposes of this book is to bring visible speech to the attention of those interested in the problems of deafness so that translator design may benefit from direct experience by the users. It is hoped also that interest in the solution of the many technical problems may be stimulated.

Chapter 4

The Training Program

Although it was apparent from the first that anyone could learn to recognize a few word patterns, it seemed desirable to study the legibility of these visible patterns of speech more thoroughly before venturing to suggest that they might be of use to the deaf. There were additional questions to be answered concerning whether it would be possible to learn to read word patterns as rapidly as they occur in ordinary conversation and as spoken by different individuals, and whether learning rates and the requirements of the visual reading task made visible speech a practicable tool. It was primarily to answer these questions that a training program was initiated.

Normal hearing class

The original visible speech class was made up of five young women with normal hearing. A normal hearing group was selected for two reasons: first, it seemed inadvisable to risk the possibility of arousing false hopes among the deaf at this stage of the development; and second, it was desirable to have a normal hearing group available as potential instructors for further experimental training. Initially, the specific objectives of the program were to investigate the legibility of the patterns and to study the teaching and learning requirements.

Before beginning formal class instruction, it was necessary to find out which elements in these patterns of speech were important and which were unimportant, and to adjust the training approach accordingly. Therefore, the first steps following initiation of the training project in September, 1943, were directed towards an investigation of three problems: first, how the various sounds were translated by the sound spectrograph; second, the ways the patterns of similar sounds differed; and third, the visible characteristics of the patterns which would assist in the recognition of sounds and words.

For this purpose, patterns of the basic sound units of English were recorded in both male and female voices. Combinations of these sound units into syllable and word forms also were recorded and studied in order to determine the transitional effects, and to associate the visible patterns with the movements of the articulators as sounds are spoken in succession. During this period, a number of individual conferences were held with the members of the visible speech class. Notes regarding their speech characteristics were filed for later reference, and each member of the class began a recording program to obtain spectrograms, in her own voice, of about 1400 of the most frequently used syllables.

Class schedule

The class met for the first time on September 21, 1943, and a schedule was set up which included two hours of group instruction daily. One hour was devoted to a discussion of phonetic principles and the relationships of articulatory movements to visible speech patterns. The other hour was spent in learning to read the patterns and in testing for the purpose of evaluating achievement. A picture of the members of the original class in visible speech is shown in Fig. 1.

In addition to the two-hour class period, one hour daily was assigned for study, and the remaining available time was spent in making recordings and in assisting with the preparation of instructional material. During the spring of 1944, a sound spectrograph was assigned full time to the experimental training program. From then on, considerable time was used in assembling speech spectrograms to form long reels suitable for



Fig. 1 — The first experimental training group reading projected sound spectrograms.



Fig. 2 — Class group using the first direct-type visible speech translator.

projection at different speeds, and in constructing vocabulary lists for use in connection with the training program. When the first direct translator was ready for class use in the fall of 1944, the class schedule was revised to allow one hour for group work with this instrument and one hour for continued study of the spectrographic material. Within a few months, all of the class instruction was shifted to the translator.

Fig. 2 shows the visible speech class at work with the cathode-ray translator. The translator is at the right of the picture, and the patterns appear within the hood that extends above the small table. The two soundproof booths at the rear of the room permitted a two-way conversation between the occupants, who used only their eyes to "hear." The receivers had been removed from the telephone handsets so that when the booth doors were closed one girl conversed with the other only by way of the translator screen. Members of the class outside the booths were able to watch both sides of the conversation, and the instructor could break in at any time to comment or to correct errors.

Although the reading of visible speech patterns in their present form requires close attention, the class members have been able to converse for periods of two hours or

more without evidence of visual strain or general fatigue. This is true even though the brightness of the patterns made by the cathode-ray translator requires that the room be partially darkened. However, much more experience is needed to determine the visual task involved in reading visible speech. Decisions that have been made concerning the number of words visible on the screen, the size and color of the patterns, and the visual angle employed, may need to be modified in the light of further information. An attempt was made in selecting the height of the present patterns to eliminate the necessity of vertical eye movements. The optimum eye span, the number of fixations and regressions, and the left-to-right direction of the horizontal eye movements are under investigation. The relationship between ordinary silent reading and the reading of visible speech cannot be studied readily until more people have learned to read visible speech.

Deaf member added to training class

In December 1944 a congenitally deaf engineer was added to the training group. He is a proficient lip reader and had obtained the degree of Master of Science in engineering from one of the large universities. We were interested in having someone of his qualifications associated with the project, in order to obtain a better understanding of the training problem as applied to the severely deaf. In the first interview, he could not be told about the scope of the investigation owing to the military classification then effective. When the project was described as one concerned with the development of a possible aid for the deaf he was obviously disappointed, saying he lacked faith in hearing aids as applied to his case. However, he did join the group and, when he understood the possibilities of visual hearing, his enthusiasm was boundless. His presence in the class provided an opportunity to obtain some evaluation of the use of the translator in speech correction. Although the instrument does not portray pitch variations, it has been of material assistance in improving his speech during the past year. The translator has also provided a means for this congenitally deaf engineer to make and receive telephone calls. It may be of interest that he carried on his first telephone conversation within the Laboratories building on February 2, 1945, and that he received his first telephone call from the outside on August 28, 1945.

Several young women were added to the training group during 1944, and the progress of these newer members is of particular interest because it is probably the best available indication of the possible rate of progress in training. The progress chart in Fig. 3 shows the rate of increase of vocabulary for three cases. Curve A represents "Hours of Training" versus "Number of Words" for the case of the congenitally deaf engineer. Curve B is for two girls that started training later, and Curve C is for two more girls starting at a still later date. Within the range of experience with these

trainees, their rates of progress seem to be about the same. The words used in this training program were selected from several well known lists.*

It should be mentioned that the rate of acquiring a vocabulary and its size do not give a complete indication of reading proficiency or pattern legibility. The existing visible speech translators, while usable, are still far from perfect. Important pattern details are not sufficiently emphasized, and there may be present and overemphasized a good deal of material that does not contribute significantly to legibility. The discrimination tests mentioned in Chapter 3 indicate that the best of present translators is about equivalent to a rather poor telephone circuit. When conversing in a familiar vocabulary, and talking slowly and distinctly with someone we know well, the exchange of information over such a marginal telephone circuit is not too difficult. But if we are called upon to talk in a new and not too familiar vocabulary, at faster rates, and with someone whose voice is not familiar, the exchange is not likely to go so well. Although it might be possible to accumulate an extensive vocabulary in an unfamiliar foreign language, even over a difficult telephone circuit, it would probably be impossible to converse rapidly and without attention to clear enunciation.

The visible speech class members were able to converse satisfactorily among themselves by talking clearly and at a fairly slow rate. Within the limits of their vocabularies, they were able to carry on conversations with about the same facility as a similarly advanced class in some foreign language. When new and entirely unfamiliar words were displayed on the translator screen, the more experienced students usually were able to read the words after a few repetitions. The difficulties in this case illustrate the analogy to a marginal telephone circuit over which it may be necessary to repeat an unfamiliar word before it can be understood.

Pattern development needed

The training experience indicates then that there is a firm basis underlying the legibility of the visible speech patterns, but that the present translators are considerably short of perfection. To improve pattern legibility, the designers must rely for guidance upon information obtained in learning to read visible speech and from tests

---

* GATES' WORD LIST FOR PRIMARY GRADES. *Arthur I. Gates*, Teachers College, Columbia University, 1935.

TEACHERS WORD BOOK OF 20,000 WORDS. *Edward L. Thorndike*, Teachers College, Columbia University, 1932.

TEACHERS WORD BOOK OF 30,000 WORDS. *Edward L. Thorndike* and *Irving Lorge*, Teachers College, Columbia University, 1944.

THE WORDS AND SOUNDS OF TELEPHONE CONVERSATIONS. *N. R. French, C. W. Carter, Jr.*, and *Walter Kornig, Jr.*, American Telephone and Telegraph Company, June, 1930.

BASIC ENGLISH AND ITS USES. *I. A. Richards*, Norton & Company, New York, 1943.

of the discrimination type. Neither the determination of requirements nor their fulfillment is likely to be a simple matter. One very important responsibility rests upon those closely associated with the problems of deafness. These specialists will need to study the application of visible speech to their problems, and to be in a position to tell the engineer what they think is needed. From now on, if the development is to be completely successful, it must proceed on a cooperative basis.

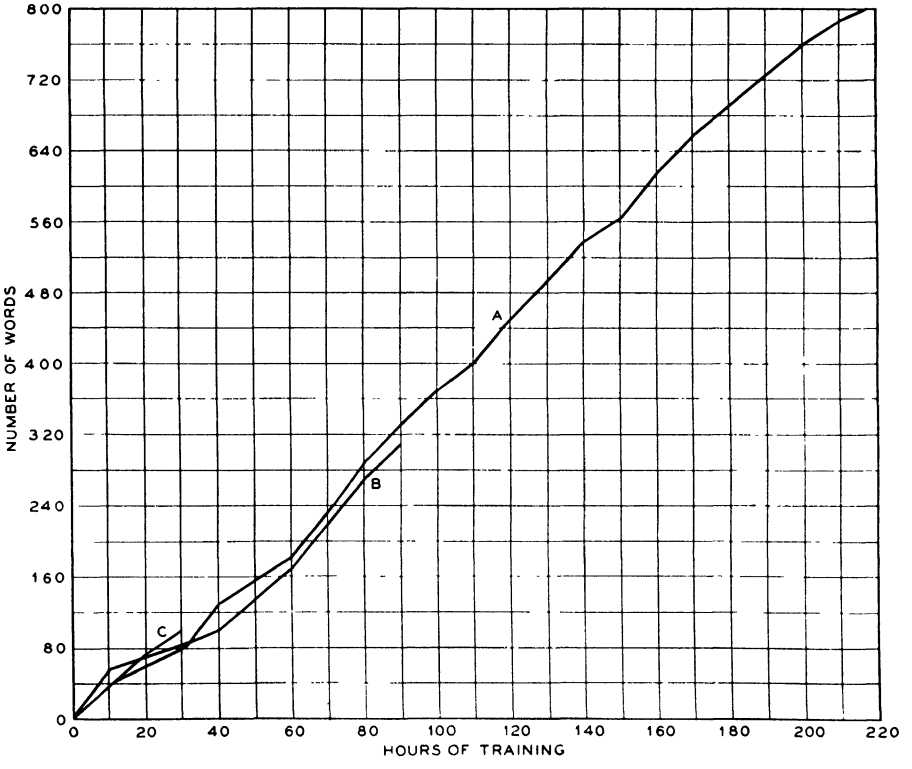


Fig. 3 — Progress chart. Case A — Congenitally deaf engineer. Cases B and C — Two groups of two girls each.

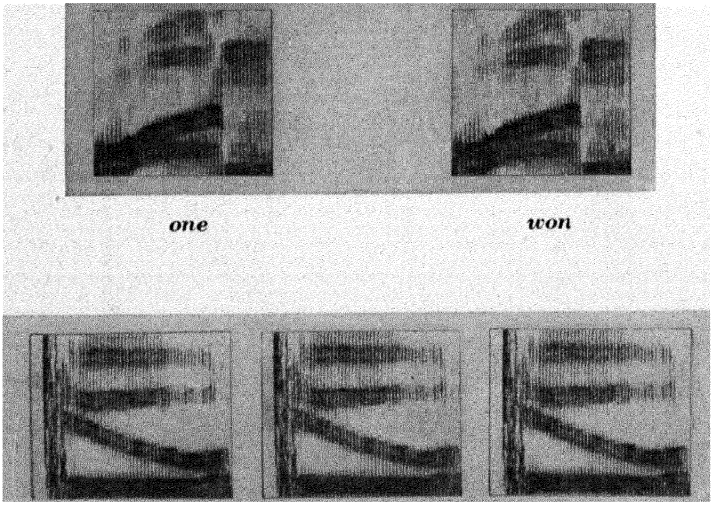
Chapter 5

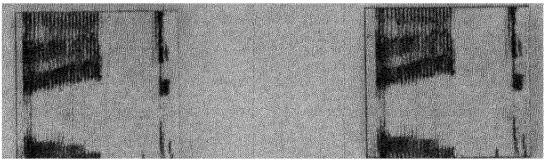
The Basic Principles of Visible Speech

The New ABC's

The background of the visible speech development, the methods used in translating speech sounds to visible patterns, and the early experiences in reading the patterns, have been presented in the preceding chapters with both the general and the professional reader in mind. This chapter has been written to meet the special interests both of those individuals who do, and those who do not, intend to learn to read visible speech. To the former, it will serve as an introduction to the lesson chapters where the significant characteristics of sound and word patterns are discussed in detail in progressive study units. To the latter, it will provide a general summary of the basic concepts underlying the legibility of visible speech. In order to accomplish these purposes, this chapter includes discussions of the basic speech processes, their relation to the visible characteristics in the patterns of speech sounds, and the use of these characteristics in reading the patterns. Special terms used to describe the patterns also are introduced.

It seems important to recognize from the beginning, that visible speech patterns are natural phonetic symbols and correspond, therefore, to the way *sounds*, *syllables*,





eight ***ate***

and *words* are spoken, rather than to the way in which they are spelled. Consequently, as illustrated in the accompanying figures, words that are spelled differently may have visible patterns that look alike.

The patterns shown above look alike for the words that are spoken alike, regardless of the spelling and meaning of those words. Some sounds have a large number of possible spellings. An example is the *a* as in *ate*, which is illustrated by the following key words.

a (*ate*)	***au*** (*gauge*)	***ei*** (*eight*)
ai (*pain*)	***ao*** (*gaol*)	***ea*** (*break*)
ay (*day*)	***e*** (*eh*)	***ey*** (*they*)

In reading visible speech it is important to remember that *words that sound alike, look alike*, no matter how they are spelled.

Some system of printed symbols was required in order to describe the phonetic relationship between the visible speech patterns and the spoken or printed words that were used in obtaining them. It has seemed advisable to adopt the symbols of the International Phonetic Alphabet for this purpose. Their use affords the accurate designation of a wide range of phonetic values in several languages. To meet the needs of those who are unfamiliar with this alphabet, common American key words have been used with the symbols throughout the text. In addition, an appendix has been provided listing for each of the different sounds, the I.P.A. symbols, the symbols and diacritical markings used in Webster's Dictionary, and common American key words.

The lesson chapters contain a large number of illustrations and it was essential to use some means for indicating on these patterns, the particular details mentioned in the text. The pointer scheme selected seemed the least offensive and the most efficient of the several considered when all the requirements were taken into account. The various "arrowhead" shapes adopted are shown in Fig. 1.



Fig. 1 — "Arrowhead" shapes as they appear on the patterns and in the text.

When the arrowheads point upwards or downwards, to the right or to the left in the text, they point in the same direction in the illustrations. They should be thought of as indicators designating particular regions or parts, rather than specific points, in the patterns under discussion.

Diagrammatic profile pictures are used to illustrate the positions of the tongue, lips, and other articulators that are used in producing the speech sounds. These profiles are intended to acquaint the reader with the relationship between the speech movements and the visible speech patterns. It is highly desirable to develop a kinaesthetic awareness of the speech movements associated with the production of speech sounds in order to assist in visualizing and recalling their patterns.

The articulatory positions which appear in the profile diagrams have been selected because they are the positions that a majority of American speakers use in producing the sounds of our language. Of course, any one individual will tend to vary a little from some or all of these positions, and when his production of the sound differs, his visible speech patterns will vary too. Fortunately, as is explained subsequently in this chapter, these variations usually are not sufficient to cause difficulty in reading visible speech.

Basic Speech Processes

The speech mechanism may be regarded, for our present purpose, as composed of the five parts or processes shown in Fig. 2. Any one part of the mechanism may be described only as static and non-functioning, since in speech all of the parts interact as a highly coordinated unit. While a detailed description is outside the scope of this book, the reader will find that some acquaintance with the basic elements of speech production is necessary to the understanding of the principles underlying visible speech. Fundamentally, one speaks by modulating the flow of air which passes from the lungs through the windpipe, vocal cords, throat, nose, and mouth cavities.

The term *modulation* is used to describe how speech is produced because it expresses concisely what happens to the air as it flows from the lungs. As used here, the term *modulate* means to alter, vary, or regulate the flow of air. There are four types of modulation in speech which are of particular importance in reading visible speech patterns. These are start-stop, vocal cord, frictional, and cavity modulation. The first three primarily affect the flow of air from the lungs. Start-stop modulation starts and stops the flow. Vocal cord and frictional modulation convert the flow into audible sound waves. Cavity modulation varies the quality of the sound.

The start-stop modulation, which may be produced either by the vocal cords or the articulators, occurs at relatively low rates, that is, less than 10 times per second. The corresponding frequencies are not heard as tones but are evidenced as pauses in the flow of speech.

Vocal cord modulation is produced by the vibrations of the vocal cords which periodically interrupt the flow of air. This results in sound waves which, when modulated by the vocal cavities, give the characteristics of *voicing* to speech. Vocal cord modulation occurs at fundamental frequencies between approximately 75 cycles per second (low pitch) and 500 cycles per second (high pitch). Because of the nature of the modulation, not only the fundamental tone but a large number of harmonic overtones of the fundamental are produced.

Frictional modulation is characteristic of the group of sounds like the ***s***, ***f***, and ***sh***, that are known as fricatives. It is produced by the articulators being placed close enough together to form a small opening or constriction through which the air must flow. Frictional modulation, caused by turbulent air flow in the constriction, produces simultaneously a wide range of frequencies that are randomly spaced and inharmonically related.

Cavity modulation occurs in the coupled cavities of the throat, mouth, and nose, and acts on the overtones produced by both vocal cord and frictional modulation. These coupled cavities suppress some of the overtones, which results in an apparent reinforcement of other overtones. In other words, these cavities have the property of selective transmission and radiation. The frequency regions in which the overtones appear to be reinforced are referred to as vocal resonances. They are determined by the shapes and sizes of the connected cavities, and are changed by the movements of the articulators, that is, the lips, jaw, tongue, and soft palate. As sounds are formed in succession, the positions of the articulators undergo changes at relatively slow rates, and thus modulate the audible overtones in speech. Since the surfaces of the articulators form the walls of the resonance cavities and change their size and shape, the modulation that takes place in the resonance cavities may be thought of as one type of articulatory modulation. It is important to mention that this cavity modulation affects the sounds produced by frictional modulation in the front of the mouth much less than it does those produced by vocal cord modulation.

In speaking, combinations of all the types of modulation mentioned above are used. Stop, frictional, and cavity modulation are combined in producing the sounds ***p***, ***t***, and ***k***. These sounds are usually called *stop* or *plosive* sounds. The term *stop* designates one phase of such sounds because the air flow is momentarily stopped in their production. The term *plosive* applies both to the stopping and releasing of the air flow. The *puff* which characterizes the quick release of the stopped air flow is fricative in nature. Thus, these sounds may be called more accurately stop-frictional sounds because they are produced primarily by a combination of stop and frictional modulation. When vocal cord modulation is added, the voiced stop sounds ***b***, ***d***, and ***g*** are produced.

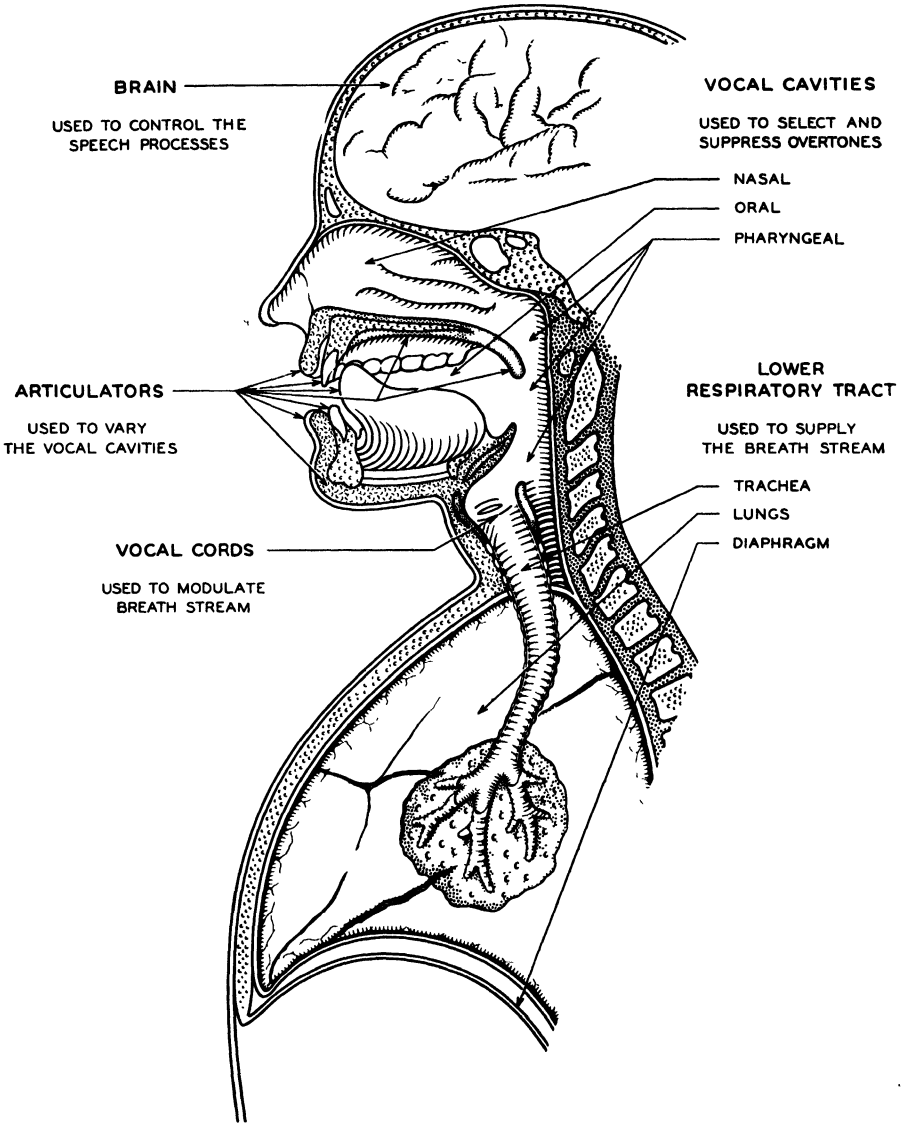


Fig. 2 — Schematic diagram of the basic speech processes.

When vocal cord, frictional, and cavity modulation occur together, the voiced fricatives like *z*, *v*, and *zh* are produced.

Only vocal cord and cavity modulation occur in the production of the vowels and vowel-like sounds.

There are a number of additional kinds of modulation such as stress, inflection, vibrato, and vocal quality, which determine general speech characteristics. Although they are of considerable interest in speech and voice studies, they play a relatively minor role at the present time in the reading of visible speech.

The Six Sound-Pattern Groups

Since visible speech is a natural phonetic translation of speech sounds, the nature of the visible speech patterns is closely related to the manner in which the speech sounds are produced. Consequently, when the sounds are classified either according to their visible speech patterns or to the manner of their production, they fall naturally into the six groups that are used generally in phonetic classifications.

We shall discuss first, the features which are characteristic of each of the six sound-pattern groups. Then we shall examine in greater detail, the means by which patterns within each group are differentiated. In the discussion, a number of terms such as *gap*, *fill*, *spike fill*, *voice bar*, *resonance bar*, and *baseline* have been introduced in order to describe certain characteristic features of the patterns. They are defined and illustrated in the following paragraphs and they will be used throughout the text.

Group 1. The voiceless stop sounds — p (*pay*), t (*to*), k (*key*)

These sounds are produced primarily by a combination of stop and frictional modulation. The sounds are made by stopping the breath flow at some point in the articulatory tract, building up breath pressure, and then rapidly releasing the breath. The sound p as in *up* is an example.

The diagrammatic profile shows the compression of the lips which produces the stop for the p. The stop is shown in the pattern by a blank space or *gap* ∧. Whenever



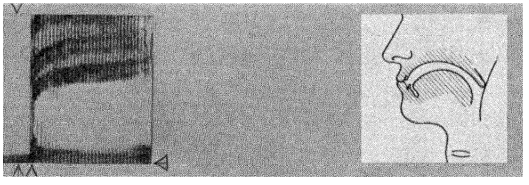
up *Profile of* p

the breath flow is stopped, a similar gap or blank space will appear in the pattern. The release of breath when the lips are opened suddenly, produces the *spike fill* ∧ in the pattern. The irregular vertical striations which result from frictional modulation are called a *fill* because they tend to fill in a segment of the pattern. When stop modulation precedes frictional modulation and the breath flow is released suddenly, the fill usually is narrow and begins abruptly. Such a fill is called a *spike fill* or more simply a *spike.*

The patterns of the voiceless stop sounds are characterized by *gaps* followed by *spikes.*

Group 2. The voiced stop sounds — b (*be*), d (*day*), g (*go*)

These sounds are produced primarily by a combination of stop, vocal cord, and frictional modulation. The sounds are made by stopping the voiced breath stream at some point in the articulatory tract, building up voiced breath pressure, and suddenly releasing the voiced breath. The sound b as in *be* is an example.



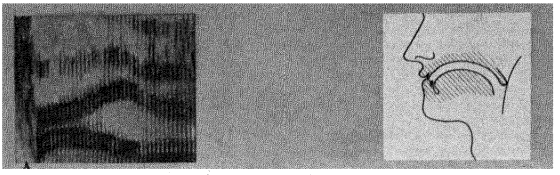
be ***Profile of* b**

Curved lines are used in the diagrammatic profile to represent the vocal cords in the larynx or Adam's apple. The vibratory modulation which results from vocal cord action produces the narrow bar ∧ on the baseline ⋖ of the pattern. Because this bar is present whenever a voiced consonant is made, or whenever vocal cord modulation is present in a consonant, it is called a *voice bar* or *voicing bar.* The profile also shows the lips compressed to stop the breath which is shown on the pattern by the *gap* ∨ above the voice bar. Such a gap-voice bar combination is called a *voiced gap.* The sudden release of the breath produces the *spike* ∧ which follows the *voiced gap.*

The patterns of the voiced stop sounds are characterized by *voiced gaps* and *spikes.*

Group 3. The voiceless fricative sounds — h (*he*), f (*for*), θ (*thin*), s (*see*), ʃ (*she*)

These sounds are produced primarily by frictional modulation. The sounds are made by forcing a continuous breath stream through a small or restricted opening. The restricted opening may occur at one of several places in the articulatory tract or at the vocal cords. The sound f as in *five* is an example.



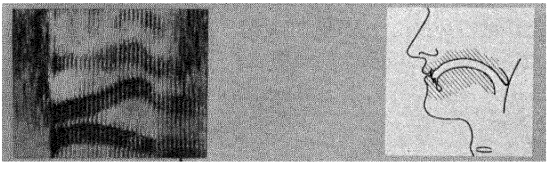
five *Profile of* **f**

The diagrammatic profile shows the restricted opening between the lower lip and upper teeth through which the breath stream is forced for the **f**. This frictional modulation of the breath produces the irregular vertical striations on the pattern which are called a *fill* Λ. Because the breath may be emitted slowly and continuously, the *fills* for the voiceless fricative sounds are generally wider than the spike fills for the stop sounds.

The patterns of the voiceless fricative sounds are characterized by *fills.*

Group 4. The voiced fricative sounds — ɦ (*ahead*), v (*vote*), ð (*then*), z (*zoo*), ʒ (*azure*)

These sounds are produced primarily by a combination of vocal cord, frictional, and cavity modulation. The sounds are made by forcing a continuous and voiced breath stream through a restricted opening which may occur at one of several places in the articulatory tract or at the vocal cords. The **v** as in *five* is an example.



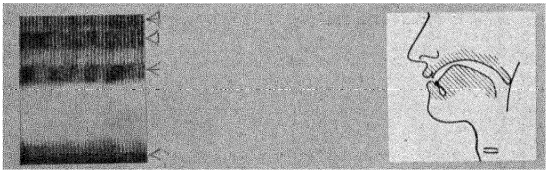
five ***Profile of* v**

The curved lines in the profile represent the vocal cord action used to produce the voiced breath stream for the **v**. This vibratory modulation is shown by the *voice bar* or *voicing bar* along the baseline of the pattern. The profile also shows the restricted lip-teeth opening through which the breath is forced for the **v**. The frictional modulation thus produced, results in the *fill* Λ which appears above the voice bar. When the voicing for the **v** becomes very strong in relation to the fricative element, the vocal resonances of the **v** become more evident. When the voicing is weak and the frictional element is relatively strong, the *fill* is more prominent and only the voice bar is present. The combination of voice bar and fill often is called a *voiced fill.*

The patterns of the voiced fricative sounds are characterized by *voice bars* and *fills,* or by *voiced fills.*

Group 5. The vowel and vowel-like sounds — i (*eve*), ɪ (*it*), e (*hate*), ɛ (*met*), æ (*at*), a (*ask*), ɑ (*father*), ɒ (*not*), ɔ (*all*), o (*obey*), ʊ (*foot*), u (*boot*), ɝ (*word*), ɜ (*word*), ʌ (*up*), ə (*about*), j (*you*), w (*we*), l (*let*), r (*read*), m (*me*), n (*no*), ŋ (*sing*).

The vowels and vowel-like sounds are produced primarily by a combination of vocal cord and cavity modulation. They are produced by the passage of the voiced breath through different combinations of the coupled vocal cavities. These cavities are modified and combined or coupled by the action of the articulators. The vowel i as in *eve* is an example of these sounds.



i (*eve*) ***Profile of* i**

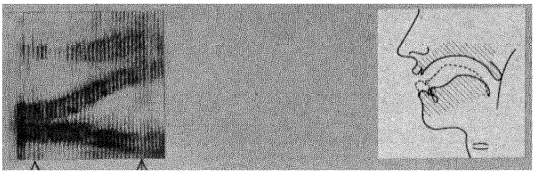
The vocal cord action used to produce the voicing is represented in the diagrammatic profile by the curved lines at the larynx. The profile also shows the positions of the articulators as they form the vocal cavities used in saying the sound i (*eve*). The particular combination of vocal cord and cavity modulation used for the i, produces the vocal resonances which are characteristic of this vowel. These resonances are shown in the i pattern by the horizontal *resonance bars.* For convenience in referring to them, the *bars* are numbered beginning at the baseline as *bar 1* ≺, *bar 2* ≼, *bar 3* ≺, and *bar 4* ≼. The positions of the bars correspond to the frequencies at which the vocal resonances occur. The regularly spaced vertical striations which appear throughout the pattern may be ignored in reading. They correspond to the pitch of the voice and, like the other representations of voice quality and inflection, are not used at present in reading the patterns.

The patterns of the vowels and vowel-like sounds are characterized by combinations of *bars.*

Group 6. Combinations of sounds — eɪ (*say*), aɪ (*I*), ɔɪ (*boy*), aʊ (*out*), oʊ (*go*), ɪu (*new*), hw (*when*), tʃ (*church*), dʒ (*judge*).

Sounds within a group, or sounds from different groups often are combined in words. Such combined sounds as tʃ (*church*), and dʒ (*judge*) are used together so frequently that they may be considered as separate sound units. When two vowels are combined as in aʊ (*out*) or aɪ (*I*), the combination is called a diphthong. The diphthong aɪ (as in *I, eye,* or *five*) is an example of a combination of two Group 5 sounds.

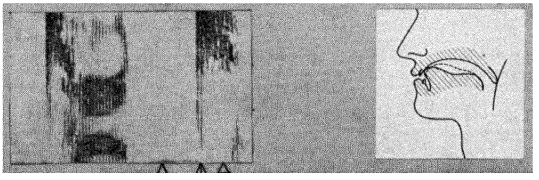
The diagrammatic profile shows the positions of the articulators that form the vocal cavities for the two vowels, a (*ask*) and ɪ (*it*), which are combined in the aɪ. Vocal



aɪ (*I*) ***Profile of* aɪ**

cord action is represented by the curved lines in the larynx of the profile. The vocal cord-cavity modulation is shown by the *bars* in the pattern of the **aɪ**. The bars curve from their positions for the **a** (*ask*) ʌ to their positions for the ɪ (*it*) ʌ.

The combination of the **t** (*to*) and **ʃ** (*she*) as in *church* is a very common example of **a** Group 1 — Group 3 combination.



church ***Profile of* tʃ**

The profile illustrates the positions of the articulators used in making the **tʃ** (*church*) combination. The stop-frictional modulation for the **t** is shown on the pattern by the *gap* ʌ followed by the *spike fill* ʌ. The frictional modulation for the **ʃ**, as shown by the *fill* ∆, follows directly after the **t**. In saying this and similar combinations, the movement often is made so quickly from one articulatory position to another, that the breath for the first sound seems to be released along with the breath for the second sound. The many other frequently used combinations of sounds such as **st** (*stay*), **sk** (*ask*), **pl** (*play*), **tr** (*tree*) and **sl** (*sleep*), are omitted generally from recognized phonetic alphabets, but they are of considerable importance in reading visible speech patterns.

The patterns of combined sounds are characterized by combinations of the conspicuous features of their separate sounds.

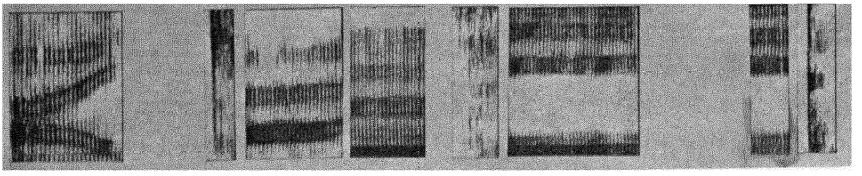
There is a tendency in normal speech to run words together, and not to separate each word from the next by a definite space as in writing. At the points where words or syllables adjoin, there may be sound combinations that are never found within syllables as, for example, **ð ð** between the words "*with this*". There are many of these between-syllable combinations.

In addition, numerous sound combinations may occur both within a syllable, as the **st** in *past*, or between two adjoining syllables, as the **st** in *pastime*. The main differ-

ences between the patterns made by the combined sounds in these two cases, are found usually in *timing* and *stress.* The amount of time given to sounds depends upon how fast a person speaks, and upon the stressing and unstressing of syllables. Between-syllable sounds sometimes may be said as quickly as sounds in a syllable, but because of stress, the pattern of a sound combination occurring within a syllable generally makes a pattern different from that of the same combination occurring between two adjoining syllables.

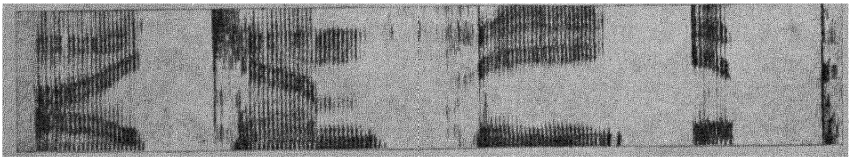
Sounds Influence Each Other

Each of the sounds in the six pattern groups has a distinctive visible pattern when it is spoken by itself. It would be convenient for reading if each sound always had only one pattern, but as we combine sounds into syllables and words in connected speech, one sound influences and is changed or influenced by the sound with which it is combined. All sounds are not influenced in the same way, and for that reason the visible speech reader must become familiar with the principles basic to sound combinations.



I	*c*	*a*	*n*	*s*	*ee*	*i*	*t*
aɪ	k	æ	n	s	i	ɪ	t

Each of the vowel and vowel-like sounds in the sentence, *I can see it,* is shown in the above picture as it appears when spoken separately. The bars for the vowels, æ (*can*), the i (*see*), and the ɪ (*it*), are straight because they were made with the articulators in fixed or steady state positions.



I	*can*	*see*	*it*
aɪ	kæn	si	ɪt

When the sounds are combined to make words, and the words are combined to make a sentence, the influence of the sounds on each other is apparent. Compare the æ (*can*), i (*see*), and ɪ (*it*) as they are spoken in the sentence with the same vowels as they were said separately. The beginning of bars 2 and 3 of æ (*can*) are curved by the **k**, and the ends of these bars are curved by the **n**. Bars 2 and 3 of i (*see*) are curved by the **s**, and the ɪ (*it*) bars are changed by the **t**.

These influences appearing in the patterns, are produced when the articulators change from positions characteristic of one sound to positions characteristic of the succeeding sound. The patterns indicate that the articulators spend about as much time in the transitional movements as they spend in the characteristic or steady state positions. Because these transitional influences are produced by the movements of the articulators, the patterns that we make in moving from one sound to another help the reader of visible speech to associate the appearance of the patterns with the feeling of the movements of the articulators. The basis of how sounds influence each other is in the way the resonating cavities change in shape and size as we say one sound after another in pronouncing words. When a bar moves from one position to another in a pattern, it means that the vocal resonance associated with that bar is changing from one frequency region to another.

Although all sounds are changed to some extent by the sounds that come before and after them, it is important to remember that the same sounds said together in the same way will always look the same.

The Hub—The Key to How Sounds Change Each Other

The shapes and positions of the resonance bars are basic to the reading of visible speech. Of the three resonance bars that are frequently present in sound patterns, the second bar is the one to which we shall refer most often. Bar 1 (the bar nearest the baseline) is limited in movement to a small interval at the bottom of the pattern, while bar 3 is both limited in its movement and frequently weak or absent from the pattern. Because bar 2 is changed in position and shape more than either of the other bars, and because of the more consistent manner in which the second bars of sounds are connected, bar 2 is considered the most significant bar to watch in identifying sound patterns and sound combinations. But it should be emphasized that bar 2 derives its significance from its relationship with the first and third bars. Although the major identifying clues are found in the second resonance bar, we do not read by the second bar alone. We read visible speech by looking at the entire area of the pattern, from left to right, as though we were reading successive chords on a sheet of music. Just as in reading music we see the treble and bass clefs simultaneously, so in reading visible speech we see the two principal kinds of clues. These clues are the individual sound patterns and the way the sound patterns are combined. The total visual impression is a

configuration which is related both to the baseline and to the top of the pattern. The importance of bar 2 lies in the fact that it plays a greater role in changing the total configuration than do the other bars.

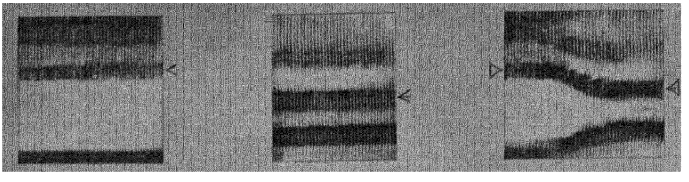
The hub is the position of bar 2.

Due to the frequency of occurrence of the phrase, *the position of bar 2*, in this and subsequent chapters, we shall use the term *hub* to denote the position of bar 2 to avoid repetition of the entire phrase. *The hub is defined as the visible or hidden position of bar 2 of any sound when the sound is made alone.* The definition of the hub may be illustrated by several representative vowel and consonant patterns.

The hub may be visible or hidden

Hubs of the vowels

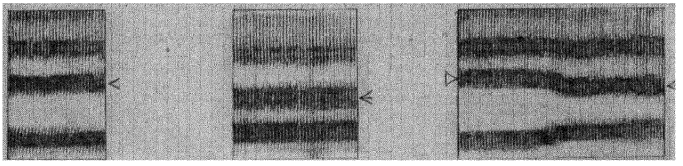
Since it is frequently difficult for an untrained ear to detect the individual vowels which are combined in common diphthongs like eɪ (*pay*) or aɪ (*I*), we have selected less common vowel combinations to illustrate the hub principle. The following patterns show (1) that the hub of any vowel is the visible position of its bar 2 when the vowel is made alone; (2) how the hubs are connected when vowels are combined.



i (*eve*) æ (*at*) i (*eve*) *combined with* æ (*at*)

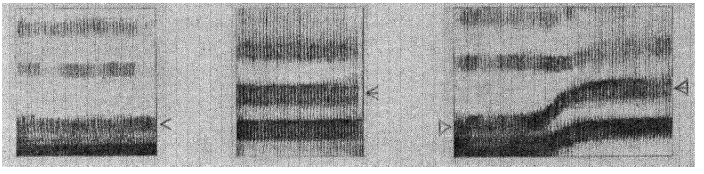
When the vowel i (*eve*) is made, its hub or second bar is in the position indicated at ≺. When the vowel æ (*at*) is made, its hub or second bar is in the position indicated opposite ≺. As the i is combined with the æ, bar 2 of i curves downwards from the i hub ⊳ to the æ hub ≺. The curve is downwards because the i hub is higher than the æ hub. Notice that bar 2 looks like a road that connects one place with another place on a map. The *places* connected are the hubs of the vowels. The curves in the *road* mean that a transitional sound is made as the first vowel hub is connected with the second vowel hub. The pattern of the transitional sound reflects the change in the vocal resonances as the articulators move from the i position to the æ position.

The hub of the vowel ɛ (*bet*), in the illustration at the top of the next page, is opposite ≺ and the hub of the vowel æ (*at*) is opposite ≺. When these two vowels are combined, bar 2 curves downwards from the ɛ hub ⊳ to the æ hub ≺. Again the curve



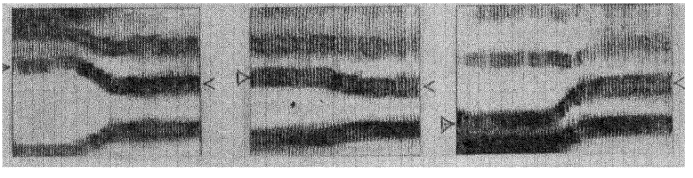
ɛ (*bet*) æ (*at*) ɛ (*bet*) *combined with* æ (*at*)

is downwards because the ɛ hub is higher than the æ hub. Notice, however, that the positions of the two hubs are not very far apart. As the two vowels are combined their hubs are connected by a transitional sound which is neither an ɛ nor an æ.



u (*boot*) æ (*at*) u (*boot*) *combined with* æ (*at*)

The hub of the vowel u (*boot*) is opposite ≺ and the hub of the vowel æ (*at*) is again indicated at ≺. When these two vowels are combined bar 2 curves upwards from the u hub ⋗ to the æ hub ≺. The curve is upwards in this combination because the u hub is lower than the æ hub. Because the positions of the two hubs are quite far apart, there is a marked transitional curve of bar 2. This distance between the hubs indicates that in the transition from the u position to the æ position there is a considerable shift both in the vocal resonances and in the articulatory positions.



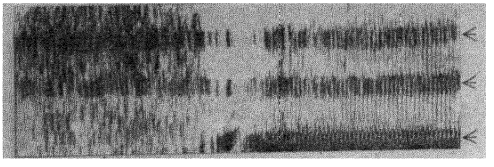
i (*eve*) *combined with* æ (*at*) ɛ (*bet*) *combined with* æ (*at*) u (*boot*) *combined with* æ (*at*)

Compare the combinations of vowels which you have seen separately in the preceding illustrations. Notice that (1) the æ hub ≺ is in approximately the same position in each pattern, (2) the i hub ⋗ is higher than the ɛ hub ⋗ and the ɛ hub is higher

than the **u** hub ⩾, (3) there are differences in the direction and extent of the curves between the i–æ, the ɛ–æ, and **u**–æ combinations. It should be noted that the transitional curve of bar 2 between two vowel hubs becomes steeper as the rate of speaking increases. The preceding illustrations were made at a normal rate of speaking.*

Hubs of the consonants

We will see next that consonants influence the positions and shapes of the bars of the vowels with which they are combined in much the same manner as one vowel influences another vowel. The reason is that the consonants have potential vocal resonances which may or may not produce visible bars in their patterns. These potential vocal resonances of a consonant do, however, clearly influence the bars of the vowel and vowel-like sounds with which it is combined. *The hub of any consonant is the visible or hidden position of its bar 2.* When the hub of a consonant is hidden, its position is indicated by the way the second bar of associated sounds tends to curve toward the area in which bar 2 of the consonant would be located if it were visible.



s (*see*) **z** (*zoo*)

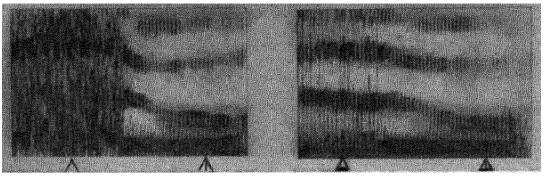
If we make the **s** (*say*) sound ⋀ and then, without in any way changing the shapes or sizes of the resonating cavities, vibrate the vocal cords, resonance bars will appear in the **s** fill ⩽. This shows that the resonance conditions corresponding to the **z** bar pattern were present when the **s** sound was made. The hidden bar 2 position of **s** is the same as the visible bar 2 position of the voiced **s** or **z**. The bars are seldom as visible in the pattern of **s** as they are in the pattern of the voiced **s** or **z** because:

1. The **s** is made without vibration of the vocal cords.
2. The position in the front of the mouth from which both the **s** and **z** sounds are made, permits only a slight resonation of the voiceless breath.

Frequently, the **s** is resonated in a cavity formed between the lips and the teeth. When this occurs, shadows of upper bars may be visible on the pattern of **s**.

In the visible speech patterns of the words *sue* and *zoo* on the next page you can see that the hidden bar patterns of **s** (*sue*) ⋀ influence the bar patterns of the vowel **u**

* Unless otherwise indicated, all of the patterns shown in this book were made at a normal rate of speaking.



sue *zoo*

(*sue*) ⅄ in the same way that the visible bar patterns of **z** (*zoo*) Λ influence the vowel **u** (*zoo*) Λ. The positions and movements of the articulators are the same in saying *sue* and *zoo*. Since the hub of the **s** is in the same position as the hub of the **z**, the **s** and **z** have the same influence on the vowel **u** (*sue*, *zoo*) or on any other vowel with which they are combined. It should be mentioned that as sounds are voiced more strongly their resonance bars tend to become more prominent, and the transitional influences are more easily observed.

There are many consonants in our language that are made in pairs, as for example, the **s** (*sue*) and **z** (*zoo*). This means that normal speakers of English produce both sounds in the pair from approximately the same articulatory positions although one sound of each pair is *voiceless* and the other is *voiced*. The hidden bars of the voiceless sound are usually in the same positions as the visible bars of its paired voiced sound. This is true because the potential vocal resonances are the same for the voiceless and voiced sounds of each pair of consonants.

Although the significance of bar 2 has been stressed somewhat in the explanation of the hub concept, we cannot emphasize too strongly that in reading visible speech it is the total pattern that contains the information or meaning, and that the significance of bar 2 lies in its relationships to the other parts of the pattern.

The hub of any sound has been defined as the visible or hidden position of bar 2 when the sound is made alone.

Essential Characteristics of the Hub Principle

In reading visible speech we must observe and be able to recognize:

1. The characteristic pattern of a sound.
2. The way a sound influences, and is influenced by, other sounds.

The ability to recognize characteristic sound patterns enables us to read the patterns made by combined sounds. In other words, before we can determine how sounds influence each other, it is necessary to know what the sounds look like when they are made separately. The total sound pattern provides the needed reference or standard by which we identify the sound in combinations, but it is the hub which is the key reference point.

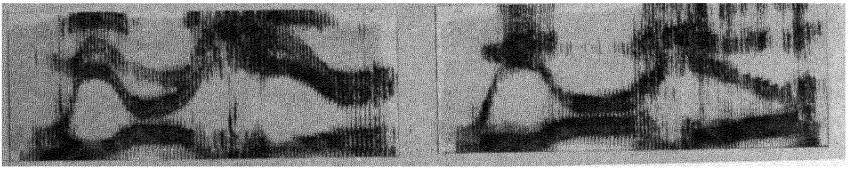
It would be logical to infer that the importance of the hub concept can be demonstrated best by combinations of sounds, for it is in the patterns of combined sounds that we can see how the hubs of sounds are connected. In examining the essential characteristics of the hub we must remember that the hubs of the sounds made separately establish the reference points used in reading the sounds made in combination.

Different speakers have approximately the same hubs for the same sounds

The hubs of the same sounds, made by different speakers, are recognizable as the same because the similarities in the oral usage of English are greater than the differences. This is true even though no two people have identical voices, natives of various sections of the country have differences in pronunciation, and foreign accent is prevalent. The explanation of why the speech of different individuals is more alike than it is different is to be found in a combination of the following factors:

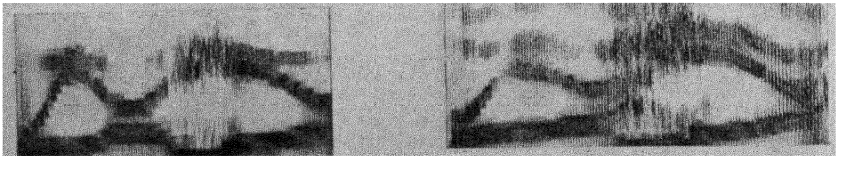
1. The general structural similarity of vocal tracts and articulators (anatomical).
2. The limitation of articulatory positions that will permit an easy and rapid transitional movement from one sound to another (physiological).
3. The general uniformity of the speech imitated by the individual during the learning process (psycho-sociological).
4. The modification of acquired speech habits to eliminate outstanding and handicapping differences from the general norm of understandability (psycho-sociological).

Since visible speech is a translation of audible speech, the similarities present in the audible speech of individuals are present also in their visible speech. Therefore, in the speech of the majority of speakers, the hubs of the same sounds are approximately the same.



We are here
(General American)

We are here
(English)



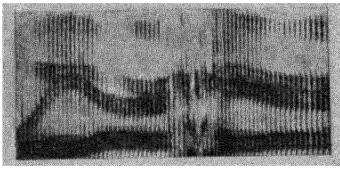
We are here
(Eastern)

We are here
(Southern)

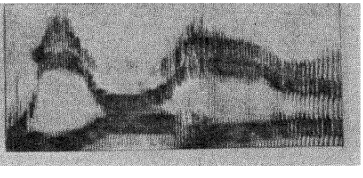
Look at the sentence, *We are here*, on the preceding page, as spoken by four individuals using General American, Eastern, Southern, and English speech. You can see some differences in the hubs, yet the four samples of speech are more alike than they are different. In visible speech, as in hearing, we must become accustomed to the individual's manner of speaking.

Differences in articulation do make differences in the hubs and patterns, but if a person can be understood when his speech is heard, his visible speech patterns can be read. Conversely, if a person's articulation is so unusual that he cannot be understood when he is heard, his speech patterns will be equally difficult to read.

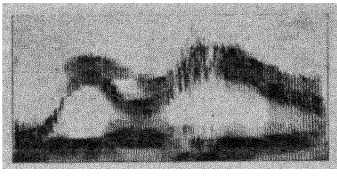
Although no two people have voices that are exactly alike, visible speech in the form in which we are reading it at present does not emphasize variations in voice pitch or quality. Even though some speakers produce patterns with four resonance bars and others have three-bar patterns, the relative positions of the bars, and therefore of the hubs, are recognizably similar. Voice quality may change the appearance of the resonance bars, but it does not tend to alter their relative positions or their hubs. For these reasons differences between male and female voices do not greatly change the patterns. It does not matter whether the speaker be a tenor, bass, soprano, or contralto. The relative hubs of the sounds are not changed enough by the pitch or the quality of voice to cause difficulty in reading.



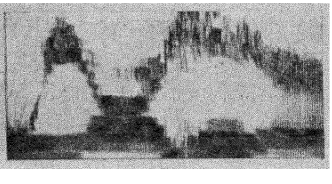
We are here
(***male low pitch***)



We are here
(***male high pitch***)



We are here
(***female low pitch***)

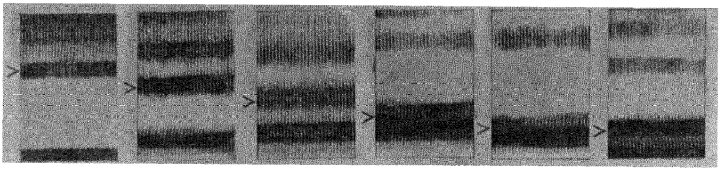


We are here
(***female high pitch***)

Look at the sentence, *We are here*, as spoken by a bass, a tenor, a contralto, and a soprano voice, and notice that changes in pitch and voice quality do not alter the hubs of the sounds very much.

Since similarities in articulation of English by different people are greater than the differences, and since voice pitch and quality do not greatly change the positions of the resonance bars, different speakers have approximately the same hubs for the same sounds.

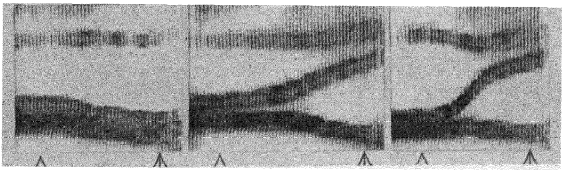
Each vowel has a different hub



i (*eve*) ɛ (*bet*) æ (*at*) ɑ (*father*) ɔ (*all*) u (*boot*)

An inspection of the above vowel patterns will show that the hub ≻ or bar 2 of each vowel is in a different position. Notice that the position of bar 2 is lowered for each vowel in the series starting with i (*eve*) and ending with u (*boot*). As bar 2 shifts its position with each vowel, bars 1 and 3 also change their positions, but not as much as bar 2 does. If you make the vowel sounds in the series above and feel the differences in the positions of the articulators, you can associate the way each of the vowels is made with its pattern. As the vocal resonances are changed by the movements of the articulators, the resonance bars change their positions on the patterns. Each vowel has a different pattern because it is made with different vocal resonances. Since the position of bar 2 is different for each of the vowels, each vowel has a different hub.

Each diphthong has two hubs

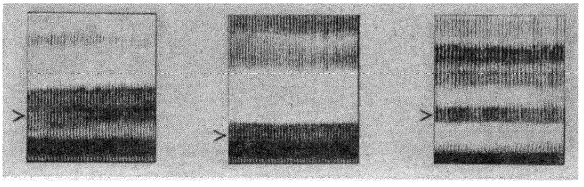


aʊ (*out*) aɪ (*I*) ɔɪ (*boy*)

You have seen the diphthong eɪ (*pay*), and know that a diphthong is a combination of two vowels which occurs frequently in words and syllables. In the diphthongs aʊ (*out*), aɪ (*I*), and ɔɪ (*boy*), notice the two vowels that make up each diphthong. The first vowel is shown at Λ, and the second vowel at Λ. Since a diphthong is a combination of two vowels, it has two hubs. The position of bar 2 of each vowel in the diphthong is the hub of that vowel. If you look at the patterns of the aɪ (*I*) and ɔɪ (*boy*),

you will see that the second vowel is the same in both diphthongs and that, therefore, their second hubs are the same. However, the first vowels are different and so are their hubs. The difference between the first and second hubs of ɔɪ (*boy*) is greater than the difference between the first and second hubs of aɪ (*I*), and that is why the transitional movement between the hubs is more pronounced for ɔɪ than for aɪ. Say the diphthongs aloud and try to associate the articulatory movements with the patterns of the diphthongs.

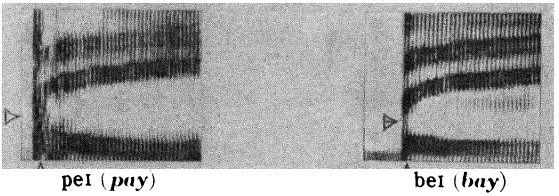
Each vowel-like sound has a different hub



r (*read*) l (*leave*) m (*me*)

Sounds like the r (*read*), l (*leave*), and m (*me*), are examples of a group of sounds which we have classified as vowel-like sounds because their patterns, like the vowels, contain resonance bars without fills. The total patterns of r (*read*), l (*leave*), and m (*me*) are each different since each vowel-like sound is made with a different combination of vocal resonances. Therefore, the resonance bars of each sound are in different positions, and the hubs of r (*read*) ≻, l (*leave*) ≻, and m (*me*) ≻ are different. Make the r, l, and m sounds and try to associate first, the vowel-like sound with the appearance of the resonance bars, and second, the positions of the articulators with the patterns of the sounds.

The voiceless and voiced sound of each pair of consonants have the same hub



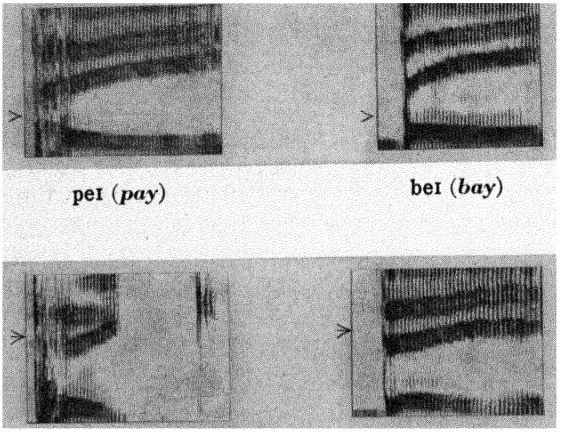
peɪ (*pay*) beɪ (*bay*)

The stop consonants p (*pay*) and b (*bay*) are made from substantially the same positions of the articulators. Therefore the hub of the p ⋗ is in the same position as the hub of the b ⋗. The pattern of the p ⋀ sound does not look like the b ⋀ pattern because,

1. The p sound is voiceless and makes a voiceless spike fill.
2. The b sound is voiced and makes a voiced spike fill.

Since the p and the b have the same hub, their influence on sounds with which they are combined is the same. Because the p is voiceless and the b is voiced, the transitional pattern between the p hub and the vowel is a little different from that between the b hub and the vowel. Notice, however, that the influence of the p on the eɪ bars follows the same direction as the influence of the b. Say the words *pay* and *bay* aloud, and try to associate the articulatory movements with the corresponding parts of the patterns.

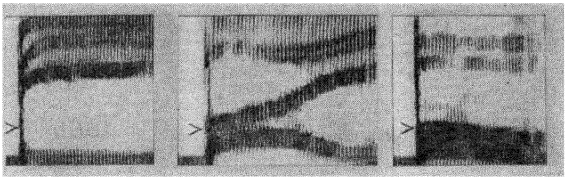
The hub of each pair of consonants is different



Compare the hub of the syllables p (*pay*)≻ and b (*bay*)≻ with the hub of t (*take*) ≽ and d (*day*) ≽. The p–b hub is lower on the pattern than the t–d hub, and therefore the influence of p–b on the eɪ sound is greater. The difference in the positions of the hubs of the two pairs of consonants is due to the differences in the way the sounds are made. Say *pay–bay* and then *tay–day* aloud and feel the differences in the movements of the articulators in saying the two pairs of words. Consonants that are articulated alike, as are the p and b, will have the same hubs. Consonants that are not articulated alike, such as the p and t, will have different hubs. Since each pair of consonants is articulated differently from every other pair of consonants, the hub of each pair of consonants is different.

The hubs of some sounds are relatively fixed

The hub of a sound is considered to be relatively fixed when the sound is articulated in a similar way regardless of its position in a syllable, or the vowel or consonant with which it is combined. Although it is impossible for an individual speaker to articulate a sound twice in exactly the same way, the variations in a sound made from a relatively fixed position are not aurally and visually significant. For example, fixed sounds like the **p** (*pay*), **b** (*bay*), and **m** (*may*) do not vary much aurally or visually because they are made with a closed lip position which cannot be modified sufficiently to change the sound patterns significantly. A fixed sound is comparatively easy to recognize and to read because its patterns are consistently similar. In addition, the fixed sounds influence other sounds more and are influenced by other sounds less than sounds made from more variable positions.



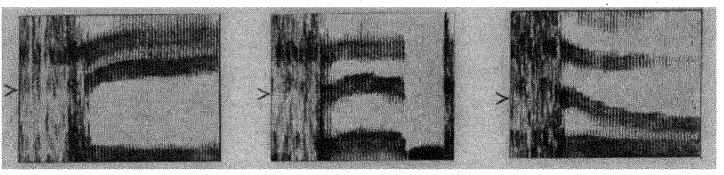
bi (*bee*) **baɪ** (*buy*) **boʊ** (*bow*)

The hub of the **b** > in combination with the **i** (*bee*), **aɪ** (*buy*), and **oʊ** (*bow*) is fixed. This is shown by its influence on the different vowels. Because the distance between the **b** and the **i** hubs is greater than the distance between the hubs of the **b** and **oʊ**, the influence of the **b** is greater on **i** than on **oʊ**.

Although the hubs of some sounds shift slightly when the sounds are combined with other consonants or with vowels, they *do not change sufficiently* to cause an easily recognizable difference in either the position of the hub or the influence of the sound on other sounds. An individual speaker may make one of these sounds from several slightly different articulatory positions without producing a significant change in either the acoustic or visible patterns of the sound. Therefore, such sounds as the **s** and **z**, the **t** and **d**, and the **ʃ** and **ʒ**, have been classified as sounds with *relatively fixed* hubs until such time as more detailed information concerning the extent of their variations enables us to construct a more precise classification.

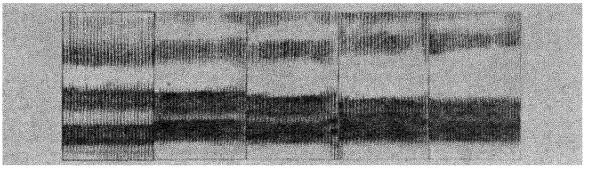
Notice the very slight shift of the **s** hub >, in the figure at the top of the facing page, from a high hub position with the **i** (*see*), to a lower hub position with the **u** (*sue*). The change in the **s** hub and pattern results from the shift of the articulators towards the position of the vowel or consonant with which the **s** is combined.

The vowels, also, are considered to be relatively fixed sounds because, under normal conditions of articulation, a specific vowel sound may be repeated only from the same



si (*see*) sɛd (*said*) su (*sue*)

or a similar combination of articulator positions. Theoretically it is impossible to produce a vowel twice in exactly the same way, but for practical purposes we recognize a vowel as i (*eve*) as long as it is more like the i than another vowel. Therefore any individual vowel is believed to be produced from a relatively fixed position which can be varied only to the extent that it does not produce another vowel. The exact point of departure from one vowel to another is not as important in reading visible speech as it is in phonetic analyses. Any number of sounds may be produced between two recognized or standard vowels in the vowel series.



æ (*at*) ɑ (*father*)

The illustration above shows three of the vowel sounds which it is possible to produce between the standard vowels, æ as in *at*, and ɑ as in *father*. For speech reading it is sufficient to read the variations of the vowels approximately as they appear.

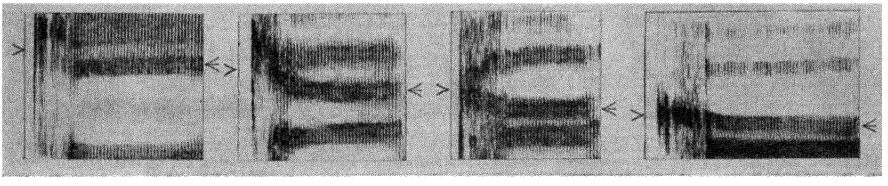
The hubs of some sounds are variable

The variable sounds are sounds that are generally produced from a slightly different articulatory position for each vowel or consonant with which they are combined. These changes in articulatory position are towards the vowel or consonant with which the sound is combined, and they result in a sound which differs noticeably in both its acoustic and visible patterns with each combination. An individual speaker may produce a larger or smaller number of these variations of a sound depending usually upon the section of the country from which he comes and upon the speed of his articulation.

A variable sound may be produced from one articulatory position when combined with sounds made in the front of the mouth, from another position with sounds made

in the back of the mouth, and from still another position with sounds made in the middle of the mouth. Or there may be a series of positions used to produce the sound when it initiates and another series when it terminates a syllable. Therefore a variable sound does not have a single hub but a range of hubs. In the same way that we have learned to hear and to accept the variations of a sound as though it were one sound, we learn, in visible speech, to read the visible variations of a sound as one sound with one meaning.

The way in which the hub of a variable consonant shifts to meet the hubs of the sounds with which it is combined is illustrated in the four **k**-*vowel* syllables.



ki (***key***) **kæ** (***cat***) **kʌ** (***cup***) **ku** (***cool***)

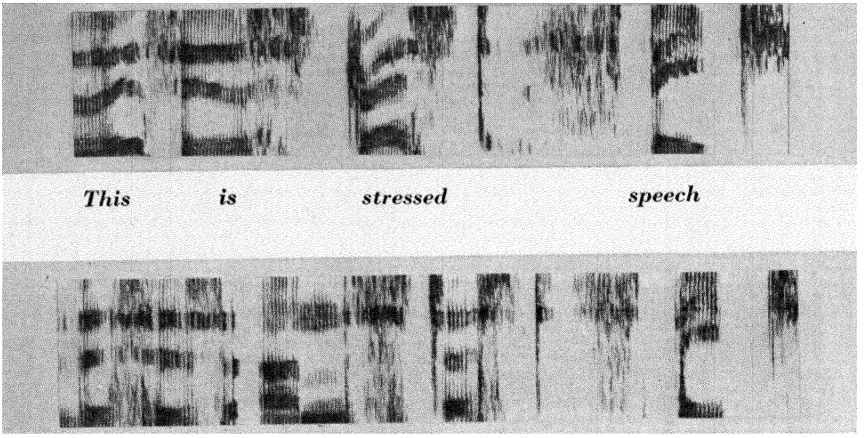
Notice that the hub of the **k** > moves with the hub of the vowel < from the top of the pattern in **ki** (*key*) towards the bottom of the pattern in **ku** (*cool*). Say the syllables aloud and try to feel the surface of the tongue touch the roof of the mouth a little farther back for the **k** with each syllable from **ki** to **ku**. We shall study the variable sounds in greater detail in the subsequent lesson units.

How Stressed and Unstressed Sounds Look

Stress is used to give emphasis to particular syllables or words. It results from changes in loudness, in pitch, and in the amount of time given to the syllable or word. The use of stress is one of the most significant differences between speaking and writing, since a speaker is enabled to say a word or sentence in many different ways with as many different meanings. It is difficult to illustrate the results of stressing or unstressing without an assumption of a certain amount of familiarity with the sound patterns. For that reason, the detailed discussion of stressing has been included in the lesson units that follow, and we present here only the general principles underlying the visible speech representations of stressing and unstressing.

In the visible speech patterns made on the spectrograph, stress is shown by:

1. The length of the picture, which is the duration or the time taken to say the syllable or word.
2. The shade of darkness of the patterns, which is related to the volume or loudness.
3. The absence of part of the pattern from the picture, which is the part of the sound too weak to be recorded.
4. Shift in the patterns from their stressed appearance due to variations of the articulators from the positions used in stressing.



Compare the pattern of "*This is stressed speech*" with the pattern of "*This is unstressed speech.*" Notice that:

1. The length of the individual sounds and words is greater in the stressed than in the unstressed sentence.
2. The patterns are darker in the stressed than in the unstressed sentence.
3. Every sound is fully shown in the stressed sentence. In the unstressed sentence parts of some of the sounds are missing.
4. In the stressed sentence, the positions of the vowel bars reach the place on the picture that they have when the sounds are made alone. Can you see that some of the vowel bars in the unstressed sentence do not reach the positions they have in the stressed sentence?

The positions of the vowel bars in the unstressed sentence show that the articulators did not reach the positions necessary to make the vowel as it is when it is stressed or made alone. The reason is that the articulators did not complete the movement to the usual vowel position. It takes time for the articulators to move from one position to another. If a short cut is taken to save time or effort in saying a word, some of the sounds in the word may be either unstressed, partly omitted, or entirely omitted.

When a vowel is unstressed in different degrees, it may show:

1. Lighter patterns with the vowel bars in their stressed positions.
2. Changes of the vowel bars from their stressed positions. When this happens, another vowel is made. The vowel made depends upon the shape and size of the resonating cavities. When vowels are completely unstressed, they sound and look like the unstressed form of the neutral vowel ʌ (*up*). Both the stressed and the unstressed ʌ (*up*) are made from a neutral position of the articulators; that is, the articulators are as relaxed as possible. The

pattern of the first sound in the word *unstressed* shows the position of the bars for the stressed ʌ (*up*) sound. When vowels are completely unstressed and resemble the unstressed ʌ (*up*), they are read as the unstressed vowel ə, the first sound in the word *about*.

3. Vowel resonance bars 3 and 4 missing.
4. A decrease in length of the resonance bars.

When a consonant is unstressed in different degrees, it may show:

1. A lighter picture of its usual pattern.
2. Some of the pattern missing. When this occurs, only the darker parts of the usual pattern are shown.
3. No visible pattern.

Good speech requires both stressing and unstressing of syllables and words. When every syllable of every word is stressed, the speech is just as abnormal as when none of the syllables is stressed. The meaning that the speaker wishes to give his words largely determines the kind and the amount of stress used. In reading visible speech, although we must be able to get the basic meaning from the patterns, we have not been concerned as yet with the fine shades of meaning that may be given by stressing and unstressing.

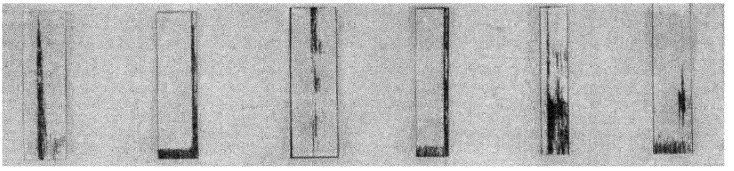
The New ABC's

In this section patterns are shown for the sounds used in speaking the American language. These patterns, in effect, constitute a new set of ABC's. Although in reading visible speech we do not read the individual sound patterns any more than we read the letters of the alphabet on this printed page, we do depend upon recognizing these ABC's in reading the combinations of sound patterns that make syllables, words, and groups of words. The ABC's are shown here with Groups 1 and 2 — the voiced and voiceless stop sounds — combined, and Groups 3 and 4 — the voiceless and voiced fricative sounds — combined.

Many sounds cannot be said naturally or easily unless they are spoken in combination with another sound. In making the patterns presented below in the ABC's of Visible Speech, all sounds except the vowels and diphthongs were spoken before the neutral vowel ʌ (*up*). The so-called neutral vowel was selected because in its production the articulators are in a relatively neutral or at-rest position, and, therefore, the vowel position does not affect sounds combined with it as much as do the other vowels. The key words which accompany the phonetic symbols in the ABC's are used also in the classification of sounds, in the lesson chapters that follow, and in the appendix. Only one pattern is shown here for variable sounds like the *h*, *l*, and *k* that can be made from several positions. Other patterns of the variable sounds are presented in the following chapters. Do not memorize these patterns, but use them as references in studying the chapters on how to read.

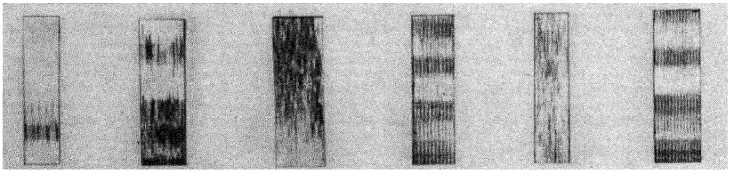
The ABC's of Visible Speech

Groups 1 and 2 — Voiceless and Voiced Stop Sounds

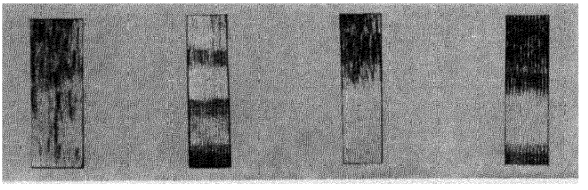


p (*pay*) b (*be*) t (*to*) d (*day*) k (*key*) g (*go*)

Groups 3 and 4 — Voiceless and Voiced Fricative Sounds

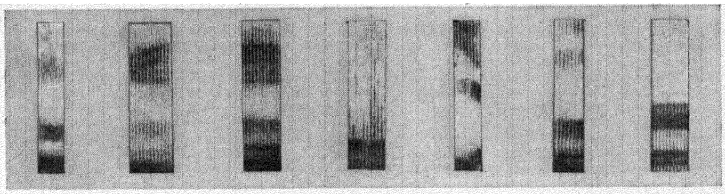


h (*he*) h (*ahead*) r (*for*) v (*vote*) θ (*thin*) ð (*then*)

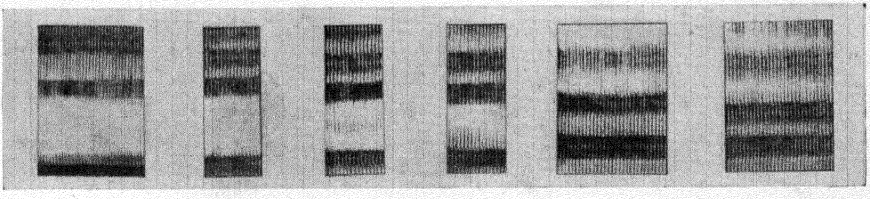


s (*see*) z (*zoo*) ʃ (*she*) ʒ (*azure*)

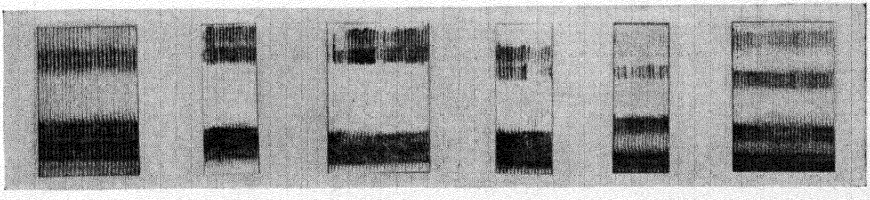
Group 5 — Vowels and Vowel-like Sounds



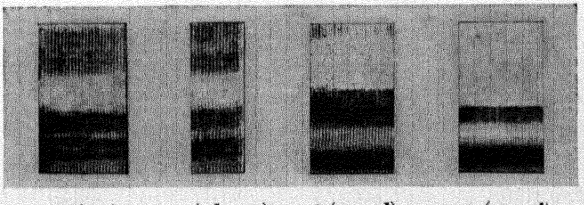
m (*me*) n (*no*) ŋ (*sing*) w (*we*) j (*you*) l (*let*) r (*read*)



i (*eve*) ɪ (*it*) e (*hate*) ɛ (*met*) æ (*at*) a (*ask*)

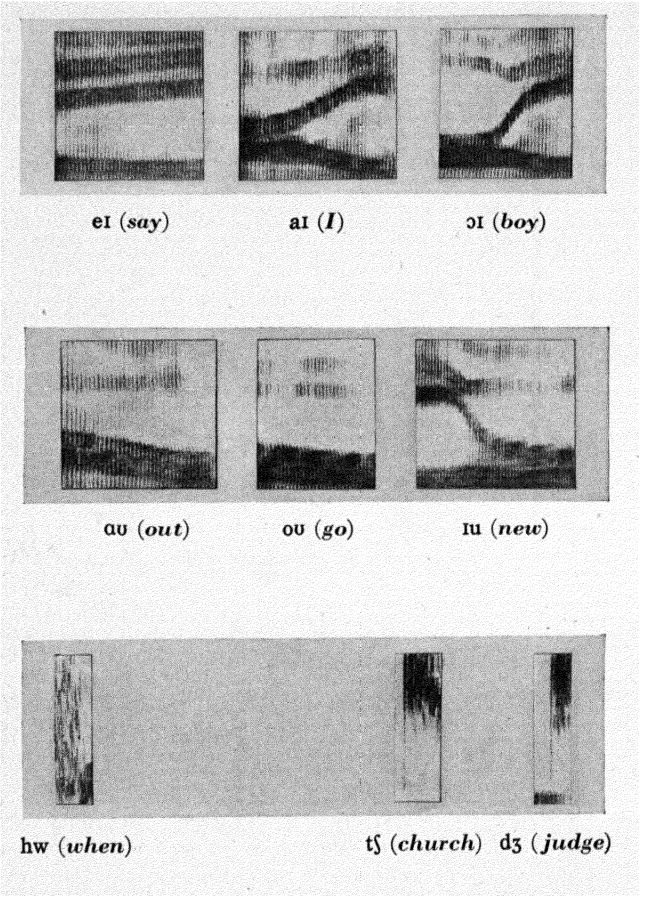


ɑ (*father*) ɒ (*not*) ɔ (*all*) o (*obey*) ʊ (*foot*) u (*boot*)



ʌ (*up*) ə (*about*) ɚ (*word*) ɜ (*word*)

Group 6 — Combination Sounds



As mentioned earlier, the above sound units represent the spoken American language, but visible phonetic alphabets can be developed for other languages. The basic principles and terminology used in describing the American sounds would apply also to the patterns obtained for other languages. Although spelling and pronunciation are at variance in many languages, we have, in visible speech, a natural phonetic alphabet in which sound patterns are related directly both to the manner in which the sounds are produced and to the aural impression.

PART TWO

LESSON UNITS

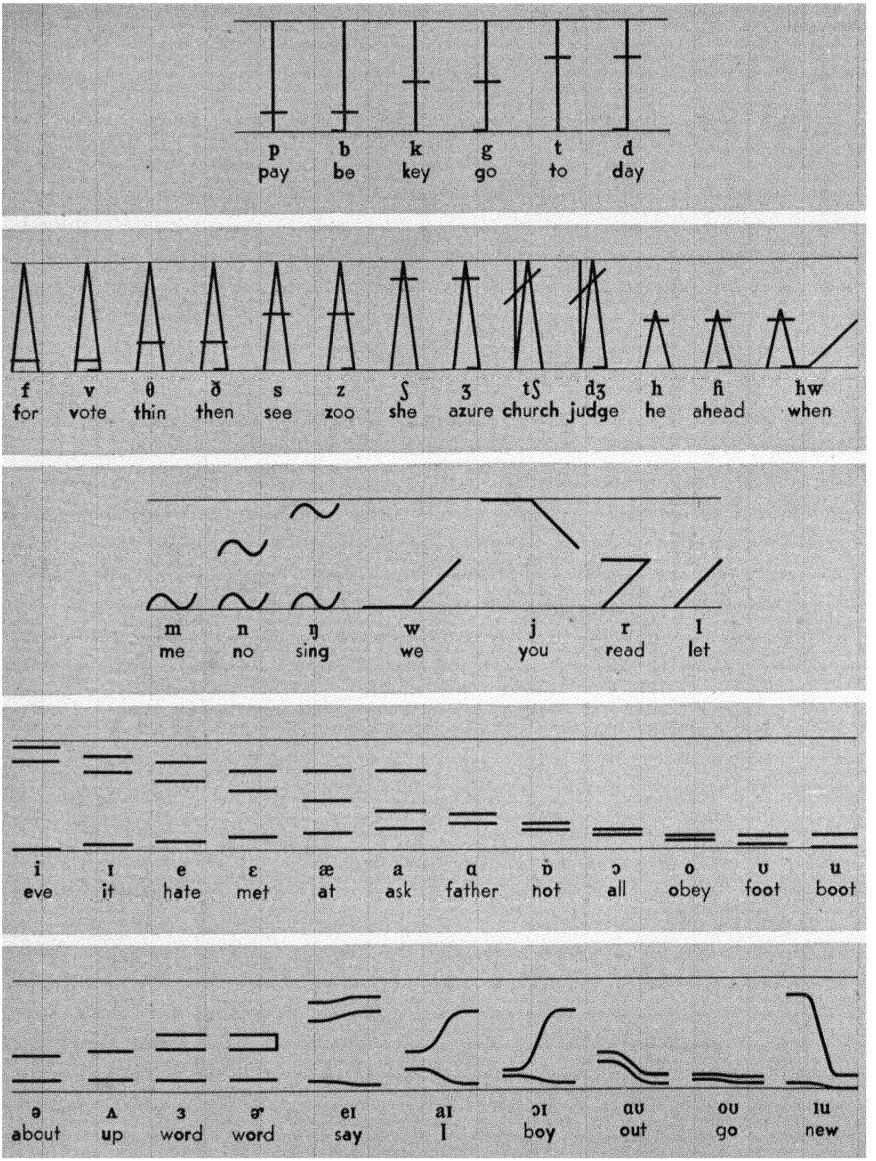
VISIBLE PATTERNS OF AMERICAN SPEECH

Chapter 6

Suggestions for Using the Lesson Chapters

The chapters immediately following this one are fundamentally lessons in reading visible speech. The lesson chapters have been organized and written for the adult reader. The materials are presented primarily for the adult with hearing or with a relatively recent hearing loss, so that the auditory impressions of sounds and words may be recalled. This does not mean that the deaf or those with a long history of hearing loss will be unable to learn about visible speech from these pages. It simply means that some of the explanations may not apply to the reader without experience in correlating auditory impressions with speech movements. For the reader who is interested only in a cursory acquaintance with the subject-content, the material will serve as additional illustrations of the basic principles which underlie the readability of visible speech. For those readers who wish to examine the patterns more critically or to learn to read visible speech, the material presented in the lesson chapters should be adequate. However, it must be remembered that a textbook alone cannot enable one to read visible speech fluently. Such proficiency requires practice in reading from a direct translator under the varying conditions of conversation, in order to acquire a familiarity with the individual variations that appear in speech. Readers interested in teaching speech to deaf children, improving and correcting speech, teaching phonetics or foreign languages, will find that the lessons are not written to meet their specific needs. However, it will be evident that much of the material can be adapted to these various fields. Aside from these qualifications, it is hoped that the content of the lesson chapters will satisfy any degree of interest ranging from casual curiosity to professional interest in specific details.

A reader who desires to learn the visible speech patterns should start with the basic principles discussed in Chapter 5. Then he may proceed at his own speed through the lesson chapters. For the most efficient results, he is advised to study the chapters in the order in which they are written. Except for Chapter 7, which presents reference patterns of the vowels as a basis for learning the words that contain the stop and fricative consonants discussed in the subsequent chapters, each chapter deals with a group of sounds as outlined in the classification of sounds, and is subdivided into lesson units dealing with the various sounds within that group. Every lesson unit introduces a few of the points essential to reading visible speech. The lessons are arranged in progressive order, and therefore it is desirable that the principles as well as the words and sentences in each unit be understood thoroughly before proceeding to the next. Familiarity with the principles and skills studied in one unit forms the



Manual Symbols for Visible Speech Adapted for Use in the Experimental Visible Speech Training Program

basis for the understanding of those that follow. Although the lesson units are all of about the same length, those at the beginning contain comparatively easy material, and the reader is acquainted gradually with the more difficult principles and patterns of visible speech.

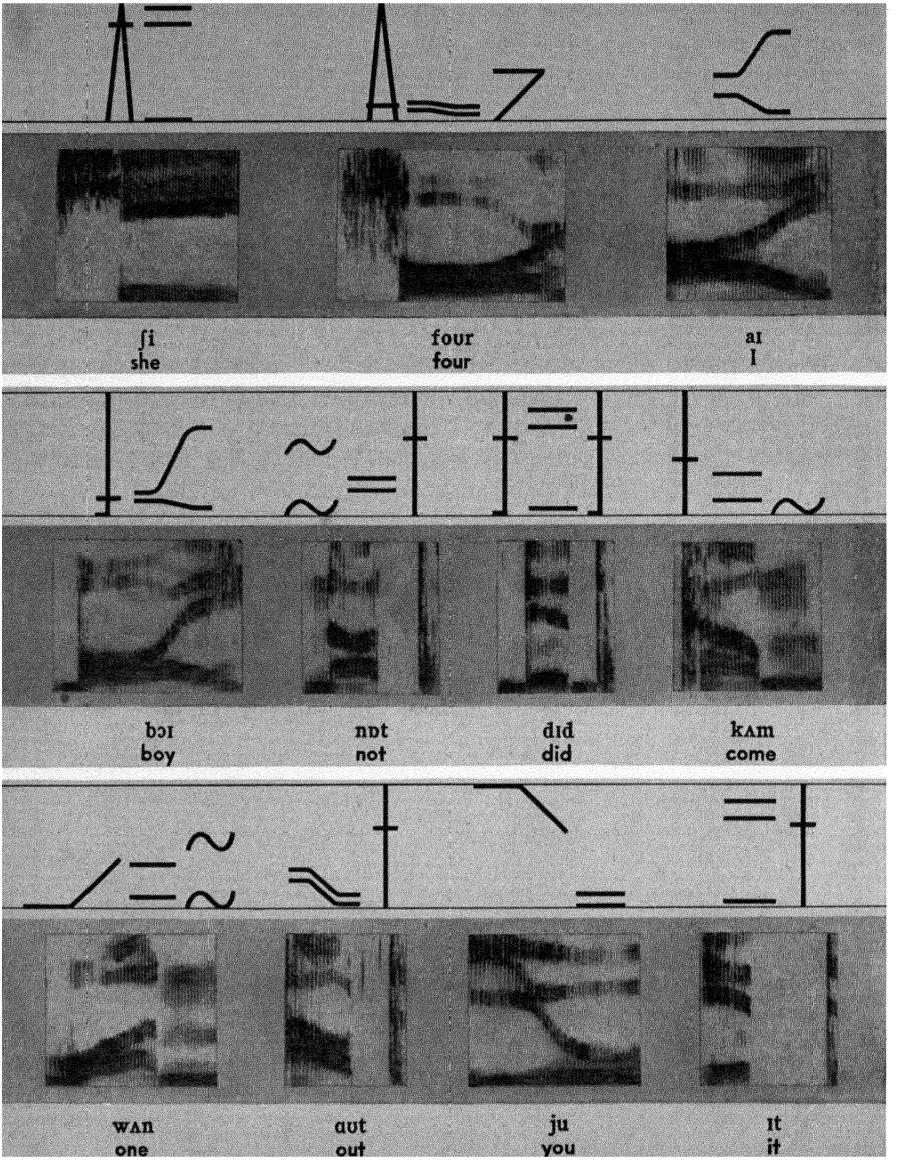
All of the lesson units are developed from the same general outline. The first part of each unit shows a profile diagram of the positions of the articulators used in making the sounds discussed in that unit. These diagrams and the exercises that are focused on them should be studied carefully because they will help to develop the awareness or the "feel" of how sounds are made. Since sounds look the way they are made, an awareness of the movements of the articulators in producing the sounds will enable the reader to associate the visible patterns with the articulatory movements. With practice, the reader will be able to tell how different speech patterns were spoken, and also to visualize how other patterns will look before he actually sees them.

When a sound has more than one spelling, the different spellings are listed so that the reader may be reminded of the differences between the spoken and written language in this respect. Since visible speech portrays the sound without relation to variations in spelling, an appreciation of the numerous inconsistencies in our written language will help to eliminate confusion.

From the very first lesson unit, the reader may begin to acquire a visible speech reading vocabulary which is enlarged with each successive study unit. The words used in these lessons have been selected largely from the first 1000 most frequently used words.* In learning visible speech, this limited vocabulary is considered only as a framework for a practical conversational vocabulary. The reader who desires to know the patterns thoroughly should apply his knowledge of visible speech principles to the visualization of the additional words presented for study in each unit, and to familiar words and phrases outside the text vocabulary. In addition, a visual recall of vocabulary previously learned, and a visualization of new words and sentences, provide an excellent means for fixing the patterns in memory.

In reviewing the sentences provided for drill, accuracy should not be sacrificed for speed in reading, and the reader is urged to guard against a natural tendency to overlook unfamiliar material. However, there is some detail in the drill words, phrases, and sentences that will be of interest only after the reader is familiar with basic speech variations. The patterns used in the lessons are the translations of the speech of the authors, of some members of the experimental training group, and of other individuals associated with the development. They contain qualities that are the result of individual variation, just as the handwriting of these same people contains characteristic shapes that convey evidence of individuality. In handwriting as well as in audible speech, these quality features may tend to reduce the general intelligibility. A hand-

*Selected from the sources given in the footnote on page 26.



Visible Speech Patterns and Manual Symbols of Words
Illustrations of the Use of Manual Symbols

writing which we call "beautiful" or "forceful" may be difficult to read, and the same may be true of speech either as it is heard or as it is seen in visible speech form. Therefore, if we undertake to learn to read visible speech, we must disregard, initially at least, all but the information-bearing clues.

These information-bearing clues are the resonance bars, the fills, the gaps, the spikes, and the baseline, which were defined in the preceding chapter. Each unit presents the unique clues by which the sounds and the combination of sounds discussed in that unit are recognized, and the reader is urged to understand how the clues are associated with the speech processes. In order to transfer reading habits to reading visible speech at a normal rate of speaking from a direct translator, it is important that only the essential clues be learned since there is not time to look at minor details.

Sounds are shown in words and words are used in sentences beginning with the first unit. The reader should always read the drill phrases and sentences from left to right in the same way that the eyes follow the lines on this page. This is a most important habit to establish, for it is the basis of the eye movements developed in reading visible speech on film and on the direct translators.

The fixation span and the duration of the fixation must also be developed with the direct translators, but the habits formed in reading the static patterns in the text may influence later reading proficiency. At present the problem of habit formation is actually one of transferring established reading habits to reading visible speech patterns. However, an incorrect transfer of habits is as easily made as a correct transfer of habits, and for this reason it seems essential that the principles involved in reading be understood. If the student has no knowledge of the problems involved in reading, it is not probable that maximum reading speed and accuracy will be attained without some supervision.

Another suggestion to aid the reader to learn visible speech patterns, is the practice of writing the drill words, phrases, and sentences in phonetic symbols that resemble the visible speech patterns. The manual symbols shown on page 60 have been adapted from the clues found essential to the recognition of the sound unit patterns. In order to simplify the manual notation, transitional influences have been omitted, and an attempt has been made to give each of the sound pattern groups a characteristic pattern. It may be noted that stop sounds are represented by vertical strokes, fricative sounds by two slanting strokes joined at the top, vowels by horizontal strokes, nasal sounds by wavy horizontal strokes, and the other vowel-like sounds by combinations of horizontal and oblique strokes. Voicing is represented by a line along the baseline. The hubs of the stop sounds and fricative sounds are represented in an approximate fashion by the placing of horizontal strokes. These symbols have been found by members of the training group to be helpful in learning sound pattern clues, in describing and explaining the patterns, and as a shorthand representation of patterns which

are to be visualized in greater detail for study. Even though some of the symbols have been modified repeatedly for greater ease and accuracy in transcription, much remains to be done to improve them. The symbols used by the training group are shown on page 60 for the consideration of those who may wish to use them. In addition to the discrete manual symbols, it often has been found desirable to draw actual diagrammatic representations of the visible speech patterns to clarify some of the more difficult points of transitional influence.

The symbolizations shown on page 62 present spectrograms of several words selected to illustrate the use of the manually written symbols and to enable the reader to compare the manual symbols with the visible speech patterns. As mentioned previously, the manual symbols are designed to show only some of the essential clues used in recognizing and reading the visible speech patterns of individual sounds. Undoubtedly, additional notations would provide a more precise and accurate manual symbolization of visible speech. But because of the experimental stage of the development, it seems desirable to postpone the refinement of the manual symbols until further improvements of the translators have been made. In the meantime, the use of some system of manual notation, such as we have presented above, will be helpful in learning visible speech.

Since the emphasis in this volume is not on teaching methods, but rather on a presentation of facts and fundamental concepts, the reader should feel free to relate the materials to his own field of interest. The authors believe that the variety of these interests makes it inadvisable to present arbitrary procedures to be followed in conjunction with the lesson chapters. It is hoped that the facts and principles of visible speech are presented in a manner which will permit the reader to draw his own conclusions as to their application to his field, and to consider the possible learning or teaching methods which might be applied most efficiently.

Chapter 7

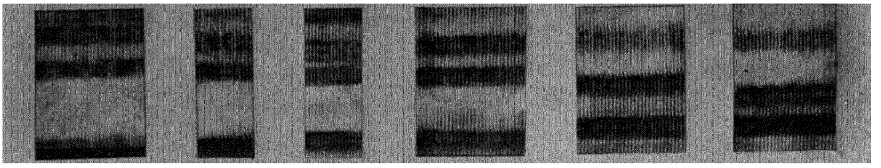
The Patterns and the Hubs of the Vowels and Diphthongs

Before beginning the specific study of the consonants, a general knowledge of the patterns and hubs of the vowels is needed. As stated previously, the pattern of a vowel made separately gives us a reference pattern by which we can recognize that vowel when it is combined with another sound. It is also true that we frequently read the consonant which is combined with the vowel by the way the vowel is changed during the transitional interval between the two. As a matter of fact, the clues presented by the way a consonant changes the shape and position of the resonance bars of a vowel are as helpful as are the characteristic patterns of the individual sounds. Unless you are familiar with the appearance of the vowel pattern when it is uninfluenced by the consonant, you cannot appreciate how it is changed in combinations. Therefore, in this chapter we hope to establish certain of the visual reference patterns used in reading visible speech. In addition, a few words are introduced to serve as a beginning visible speech vocabulary which will be enlarged in each succeeding unit. After the consonants have been explained, the vowels and diphthongs will be discussed in more detail, including diagrams of the articulatory positions and movements. Several of the final lesson units have been reserved for that purpose.

For convenience, the vowels with which we are concerned in this chapter have been classified into the three groups usually used in phonetic descriptions. The groups, which are referred to as the front, back, and mid vowels, attempt to take some account of the positions of the tongue employed in their formation.

The individual vowels result from the modulation of the voiced breath stream by different sizes and shapes of the vocal cavities. The main cavities are formed in the mouth, throat, and nose. The nasal cavity is usually completely or partially closed off from the pharyngeal or throat cavity when we produce the vowel sounds. All of the vowels are voiced sounds, and in the 3500-cycle patterns used in this book, they usually show from two to four resonance bars. The number and positions of the resonance bars in the pattern of each vowel are related to the characteristic sizes and shapes of the vocal cavities used to produce the vowel. In learning to read the vowels, it is helpful to associate the articulatory positions with the numbers and positions of the resonance bars.

Each Front Vowel Has a Distinct Pattern



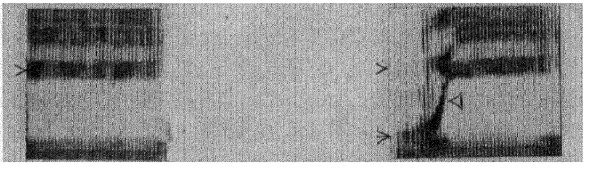
i (*eve*) **ɪ** (*it*) **e** (*hate*) **ɛ** (*met*) **æ** (*at*) **a** (*ask*)

These are the patterns of the front vowels as they appear when each vowel is said separately. Notice that no two vowel patterns are alike. The most important difference is in the positions of the resonance bars for each vowel. Bar 1 is on the baseline for **i** (*eve*). For each of the following vowels in the series, **ɪ** (*it*), **e** (*hate*), **ɛ** (*met*), **æ** (*at*), and **a** (*ask*), bar 1 is raised increasingly farther from the baseline. In **a** (*ask*), bar 1 is farther off the baseline than it is for any other front vowel.

Bar 2 in **i** (*eve*) is higher than it is for any other vowel. Observe that bar 2 is lowered progressively for each of the vowels in the series. In **a** (*ask*), bar 2 is closer to bar 1 than it is to bar 3. Bar 3 for **i** (*eve*) is just above bar 2, and it follows bar 2 in the series except that it is not lowered as much.

Say the vowel sounds **i** (*eve*), **ɪ** (*it*), **e** (*hate*), **ɛ** (*met*), **æ** (*at*), and **a** (*ask*), aloud. Feel how the tongue takes a high front position for **i** (*eve*) and how it is lowered a little for each of the other vowels when they are said in series. The teeth are almost together for the **i** (*eve*), but are farther apart for each of the vowels made after **i**. As the tongue and jaw are lowered, in making the vowels in series from **i** (*eve*) to **a** (*ask*), the position of bar 1 becomes higher and the positions of bars 2 and 3 are lowered.

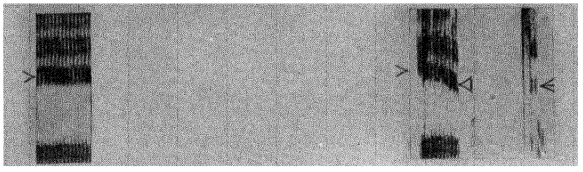
The main hub of each of the front vowels is the position of bar 2



i (*we*) ***we***

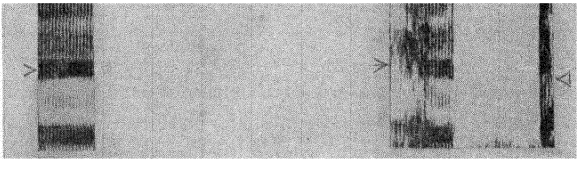
The main hub of the **i** (*we*) is the position of its second bar >, which is close to bar 3 at the top of the pattern. Bar 1 of the **i** is always flat on the baseline. When the **i** is said with another sound its pattern is altered by the transitional influence as

discussed previously in Chapter 5. In the word *we*, the hub of the **w** is in the **u** (*boot*) position ≻ near the bottom of the picture. Since the **w** is made from a moving position of the articulators, the **w** glide is shown by the pattern ≺ that connects the **u** (*boot*) hub and the hub of **i** (*we*).



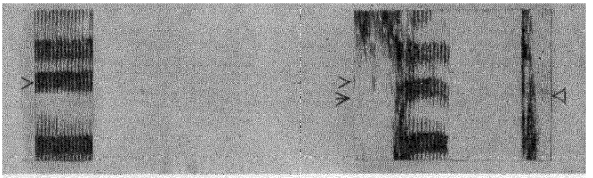
ɪ (*it*) ***it***

The main hub of the **ɪ** (*it*) is the position of its second bar ≻. In this vowel, bars 2 and 3 are further from the top of the pattern than they are in **i** (*eve*), and bar 1 is raised off the baseline a little. The **ɪ** (*it*) usually is said as a short vowel and therefore its pattern is not as long as the pattern for the **i** (*eve*). It is interesting to note that the hub of **t** in *it* ≺ is lower than the hub of the **ɪ** sound. Because the **t** follows the vowel, the **ɪ** is cut off sharply by the articulators as they stop the breath stream to make the **t**. As the tongue moves from the **ɪ** position to the **t** position to stop the breath, bar 2 of **ɪ** curves downwards ≺.



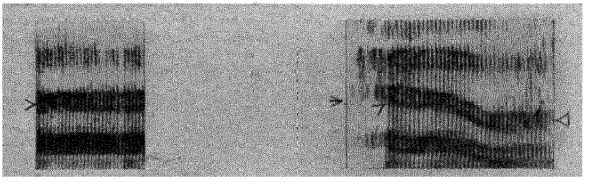
e (*hate*) ***hate***

The main **e** (*hate*) hub is the position of its second bar ≻. Bar 2 of the **e** is lower than it was for the **ɪ** (*it*) and so is bar 3. The space between bar 1 and the baseline is more evident in **e** than in **ɪ** (*it*). The hub of the **h** in *hate* ≻ is in almost the same position as the hub of the vowel and therefore does not influence the vowel pattern. The **t** hub ≺ is a little lower than the **e** hub, and the movement of the articulators from the **e** position to the **t** position curves bar 2 very slightly and also cuts off the vowel pattern sharply. The vowel **e** is used in only a few words by American speakers, and even in these words the diphthong **eɪ** (*take*) or the vowel **ɛ** (*met*) frequently is substituted. However, the vowel **e** is used by most speakers as the first element of the diphthong **eɪ**.



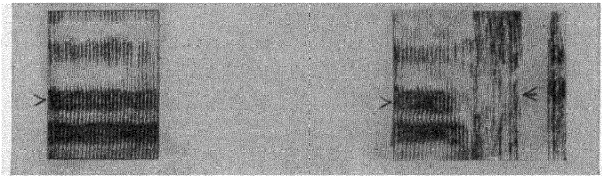
ɛ (*set*) *set*

The main hub for the ɛ (*set*) is bar 2 ≻ which is just a little above the center of the pattern. Bar 3 has moved down a little from the top of the pattern and is lower than it was for ɪ (*it*) or e (*hate*). Bar 1 is further off the baseline than it was for ɪ or e, and there is considerably less space between bars 1 and 2. In the word *set*, the hub of the s ≻ is a little lower than the ɛ hub, and the hub of t ◁ is also lower than the bar 2 position of ɛ. When *set* is said, the beginning of bar 2 curves upwards from the s hub. As the tongue moves from the ɛ position to the t position to stop the breath, bar 2 curves downwards.



æ (*have*) *have*

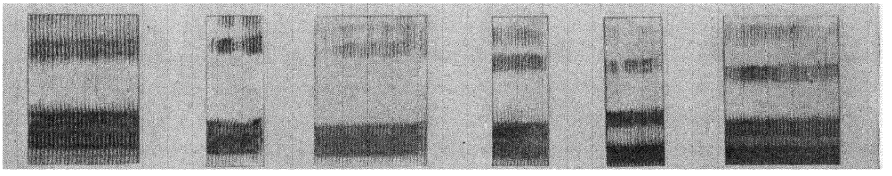
The main æ (*have*) hub is the position of the second bar ≻. In æ, bars 1, 2, and 3 are almost equally spaced on the pattern. It is worth remembering this spacing since it affords an easily recognized reference. Both bars 1 and 3 are nearer to the center of the pattern than they were in the front vowel patterns which precede æ (*have*). The hub of h in *have* ≻ is almost the same as the hub of the vowel, but the hub of v ◁ is lower than the æ hub. As *have* is said, the positions of the æ bars are not changed very much by the h, but bar 2 curves downwards towards the hub of v.



ɑ (*ask*) *ask*

Bar 2, the main hub of **a** (*ask*) > is a little below the center of the pattern. Bar 3 remains approximately in the same position it had for the æ (*have*), but bar 2 is nearer to bar 1 than it is in any of the other front vowels. In *ask*, the **s** hub < is almost in the same position as the main **a** hub >, so the end of bar 2 of **a** is not changed very much by the **s** hub. Many speakers substitute the ɑ (*father*) or the æ (*at*) for this vowel in such words as *last*, *past*, and *path*.

Each Back Vowel Has a Distinct Pattern



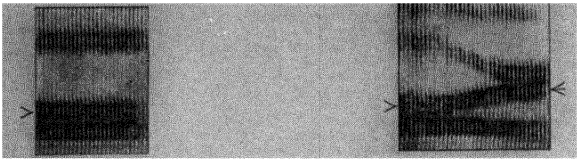
ɑ (*father*) ɒ (*not*) ɔ (*all*) o (*obey*) ʊ (*put*) u (*boot*)

These are the patterns of the back vowels as they appear when each vowel is said separately. Notice that each vowel pattern is a little different from the patterns of the other vowels. The main difference among the back vowels is in the amount of space between the baseline and the first bar. Bar 1 in ɑ (*father*) is farther from the baseline than it is in any other vowel. For each of the following vowels in the series ɒ (*not*), ɔ (*all*), o (*obey*), ʊ (*put*), and u (*boot*), bar 1 is lowered progressively. In u (*boot*), bar 1 is on the baseline.

Bar 2 in ɑ (*father*) is just above bar 1, and usually the two bars appear to be joined. In the other vowels of this series, bar 1 is lowered progressively. Bar 2 also is lowered in the vowels between ɑ (*father*) and o (*obey*). There is often a small space between bars 1 and 2, in the ʊ (*put*) and the u (*boot*). Because the third and fourth bars of the back vowels are weak, they frequently are not present, and are of little help in reading these vowels.

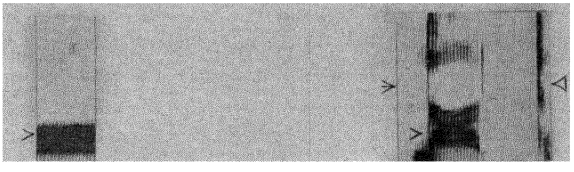
Say the vowel sounds ɑ (*father*), ɒ (*not*), ɔ (*all*), o (*obey*), ʊ (*put*), and u (*boot*), aloud. Try to associate the positions of your articulators with the patterns of the sounds. The tongue is almost flat for the ɑ (*father*), and the back part of the tongue is raised progressively for each of the vowels as you make them one after another. The teeth are farther apart for the ɑ (*father*), than for any other vowel, and they are brought a little closer together for each of the vowels from the ɑ (*father*) to the u (*boot*). Notice that as the back part of the tongue and the lower jaw are raised in making the back vowels in series from the ɑ (*father*), to the u (*boot*), the positions of bar 1 become progressively lower.

The main hub of each of the back vowels is the position of bar 2



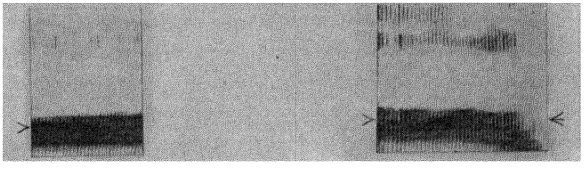
ɑ (*are*) *are*

The main hub of the ɑ (*are*) ≻ is below the center of the pattern. In ɑ, for the first time, bar 2 is just above bar 1. Bar 1 is higher than it is for the front vowels and the other back vowels, and bar 3 is in approximately the same position that it has for the æ (*have*). The ɑ hub ≻ is lower than the r hub ≺; therefore, in *are*, the end of bar 2 of ɑ curves upwards towards the r hub. The continued voicing during the movements of the articulators from the ɑ position to the r position makes the curve of bar 2 from one hub to the other.



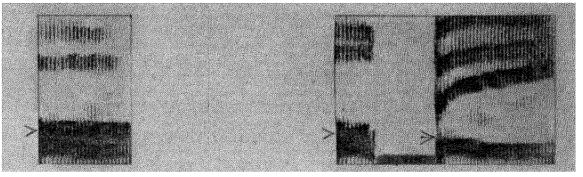
ɒ (*not*) *not*

The main hub of the ɒ (*not*) ≻ is a little lower than the hub of ɑ (*are*). Bars 1 and 2 are so close together that they appear to be joined, and both bars are nearer the base-line than they are for the ɑ (*are*). Frequently a short form of the ɑ (*are*) is substituted for the ɒ (*not*). In the word *not*, the n hub ≽ is higher than the main ɒ hub ≻. The t hub ≺ is in about the same position as the n hub. As *not* is said, the beginning of bar 2 of the vowel curves downwards from the n hub, and then the end of bar 2 curves upwards towards the t hub. The curving of the beginning and ending of bar 2 of the vowel is due to the movements of the articulators as you say one sound after another.



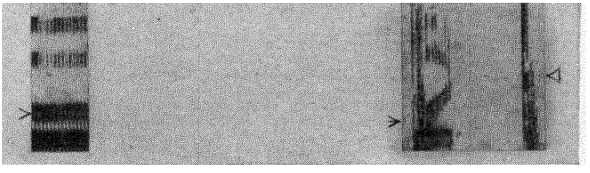
ɔ (*all*) *all*

The main hub of the ɔ (*all*) ≻ is lower than the hubs of the preceding back vowels; and, since bars 1 and 2 are close together, both bars are lowered. When bar 3 is present, it is in approximately the position it has for the æ (*have*). In the word *all*, the hub of 1 ≺ is almost level with the ɔ hub ≻. That is why the position of bar 2 of the vowel is changed very little.



o (*obey*) *obey*

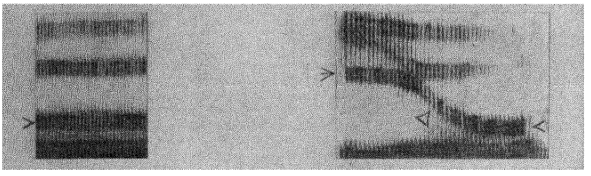
The main hub of the o (*obey*) ≻ is a little lower than the ɔ (*all*) hub; and, since bars 2 and 1 are close together, they are both lower in position than they are for ɔ (*all*). Bar 3 of the vowel o is frequently absent, but when it is present its position is a little lower than it is in ɔ (*all*). The vowel o is used by American speakers in only a few words and occurs most frequently as the first element in the diphthong oʊ (*go*). In the word *obey*, the hub of o ≻ is a little higher than the hub of b ≽, and therefore the end of bar 2 of o curves downwards towards the b hub.



ʊ (*put*) *put*

The main hub of ʊ (*put*) ≻ is a little higher than the hub of o (*obey*) but bar 1 is nearer to the baseline than it is in o. Consequently, bars 1 and 2 are usually separated in the vowel ʊ. When bar 3 is present in the ʊ pattern, its position is slightly lower than the position of bar 3 in o (*obey*). In *put*, the hub of p ≽ is lower than the hub of ʊ, and the hub of t ≺ is higher than the hub of ʊ. As *put* is said, the beginning of bar 2 of ʊ curves upwards towards the t hub. Since the ʊ is a short vowel, the result of the p and t influence is to angle bar 2 upwards from the p hub to the t hub.

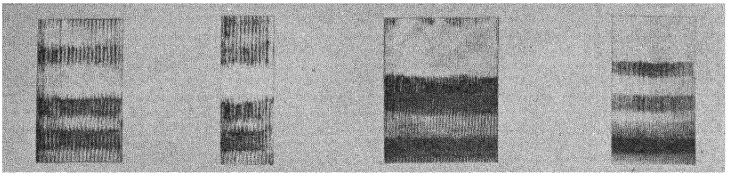
The main hub of u (*you*) ≻, shown at the top of the next page, is just a little lower than the hub of ʊ (*put*). However, bar 1 of the u (*you*) is always flat on the baseline just as it is for the i (*we*). It should be noted that there is usually a small space between



u (*you*) *you*

bars 1 and 2. When bars 3 and 4 are present in the u, they occupy approximately the same positions that they do for the ʊ (*put*) vowel. In the word *you*, the beginning of the j glide ⩾ is near the i (*eve*) position. Remember that the i hub has a higher position than any other vowel hub. The main hub of the u (*you*) ≺ has the lowest position of any of the vowel hubs. As *you* is said, bar 2 of the u is curved downwards ⩾ from the ɪ hub to the ʊ hub.

Each Mid-Vowel Has a Distinct Pattern



ʌ (*up*) ə (*about*) ɚ (*church*) (*General American*) ɝ (*church*) (*Eastern*)

The patterns of the mid-vowels ʌ (*up*), ə (*about*), ɚ (*church*), and ɝ (*church*) are different from each other and from the front and back vowel patterns. The vowel ʌ as in *up* often is called the stressed central or the neutral vowel because during its production the articulators are approximately in a position of rest. The ə as in *about* frequently is referred to as an unstressed ʌ (*up*) sound, and sometimes it is called the *schwa* vowel. Both of these vowels produce similar patterns except that the pattern of the unstressed vowel is generally lighter and shorter than that of the stressed vowel. The term *schwa* is used also by some phoneticians to represent the unstressed forms of the other vowels. The patterns of stressed and unstressed sounds have been discussed previously in Chapter 5.

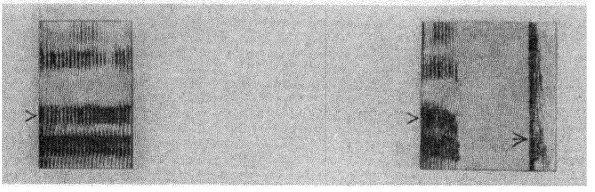
The vowelized r, as in the word *church*, has two common American pronunciations — the General American and the Eastern. A frequently used phonetic symbol for the General American r sound is ɚ and the symbol for the Eastern vowel is ɝ. The General American ɚ is one of the easiest sounds to recognize because no other vowel has bars

2 and 3 placed in the middle of the pattern, and the pattern closely resembles the pattern for the r as in *run*. Bars 2 and 3 of the Eastern ɜ are not together, and the pattern is also vowel-like in appearance.

Since the illustrations in the following chapters are made by speakers from different sections of the United States, both the General American and the Eastern r occur. The reader of visible speech should become accustomed to recognizing both forms of this sound.

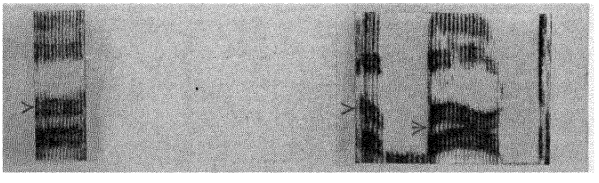
It is helpful to associate the positions of the articulators used in making the sounds with the visible speech patterns.

The main hub of the mid-vowels is the position of bar 2



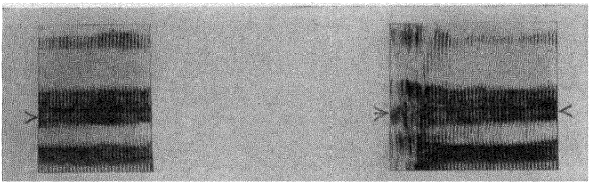
ʌ (*up*) *up*

The main hub of ʌ (*up*) ≻ is the position of bar 2 just a little below the middle of the pattern. Bar 1 is off the baseline about the same distance that it is for the vowel a as in *ask*. Bar 3 is also in the approximate position that it has for the a (*ask*). In the word *up*, the hub of the p ≽ is lower than the ʌ hub, and so bar 2 of the vowel curves downwards towards the hub of the p.



ə (*about*) *about*

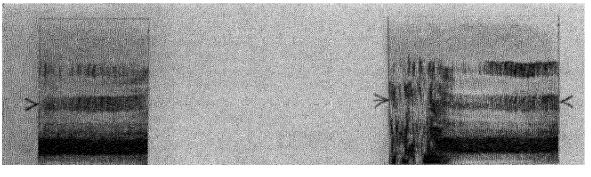
The main hub of ə (*about*) ≻ is almost the same as the hub of ʌ (*up*), although it is sometimes very slightly lower. The positions of bars 1 and 3 also correspond to the ʌ, but bar 3 is often absent. In the word *about*, the hub of b ≽ is lower than the ə hub, so that bar 2 of the vowel curves downwards towards the b hub.



ɝ (*her*) *her*

As mentioned previously, the General American ɝ (*her*) is the only sound in our language that has bars 2 and 3 placed together near the middle of the pattern. It is, therefore, one of the easiest sounds to recognize and to read. Bar 1 of the ɝ is off the baseline. The hub of the ɝ > is the position of bar 2, and is a little below the middle of the pattern like the hub of the vowel a (*ask*). The hub of the h > as it was said in the word *her* is only slightly higher than the ɝ hub < and therefore has very little influence on the vowel pattern.

The symbol ɝ has been used in this text to designate this vowel when it occurs in both the stressed and unstressed positions. Some phoneticians have used symbols which differentiate the stressed and unstressed occurrences, but the similarity in visible pattern in both cases makes such a differentiation unnecessary for our present purpose.

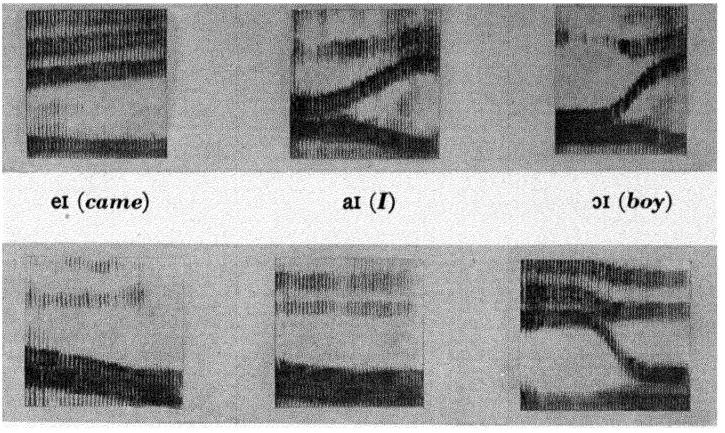


ɜ (*her*) *her*

The main hub of the Eastern ɜ (*her*) > is the position of bar 2 which is just below the middle of the pattern and is slightly higher than the hub of the General American ɝ (*her*). Bar 1 of the ɜ is in the same position as bar 1 of the ɝ. However, unlike the ɝ, the third bar of the ɜ is separated from bar 2. The hub of the h in *her* is not sufficiently different from the ɜ hub to exert much influence on the vowel bars.

Each Diphthong Has a Distinct Pattern

On the next page are shown the patterns of the most frequently used diphthongs. A diphthong usually is defined as a combination of two vowels in the same syllable, in which the speaker glides continuously from one vowel to another. Each of the diphthongs produces a distinctive pattern composed of the first vowel position from which



the glide starts, the glide, and the second vowel position at which the glide terminates. Because a diphthong is a combination of two vowels, it has two main hubs. The hub of the first vowel influences and is influenced by the sound which is spoken before the diphthong. The hub of the second vowel is related in the same ways to the sound which follows. Since the patterns and hubs of the individual vowels have been discussed previously, it is not necessary to consider each vowel of the diphthong in detail.

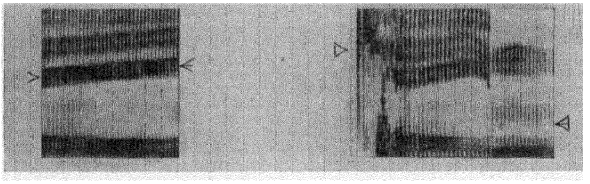
Notice that the second vowel in the diphthongs eɪ (*came*), aɪ (*I*), and ɔɪ (*boy*), is the same. It is the vowel ɪ as in *it*. Because the hub of the ɪ (*it*) is higher than the hub of the first vowel in each of these three diphthongs, bar 2 curves upwards from the first vowel hub to join the hub of the ɪ (*it*). These diphthongs are sometimes called *ascending glides.*

The second vowel of the diphthongs aʊ (*out*) and oʊ (*go*) is the ʊ as in *put.* In both diphthongs, the ʊ hub is lower than the hubs of the first vowels and therefore bar 2 curves downwards from the first vowel to the second vowel. The diphthong ɪu (*new*) is the third of the common *descending glide* diphthongs, and of the three, it has the most marked curve of bar 2.

Try to associate the positions and movements of your articulators with the pattern that each diphthong makes as you say the diphthongs aloud.

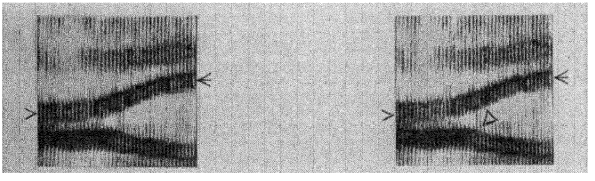
The main hub of the diphthong is the position of bar 2 of each vowel in the diphthong

The first vowel of the diphthong eɪ (*came*) is e (*hate*) with its hub indicated at ≻. The second vowel is the ɪ (*it*) with its hub at ≺. Observe the glide from the first vowel



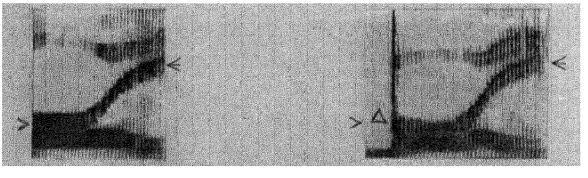
eɪ (*came*) *came*

position to the second vowel position, which is evident in bars 1, 2, and 3. In the word *came*, the **k** hub ⋗ is near the **e** hub, and therefore does not change the position of the beginning of bar 2 of the diphthong very much. However, the **m** hub ⋖ is lower than the **ɪ** hub, and so the end of bar 2 of the diphthong curves downwards towards the **m** hub. Try to recall how the **eɪ** looked in the word *eight*.



aɪ (*I*) ***I***

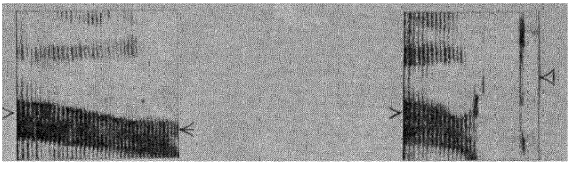
The diphthong **aɪ** (*I*) consists of the vowel **a** (*ask*) with its hub indicated at ≻, followed by the vowel **ɪ** (*it*) with its hub indicated at ≺. The word *I*, consists of the diphthong alone and therefore both the above patterns are identical. Notice how bar 2 of the first vowel curves △ into bar 2 of the second vowel to make the diphthong. Do you remember how the diphthong **aɪ** looked in the words *five* and *buy*?



ɔɪ (*boy*) ***boy***

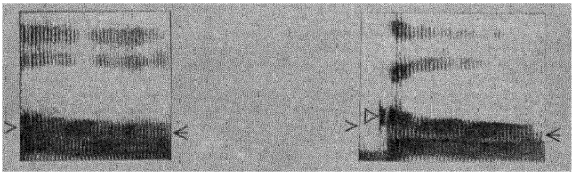
The first vowel of the diphthong **ɔɪ** (*boy*) is the **ɔ** (*all*) with its hub indicated at ≻. The second vowel is the **ɪ** (*it*) with its hub at ≺. The hub of the **b** ⋗ in the word *boy* is in almost the same position as the hub of the first vowel, and therefore the beginning of bar 2 of the diphthong is not curved noticeably. The glide from the first to the

second vowel is most prominent in the ɔɪ (*boy*), and is least marked in the eɪ (*came*). This difference is due to the amount of separation between the hubs of the two vowels that constitute each diphthong.



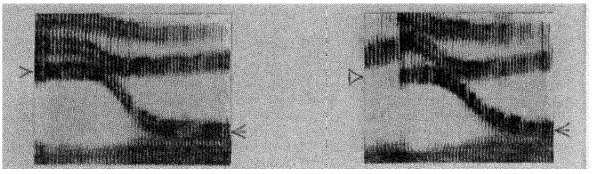
aʊ (*out*) *out*

The diphthong aʊ (*out*) combines the vowel ɑ (*father*) with its hub indicated at ≻, with the vowel ʊ (*put*) with its hub at ≺. In the word *out*, the hub of t ◁ is higher than the hub of the second vowel, and therefore the end of bar 2 of the diphthong curves upwards towards the t hub. Notice that the diphthong spoken before the voiceless stop sound t, is considerably shorter than when it is said by itself. Many American speakers initiate this diphthong with the vowel a as in *ask*, and for those speakers the diphthong should be written as au. While differences in pronunciation are important to the phonetician, they need only be noted casually by the beginning student of visible speech. If a vowel sound is recognized and read as the vowel in the vowel series that its pattern most closely approximates, the reading will be accurate enough for the practical purpose of communicating information.



oʊ (*go*) *go*

The diphthong oʊ (*go*) consists of the vowel o (*obey*) with its hub indicated at ≻, followed by the vowel ʊ (*put*) with its hub indicated at ≺. In comparing the glides in the diphthongs aʊ (*out*) and oʊ (*go*), notice that, because of the greater separation of hubs, the aʊ (*out*) glide is the steeper. The hub of g ▷ in the word *go*, is a little higher than the hub of the first vowel, and so the beginning of bar 2 of the diphthong curves downwards a little from the g hub.



ɪu (*new*) *new*

The first vowel of the diphthong ɪu (*new*) is ɪ (*it*) with its hub indicated at ≻. The second vowel is the u (*boot*) with its hub at ≺. Because of the wide difference between the hubs of these vowels, bar 2 of this diphthong evidences a marked glide which makes it quite easy to recognize. Compare this diphthong with the pattern of the word *you*, and notice that bar 3 of the ɪu (*new*) is closer to bar 2 than it is in the ju (*you*) combination. In the word *new*, the n hub ≻ is slightly below the hub of the first vowel and therefore the beginning of bar 2 of the diphthong may curve upwards a little from the n hub before gliding downwards to the second vowel.

The patterns and hubs of the vowels and diphthongs have been presented in order to establish the visual reference points used in recognizing the patterns of these sounds when they appear in combination with other sounds. The next step in studying visible speech is to learn the characteristic patterns of the stop consonants and how they influence the patterns of the vowels and diphthongs. The reading vocabulary is progressively enlarged in the stop consonant chapter and in the subsequent chapters which deal with the fricatives and vowel-like sounds. Then the vowels and diphthongs are presented for the second time, in order that the influences of the consonants and vowel-like sounds may be summarized for each vowel and each diphthong. Thus, the vowels and diphthongs are presented twice: first, as visual reference points, from which other sounds may be studied; second, as a summary of how they influence and are influenced by associated sounds.

Chapter 8

Reading the Stop Sounds

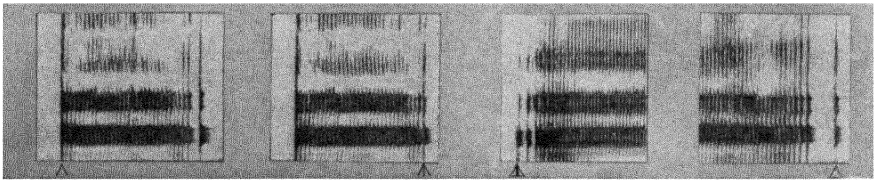
In this chapter we shall discuss the characteristic patterns by which the stop sounds are recognized and read. The six sounds included in this classification are the **p** (*pay*), **b** (*be*), **t** (*to*), **d** (*day*), **k** (*key*), and **g** (*go*). Patterns of this general group are recognized by two essential clues:

1. Voiceless or voiced stop gaps produced during the period of building up of breath pressure.
2. Voiceless or voiced spike fills produced by the sudden release of breath pressure.

The voiceless stop sounds **p**, **t**, and **k**, are produced primarily by a combination of stop and fricative modulation, but the voiced stop sounds **b**, **d**, and **g**, combine vocal cord modulation with both the stop and fricative periods of their voiceless cognates. Sometimes in rapid, continuous speech the breath stream is not stopped completely. In that case, the stop gap is either short or missing, and the vertical striations of the spike fill become the more prominent part of the pattern. Similarly the voicing that occurs during the stop period for the **b**, **d**, and **g** may be short or weak, or occasionally may be absent. When the voicing is strong, upper resonances are sometimes visible in the stop gap. The duration and volume of the voicing are shown by the length and darkness of the voice bar in the stop gap.

The transitional movement or influence of the stops on the vowels usually is more evident when the stop precedes the vowel than when it follows the vowel. Two of the factors related to this pattern difference are first, the period of the stop sound which is adjacent to the vowel; and second, the rapidity of movement of the articulators. If the stop sound is said after the vowel, the period during which the breath flow is stopped follows immediately after the vowel bars and, thus, usually only part of the transitional pattern appears. If the stop sound is made before the vowel, the period during which the breath pressure is released immediately precedes the vowel bars, and the transitional movement is apparent in the voiceless or voiced pattern that follows the sudden release of breath. The movement of the articulators from the vowel to the stop position frequently is made more quickly than the movement from the stop to the vowel, and therefore the influence of the stop following the vowel bars is not always as apparent as it is when the transitional movement either precedes the vowel or is prolonged. Because voicing continues between the voiced stop and the vowel, the influence of the voiced stop sounds often may be observed more easily than that of the voiceless stops, and also the voiced spike fills are narrower than the voiceless.

Although the glottal stop is recognized as an indication of stress, it usually is considered as a non-information-bearing sound in American speech. Since it occurs at random in combination with voiced sounds, it will be discussed briefly at this point rather than in a separate lesson unit. The glottal stop commonly is used as a way of either beginning or terminating a voiced sound. Some speakers substitute the glottal stop for any or all of the stop sounds, especially when they occur in syllabic combination with other sounds. It also appears in some dialects and in several languages, and, in defective speech, it is used frequently as a substitute for many sounds that are difficult to make. The glottal stop is produced by a complete closure of the vocal folds or cords which causes the breath pressure to be built up below the glottis. When the vocal folds are separated suddenly, thus opening the glottis quickly as vocal vibration is begun, the release of breath pressure produces the explosion characteristic of a voiced stop. When the vocal folds move suddenly from the vocal vibratory position to a complete closure, the voiced breath is cut off abruptly. Since the glottal stop is produced laryngeally, it does not involve special positions of the articulators of the upper vocal tract. It does not have a separate hub, nor does it influence the sounds with which it is combined. Therefore, there is little possibility of confusing the glottal



glottal beginning	*glottal termination*	*increase in vibration rate*	*vocal flaps*

The glottal stop used before a vowel is portrayed by an abrupt start of the pattern and often by a single spike ⋀. The glottal stop used to terminate a voiced sound appears on the pattern as an abrupt ending ⋀ of the sound. The vocal vibration which follows a glottal stop does not always begin at its normal vibratory rate, and an increase in rate of vibration may be observed in the decreasing distance between adjacent vertical striations ⩓. An extremely slow rate of beginning or terminating vocal fold vibration produces widely separated vertical striations that have been called vocal flaps ⩓.

The transitional influences of each of the sound patterns are considered specifically in word combination with three vowels: a front vowel, a back vowel, and a vowel which occurs towards the middle of the vowel series. From the illustrations and explanations in the accompanying text, the reader should be able to comprehend and to

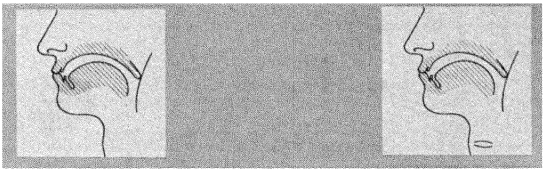
visualize the basic principles governing the transitional influences of each sound pattern unit upon the various vowels and diphthongs.

Since the patterns of the sounds vary according to the manner of their production, it is advisable to subdivide the general sound groups into smaller units for more detailed consideration.

Unit 4 presents a summary and comparison of the clues by which the various stop sounds may be recognized. This summary is intended to facilitate comparisons among the voiceless and among the voiced stop patterns, hubs, and influences.

Unit 1 — Reading the p (pay) and b (be)

How the p and b are made



Profile of p (*pay*) *Profile of* b (*be*)

The **p** and **b**, as indicated in the profiles above, are produced by:

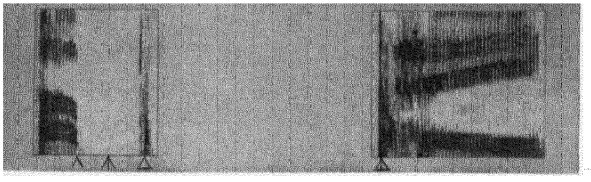
1. Stopping the breath stream by closing the lips and closing the opening between the throat and nose.
2. Building up voiceless or voiced breath pressure in the closed oral cavity.
3. Exploding the voiceless or voiced breath by quickly opening the lips.

Although the articulators are in almost the same positions for the **p** and **b**, the sounds are modulated differently. The **p** is produced primarily by stop-start and fricative modulation; but in the **b**, vocal cord and closed cavity modulation are added to both the stop-start and the fricative periods. Therefore, the **p** and **b** are referred to as cognates: the **p** as a voiceless stop or plosive sound, and the **b** as its voiced equivalent. In physiologic phonetics, the **p** is described as a bilabial voiceless stop, and the **b** is designated a bilabial voiced stop. The **p** is said without vocal cord modulation, while the **b** is produced with vocal cord modulation. Therefore the **p** is referred to as a voiceless sound, and the **b** is called a voiced sound.

The p and b may be spelled in different ways

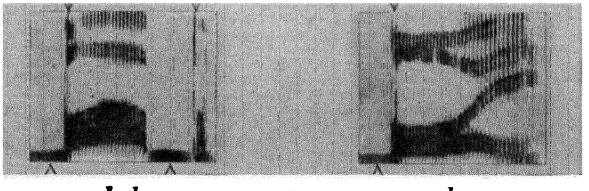
p (*pay*) *pe* (*ape*) *pp* (*happy*) *b* (*be*) *be* (*tube*) *bb* (*ebb*)

Visible patterns of the p and b



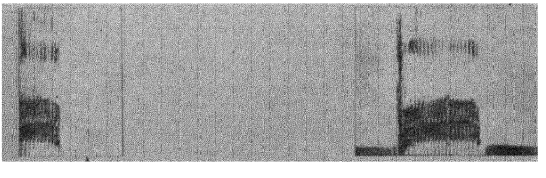
up ***pay***

Since the breath is stopped for the **p** Λ and no energy is emitted from the mouth or nose, there is a blank space or stop gap Λ on the pattern. The release of the voiceless breath pressure results in a spike fill as indicated at Λ and Λ. This spike fill is usually darker near the baseline. The puff of air which follows the release of pressure is fricative in nature, and is indicated by the presence of the vertical striations of the fill.



bob ***boy***

Whenever the voiced breath is stopped for the **b**, there is a voiced stop gap Λ on the pattern. Voicing is shown by the voice bar which appears in the stop gap and which follows the spike fill. The voiced spike fill for the **b** Ψ is narrower and sharper than the voiceless spike fill for the **p.**



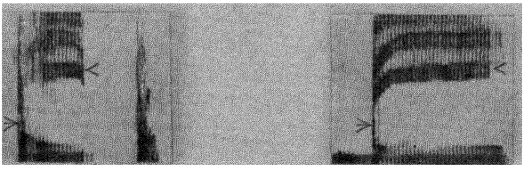
up ***bob***

During continuous speech, the **p** and **b** in the final position often are not exploded although the breath is stopped and pressure is built up. Such sounds are called *imploded* stops. When this occurs the spike fills are absent and the stop gaps look like

those indicated above at ∧ and ∨. Compare these patterns with the patterns of the exploded **p** and **b**.

Because the **p** and **b** are made from similar articulatory positions they have the same hub. While the hub is not always apparent in the consonant patterns, it is indicated clearly in the influence of the **p** and **b** on the vowels. The **p** and **b** hub is located near the baseline in approximately the position of the hub of the vowel **ʊ** (*put*), and the spike fills of the sounds are often darkest in that area.

How the p and b influence vowel patterns

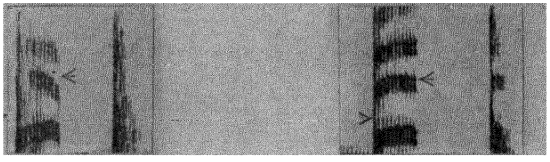


peep ***be***

The patterns of *be* and *peep* show the influence of the **p** and **b** on the vowel **i** (*eve*). The hub of this vowel is near the top of the pattern ≺, while the hub of the **p** and **b** is near the baseline ≻. In both words, note that the beginnings of bars 2 and 3 of the **i** (*eve*) curve upwards from the **p–b** hub. The endings of bars 2 and 3 also curve downwards towards the final **p** hub in *peep*. Bar 1 of the vowel **i** is always on the baseline, and the influence of the **p** and **b** on its position is not considered to be significant for reading. The transitional influences are usually more apparent in vowels preceded or followed by voiced sounds than in those combined with voiceless sounds. In addition, the difference in the duration or length of the vowels in these and other words should be observed, since it is generally true in continuous speech that a vowel is shorter when it is made in isolation or when it precedes a voiceless sound than it is when it precedes a voiced sound. The movement of the articulators from the vowel to the stop position is frequently made so quickly that the influence of the stop on the vowel bars is not as apparent as it is when the transitional movement is slower.

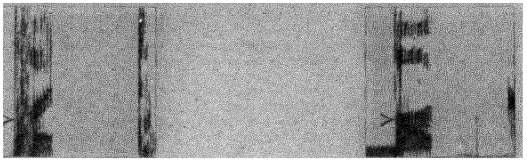
Note also that the **b** spike is closer to the vowel bars than is the **p** spike, and that the **p** spike fill is the wider. These characteristics of the **p** and **b** should be observed in the following illustrations.

In the patterns of the words *pep* and *bet*, at the top of the next page, the beginnings of vowel bars 2 and 3 do not curve up as far as they do in *peep* and *be*. That is true because the main hub of the vowel **ɛ** (*met*) is just a little above the center of the pattern ≺ and therefore is closer to the **p–b** hub ≻ than is the hub of the **i** (*eve*). The endings of bars 2 and 3 point downwards in *pep*, and again the curvature is less



pep ***bet***

marked than it is in *peep* due to the difference in the amount of separation between the hubs. Bar 1 of the vowel ɛ (*met*) is off the baseline a little, and a slight up-curve of the beginning of the bar in the word *bet* as the vowel follows the voice bar and spike of the **b** may be observed. Differences in the patterns of the **p** and **b** should be noted in these words as they were in *peep* and *be.*



put ***book***

The hub of the vowel **ʊ** (*put*) is almost in the same position as the hub of the **p** and **b** >. Consequently, the beginning of bar 2 of the vowel is not curved noticeably in the words *put* and *book.* Bar 1 of the vowel is a little off the baseline, and is very slightly curved in *book* as it connects with the voice bar and spike of **b.**

The **p** and **b** influence bars 2 and 3 in the vowel **i** as in *peep* and *be* more than they do in any other vowel, because the separation between hubs is at its maximum. The **p** and **b** influence the vowel **ʊ** as in *put* and *book* the least, because the position of the hubs is so similar.

The visible patterns of the **b** *and the* **p** *differ chiefly in that:*

1. The **b** spike fill usually is preceded and followed by a voice bar resulting in a voiced stop gap and a voiced spike fill.
2. The **b** spike fill is narrower than that of the **p.**
3. The **b** spike fill is closer to the sound that follows.

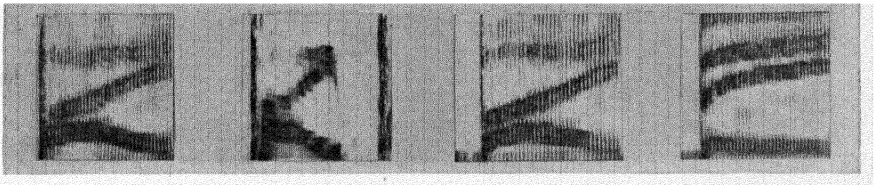
Read the **p** *and* **b** *by:*

1. The stop gap preceding the spike fill.
2. The appearance of the spike fill.
3. The influence on the vowels.

Review words

peep, page 83 *pay*, page 82 *up*, page 82 *bet*, page 84 *book*, page 84
pep, page 84 *put*, page 84 *be*, *bee*, page 83 *boy*, page 82 *bob*, (*Bob*), page 82

New words to learn

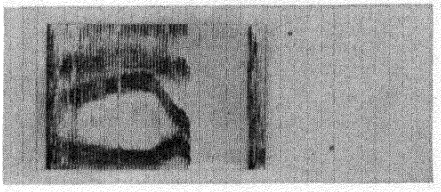


pie *pipe* *buy*, *by*, *bye* *bay*

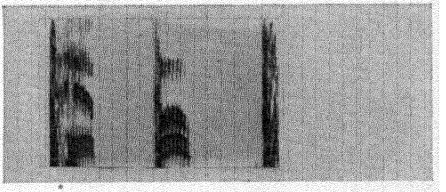
Additional words for practice in visualization

pa *pup* *puppy* *pop* *baby* *obey* *bow* (baʊ) *bow* (boʊ)

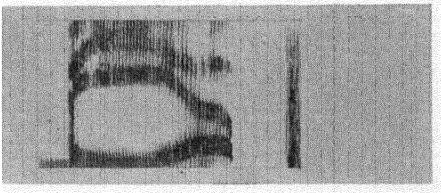
Drill phrases and sentences



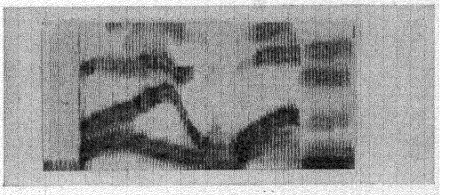
Pay up



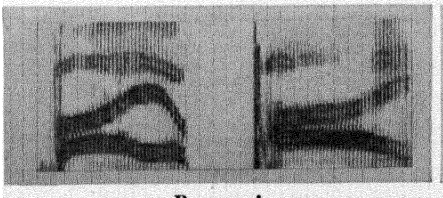
Pep up



Be up



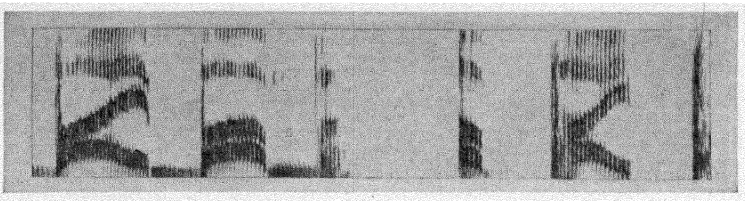
Buy one



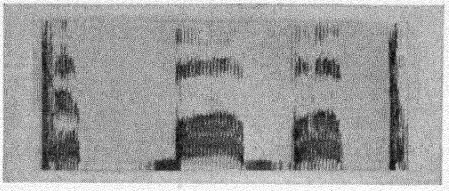
Buy a pie



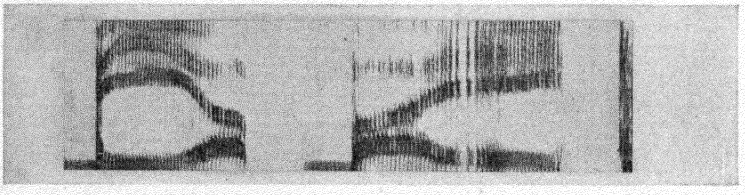
Buy a pipe



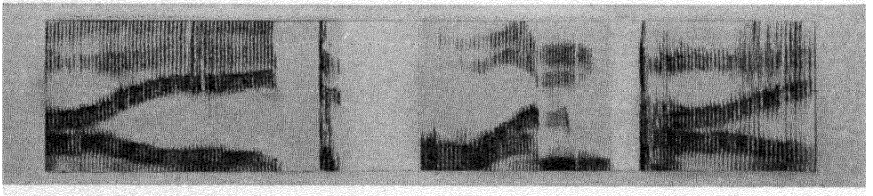
Buy Bob a pipe



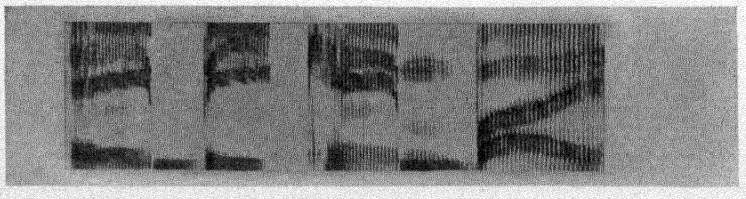
Pep Bob up



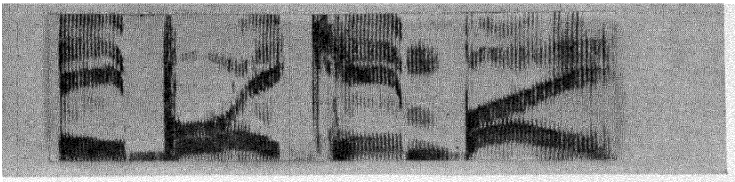
Be up by eight



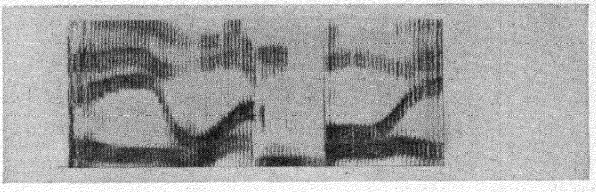
I ate one pie



A bee came by



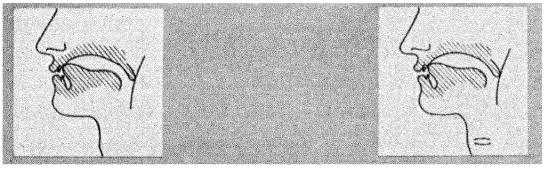
A boy came by



Pay one boy

Unit 2—Reading the t (to) and d (day)

How the t and d are made



Profile of t (*to*) *Profile of* d (*day*)

The **t** and **d** are produced by:

1. Stopping the breath stream by placing the tip of the tongue against the gums (alveolar ridge) of the upper front teeth, placing the sides of the tongue against the side teeth, and closing the opening between the throat and nose.
2. Building up voiceless or voiced breath pressure in the closed oral cavity.
3. Exploding the voiceless or voiced breath by quickly moving the tongue tip away from the gums (alveolar ridge).

The **t** and **d**, as indicated in the profiles above, are produced from similar articulatory positions yet they are modulated differently. The **t** is produced primarily by stop-start and fricative modulation; but in the **d**, vocal cord and closed cavity modulation

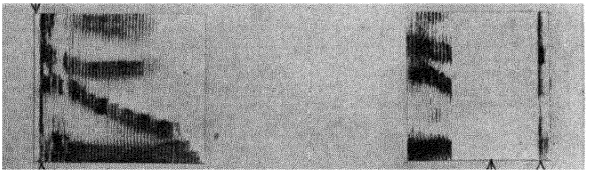
are added to both the stop-start and the fricative periods. Therefore, the **t** and **d** are referred to as cognates: the **t** as a voiceless stop or plosive, and the **d** as its voiced equivalent. In physiologic phonetics, the **t** is described as a lingua-alveolar voiceless stop, and the **d** as a lingua-alveolar voiced stop.

The t and d may be spelled in different ways

t *(to)*	***tt*** *(matter)*	***th*** *(thyme)*	***d*** *(day)*	***dd*** *(add)*
te *(rite)*	***tw*** *(two)*	***pt*** *(ptomaine)*	***ed*** *(seemed)*	***de*** *(fade)*
bt *(debt)*	***ght*** *(eight)*	***cht*** *(yacht)*		
ed *(walked)*				

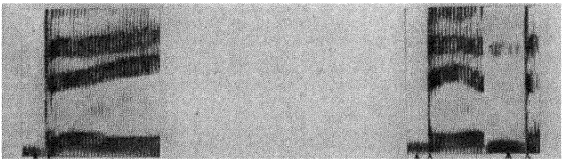
Visible patterns of the t and d

The **t** and **d** are stop sounds like the **p** and the **b**, and therefore their patterns have stop gaps and spike fills.



to *it*

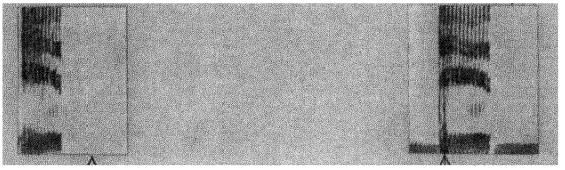
The quick release of the voiceless breath pressure for **t** makes a spike fill that looks like Λ in *to* and *it*. The stop gap ⅄ shows that the breath is stopped from the end of the vowel until the breath pressure is released at the beginning of the spike fill. Part of the stop gap Ψ is visible before the **t** spike fill in *to*.



day *did*

The quick release of the voiced breath for **d** makes the narrow, voiced spike fills Λ in *day* and *did*. Notice the voice bar ⅄ on the baseline at the bottom of the stop gap.

This voice bar extends through the stop gap and is connected with the spike fill. It usually appears on the baseline for voiced stop sounds.

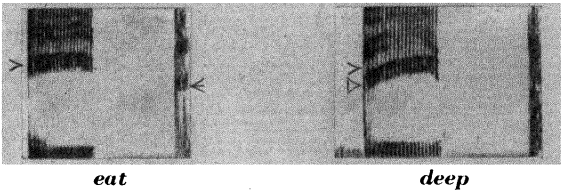


it ***did***

When the **t** and the **d** are imploded, the spike fills are not portrayed on the patterns and the stop gaps look like ∧ and ∨ in *it* and *did*. Notice that the exploded, initial **d** ∧ in *did* is closely joined to the vowel.

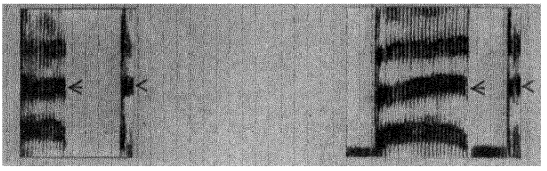
The **t** and **d** have the same hub because they are made from similar articulatory positions. The location of the **t**–**d** hub is approximately the same as the hub of **æ** (*at*). Consequently, the influence of these stop sounds on adjacent sounds can be predicted.

How the t and d influence vowel patterns



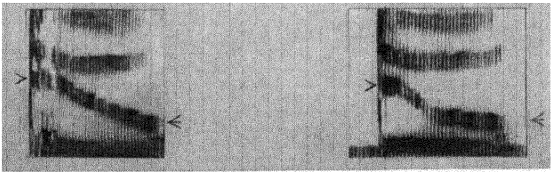
eat ***deep***

The patterns of the words *eat* and *deep* show the influence of the **t** and **d** on the vowel **i** (*eve*). The hub of the vowel is towards the top of the pattern and the **t**–**d** hub is near the center of the pattern. The difference in hub position is indicated in the curves of the vowel bars. When **i** (*eve*) precedes **t** as in the word *eat*, the rapid movement of the articulators in stopping the flow of breath to make the **t** while the vocal cords are still vibrating for the vowel causes the ending of bar 2 to point downwards from the **i** hub ≻ towards the **t** hub ≺. If the vibration of the vocal cords is stopped before or at the instant the articulators move from the vowel position, little influence is noted on the vowel bars. When **d** precedes the vowel **i** (*eve*) as in *deep*, the beginning of bar 2 curves upwards from the **d** hub ≻ to the hub of **i**. Notice that the voicing continues from the **d** to the vowel, and that this causes the influence to be more easily recognized than when the stop follows the vowel. Bar 3 in both *eat* and *deep* is curved like bar 2. Bar 1 is on the baseline and it is not noticeably influenced by the consonants.



at *bad*

The **t** and **d** have very little influence on the vowel **æ** as in *at* and *bad.* This is because the **t–d** hub $<$ is in approximately the same position as the **æ** hub $\lessdot$. In the patterns of *at* and *bad* the endings of bars 2 and 3 are almost as straight as they are when the vowel **æ** is made separately. The end of bar 1 in both patterns curves downwards towards the baseline. Compare the **b** and **d** influences on **æ** in the pattern of the word *bad.*



to *do*

The **t** and **d** influence on the back vowel **u** (*to*) is shown in the patterns of the words *to* and *do.* Bar 2 in both words curves downwards from the **t–d** hub $>$ to the hub of the vowel $\lessdot$. The steepness of the curves is determined by the rapidity of the movements of the articulators in going from the **t–d** position to the **u** position. The beginning of bar 3 in both words curves downwards and bar 1 is on the baseline uninfluenced by the **t** and **d.**

The visible patterns of the **d** *and the* **t** *differ chiefly in that:*

1. The **d** spike fill usually is preceded and followed by a voice bar resulting in a voiced stop gap and a voiced spike fill.
2. The **d** spike fill is narrower than that of the **t.**
3. The **d** spike fill is closer to the sound that follows.

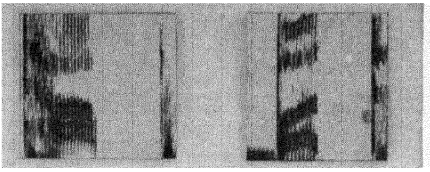
Read the **t** *and* **d** *by:*

1. The stop gap preceding the spike fill.
2. The appearance of the spike fill.
3. The influence on the vowels.

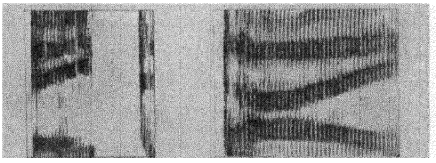
Review words

eat, page 89
bet, page 84
to, two, too, pages 88,90
it, page 88
at, page 90
put, page 84
deep, page 89
bad, page 90
day, page 88
did, page 88
do, page 90

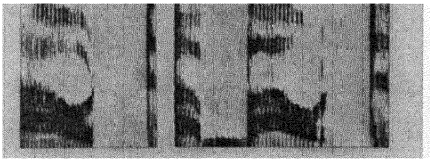
New words to learn



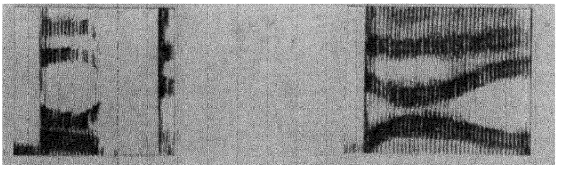
top *but* *bed* *dead*



ate, eight *tie* *add* *body*



out *about* *bird* *today*

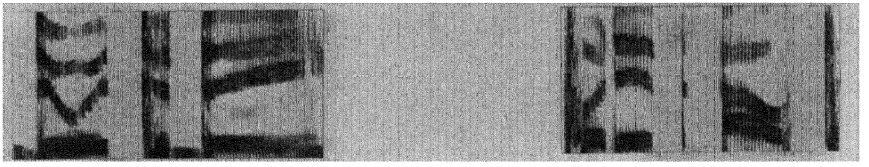


boat *die, dye*

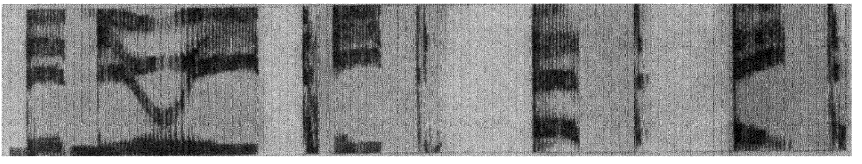
Additional words for practice in visualization

beat	*tight*	*bit*	*deed*	*dad*
ought		*tape*	*dirt*	*doubt*

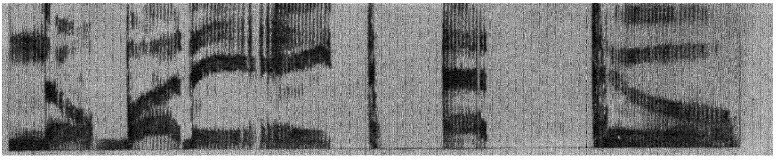
Drill phrases and sentences



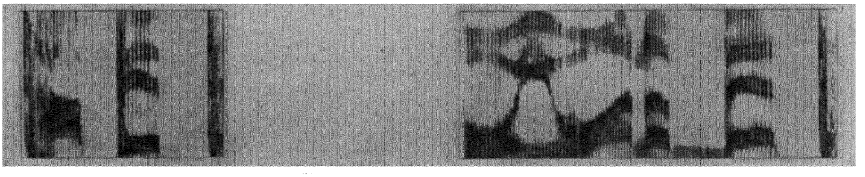
Do it today *Put it out*



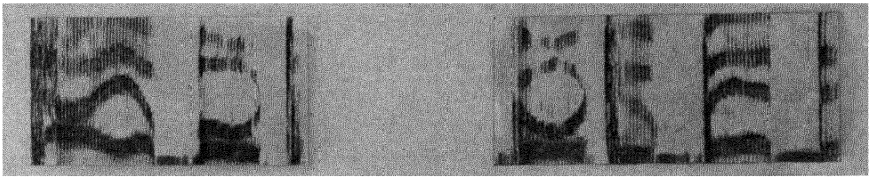
Did you eat? *Eat at eight*



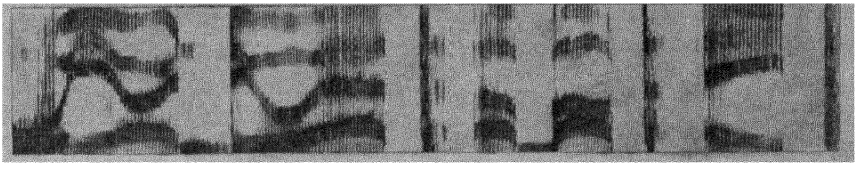
Nobody ate at two



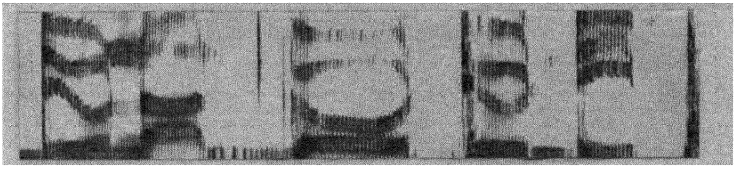
Top it *We won a bet*



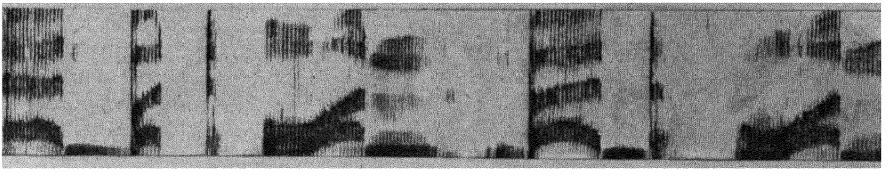
Tie a boat *Go to bed*



We are due at about eight



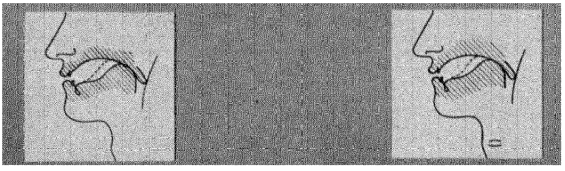
Do not go too deep



Add but one bad one

Unit 3 — Reading the k (key) and g (go)

How the k and g are made



Profile of **k** (*key*) *Profile of* **g** (*go*)

The **k** and **g** are produced by:

1. Stopping the breath stream by placing the back of the tongue against the roof of the mouth and closing the opening between the throat and nose.
2. Building up voiceless or voiced breath pressure in the closed oral cavity.
3. Exploding the voiceless or voiced breath by quickly moving the tongue away from the roof of the mouth.

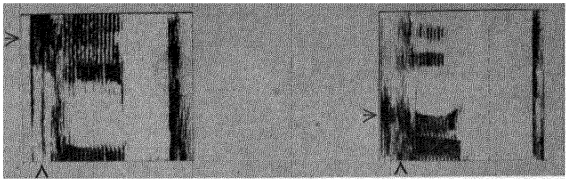
The **k** is made primarily by stop-start and fricative modulation; but in the **g**, vocal cord and closed cavity modulation are added to both the stop-start and the fricative periods. Therefore, the **k** and **g** are referred to as cognates: the **k** as a voiceless stop or plosive, and the **g** as its voiced equivalent. In physiologic phonetics, the **k** is described as a lingua-palatal voiceless stop, and the **g** is designated a lingua-palatal voiced stop. Both the **k** and the **g** are referred to as variable sounds since they may be said a little differently with each sound with which they are combined. Usually, they are made towards the front of the mouth with the front vowels and consonants, and towards the back of the mouth with the back vowels and consonants. These differences in production are apparent in the patterns and a number of **k** and **g** pattern positions may be observed in the illustrations in this unit.

The k and g may be spelled in different ways

k (*keep*)	*kh* (*khaki*)	*ke* (*take*)	*g* (*go*)	*gg* (*egg*)
ck (*back*)	*c* (*can*)	*cc* (*accord*)	*gh* (*ghost*)	*gu* (*guard*)
ch (*chorus*)	*che* (*ache*)	*cu* (*biscuit*)	*gue* (*plague*)	
qu (*conquer*)	*que* (*pique*)	*ugh* (*hough*)		

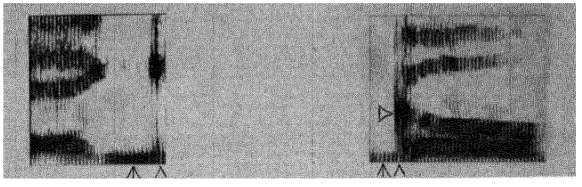
Visible patterns of the k and g

Like the **p–b** and **t–d**, the **k** and **g** are stop sounds, and their patterns are characterized by stop gaps and spike fills.



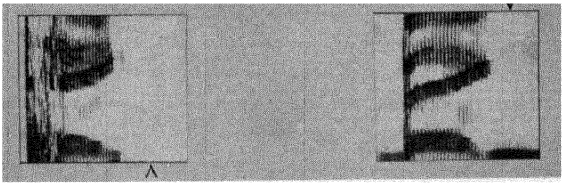
keep *coat*

The spike fill ∧ represents the release of breath for the **k** in *keep* and *coat.* Notice that the darker area of the spike ⩾ is towards the top of the pattern in *keep* and near the bottom in *coat.* The variable production of **k** with the front and the back vowels causes a shift both in the hub of **k** and in the appearance of the pattern. The stop gap, in each case, precedes the spike fill.



egg *go*

The voiced breath for the **g** in *egg* and *go* produces the voiced stop gap ⩚ which precedes the spike fill ∧. The **g** pattern, in each example, illustrates the shift in articulatory position of the **g** with a front and a back vowel. Frequently only the dark area ⩾ of the spike is visible on the pattern.



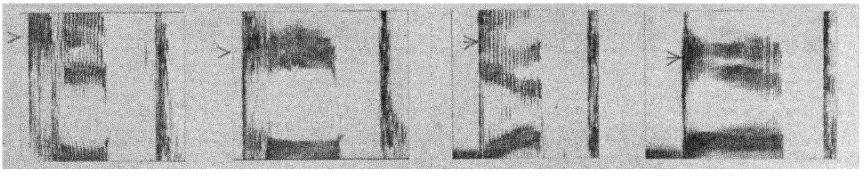
take *beg*

When the **k** and **g** are imploded, the period in which breath pressure is built up is represented by the voiceless stop gap in *take* ∧ and the voiced stop gap in *beg* ⩛. Since the breath pressure is not exploded, there are no **k** or **g** spike fills in the patterns. However, the initial **t** in *take* and the **b** in *beg* are exploded, and their spike fills should be observed and compared.

Since the **k** and **g** are variable sounds, their hub shifts depend upon the sound with which they are combined. The hub of the **k** and **g** appears near the top of the pattern with the front vowels, towards the baseline with the back vowels, and towards the middle of the pattern with the mid vowels. It is of interest to observe that the hub of the initial **k–g** is usually slightly higher than the hub of the final **k–g** with the front vowels, and that the hub of the initial **k-g** is slightly lower than that of the final with the back vowels. It seems, therefore, that there is more of a tendency for the speaker to go towards a mid **k–g** position for the final than for the initial stop. In continuous speech, the position of the **k–g** hub is affected, dependent somewhat upon stress, by the sounds which precede and follow the stop.

As in the other stop sounds, the hub is not always apparent in the patterns, but it is indicated clearly in the influence of the **k** and **g** on the vowels. The dark area of the spike fill is in approximately the same position as the hub. In certain dialects, however, the **k–g** hub may shift differently resulting in aural and visible influences which are too varied and numerous for consideration here.

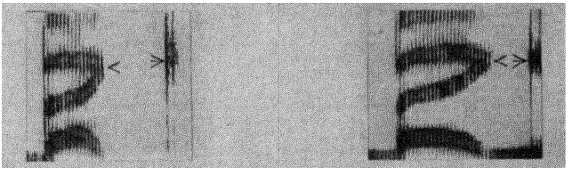
How the k and g influence vowel patterns



keep ***keep*** ***get*** ***get***
(four-bar) ***(three-bar)*** ***(four-bar)*** ***(three-bar)***

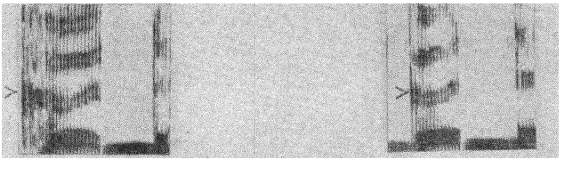
The influence of the **k** and **g** on the front vowels, **i** (*eve*), **ɪ** (*it*), **e** (*hate*), and **ɛ** (*met*), is usually most evident in vowel bars 2 and 3 in three-bar vowel patterns, and in bars 3 and 4 in four-bar patterns. In the four-bar pattern, the rounding together of the beginnings of bars 3 and 4 is apparent in *keep* but becomes progressively less marked with each of the vowels in series from **i** (*eve*) to **ɛ** (*met*). The three-bar vowel pattern of *keep* shows rounding of the beginnings of bars 2 and 3. Both the three and four-bar patterns of **ɛ** in the word *get* show a somewhat similar rounding of bars 2 and 3.

The hub of the **k** > with the **i** as in *keep* is generally towards the top of the pattern. Notice that the **k** hub in the four-bar vowel pattern of *keep* is a little higher than the **k** hub in the three-bar vowel pattern. With the vowel **ɛ** as in *get*, the hub of the **g** > is just a little above the hub of the **ɛ** and the downward curve of bar 2 of **ɛ** from the **g** hub is apparent. It is advisable to compare the influences of the **k** and **p** in *keep* and of the **g** and **t** in *get.*



back *bag*

The **k–g** influence is most marked on the vowel æ (*have*) as is illustrated by the characteristic egg-shaped closure of bars 2 and 3 < in *back* and *bag*. The hub of the æ is near the middle of the pattern, and the hub of the **k–g** > is just a little higher. Therefore, bar 2 of æ curves upwards towards the **k–g** hub. The end of bar 1 of the vowel curves towards the baseline, and in *bag* it joins the voice bar of **g**. Notice the dark area of the spike fill at the hub, and also the voice bar preceding and following the **g** spike fill. Since the **b** hub is located near the baseline, the beginning of bar 2 of æ curves upwards from the **b**, and the second bar moves diagonally upwards from the **b** hub to the **k–g** hub. The voice bar and spike fill of **b** should be compared with the illustrations of the **b** in Unit 1 and with the patterns of the **g** in this Unit.



could *good*

The hub of the k and g > shifts with the vowel **u** (*put*), as in *could* and *good*, to a position just a little above the vowel hub. The spike fills of the **k** and **g** are darker at the hub and, especially in the **k**, the upper part of the spike fill is frequently weak or absent. The rounding of bars 2 and 3 is not as apparent in the back vowels, and bar 3 is often absent or light, but the downward curve of bar 2 from the **k–g** hub may be observed. The difference between the **k–g** and the **d** influences is easily recognized, and the position of the **d** hub should be recalled.

The hub of the **k** and **g** is usually a little higher than the hub of the vowel with which it is combined. The **k–g** influence is shown in the rounding together of vowel bars 2 and 3, or bars 3 and 4. The **k–g** spike fills are darker in the area of the hub and, sometimes, are only as high as the hub.

The visible patterns of the **g** *and the* **k** *differ chiefly in that:*

1. The **g** spike fill usually is preceded and followed by **a voice bar.**
2. The **g** spike fill is narrower than that of **k.**
3. The **g** spike fill is closer to the sound that follows.

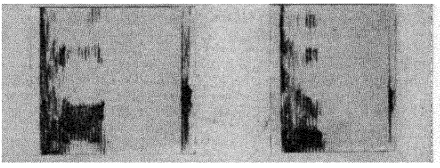
Read the **k** *and* **g** *by:*

1. The stop gap preceding the spike fill.
2. The dark area of the spike fill.
3. The influence on the vowels.

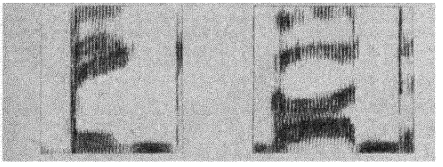
Review words

keep, page 95, 96	*back*, page 97	*egg*, page 95	*get*, page 96
could, page 97	*book*, page 84	*beg*, page 95	*bag*, page 97
take, page 48	*coat*, page 95	*good*, page 97	*go*, page 95

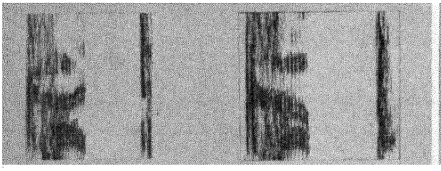
New words to learn



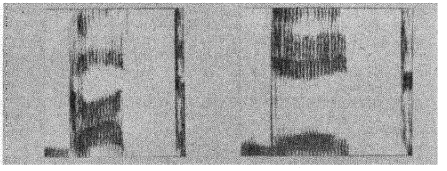
talk *took*



big *God (god)*



cut *cup*

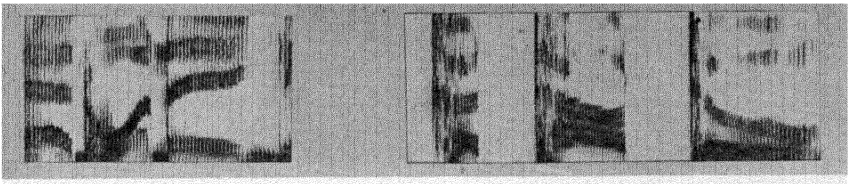


got *gate*

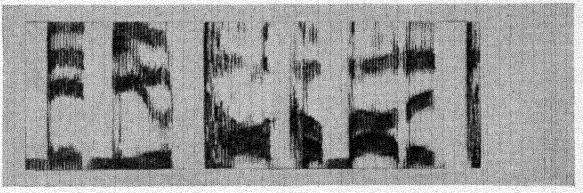
Additional words for practice in visualization

cook	*pick*	*cap*	*gap*	*dog*
cake	*oak*	*cow*	*guide*	*ago*

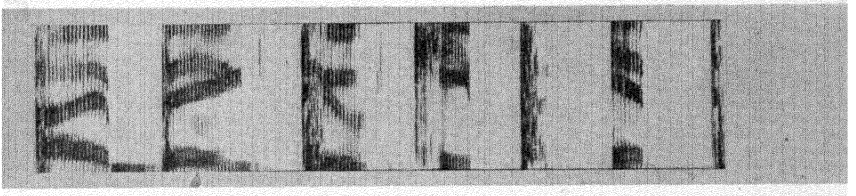
Drill phrases and sentences are shown on pages 99 and 100



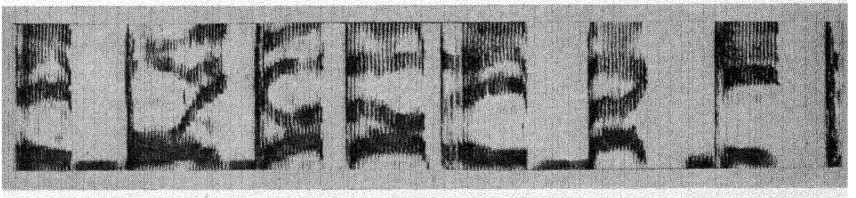
Add one egg *Cut out two*



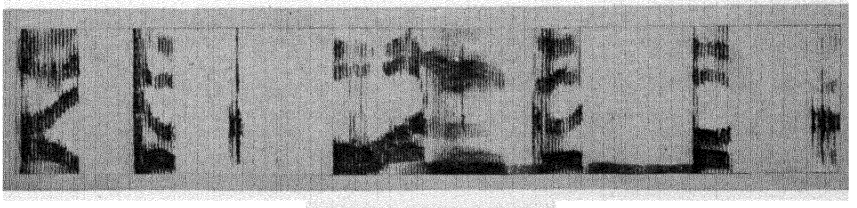
Did you talk about it?



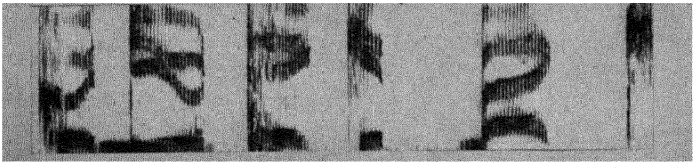
I beg to keep it



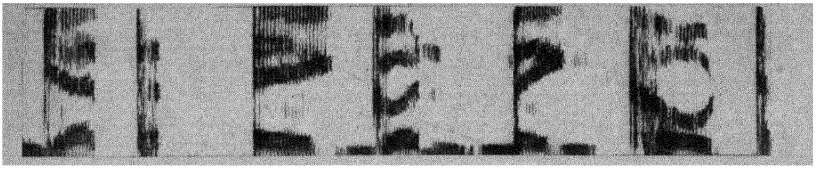
A boy got out a back gate



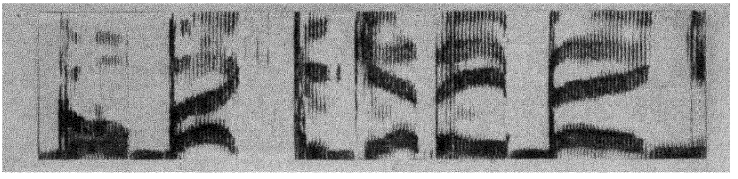
I took one good book



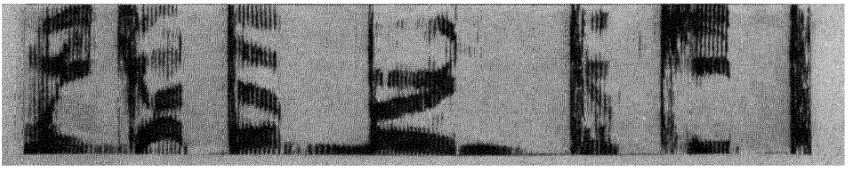
Could you take it back?



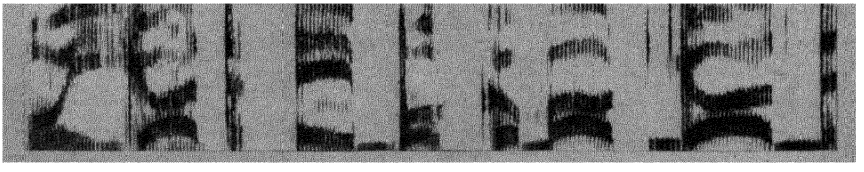
Get a good, big coat



Go back to get a bag



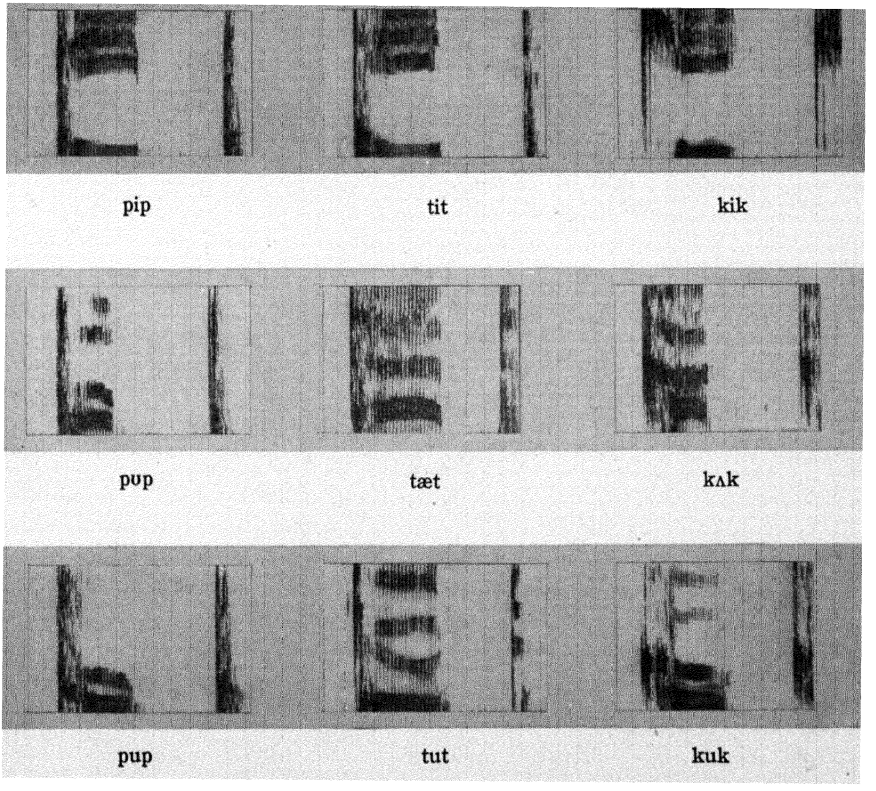
We got a bird to keep



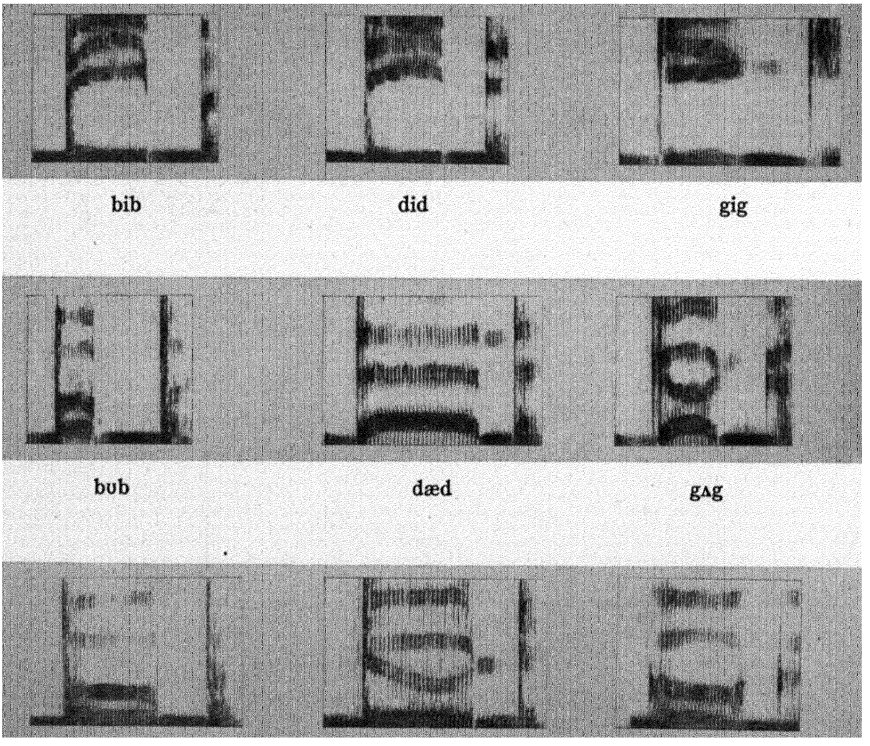
We got a book about God

Unit 4 — Comparative Review of the Stop Sounds

In an explanation of the stop sounds, (**p–b, t–d, k–g**), it is important to compare clearly the similarities and differences in their visible speech patterns. The general rule that the patterns portray the way the sounds are made should be remembered, and, in this connection, it is helpful to recall how the breath stream is modulated in producing the different sounds. The stop sounds have been presented in pairs, **p–b**, **t–d**, and **k–g**, because the two consonants in each pair are made from similar articulatory positions. However, attention also has been called to the fact that the **b**, **d**, and **g** are made with vocal cord modulation, while the **p**, **t**, and **k** are made without vocal cord modulation. The presence or absence of vocal cord modulation results in changes in the patterns by which the **p** can be differentiated easily from the **b**, the **t** from the **d**, and the **k** from the **g**. The differences between the voiced and voiceless sounds have been explained in some detail in Units 1, 2, and 3. Since each pair of



sounds, **p–b**, **t–d**, and **k–g**, is made from different articulatory positions, their patterns provide certain clues which are usually characteristic of the sounds and by which each pair may be recognized. The most important of these clues are summarized here in both chart and verbal form for comparison. The six figures on this page and the preceding present patterns of each of the stops as it is said before and after the vowels **i** (*eve*) and **u** (*boot*). The center rows of the groups of voiceless and of voiced stops contain combinations of the **p**, **t**, **b**, and **d** with vowels which they influence only slightly, and therefore vowels whose hubs somewhat approximate the hubs of the stops with which they are said. Since the **k** and **g** are variable sounds, the high or front **k–g** influences the **i** (*eve*) only slightly, and the low or back **k–g** influences the **u** (*boot*) just a little, as may be seen on the chart. Since there is no other vowel on which their influence appears to be less marked than on the **i** or **u**, the stressed neutral vowel ʌ (*up*) has been selected for the central row of **k** and **g** combinations.



Each pair of stop sounds has a different hub position

1. The **p**-**b** hub is towards the bottom of the pattern.
2. The **t**-**d** hub is near the middle of the pattern.
3. The **k**-**g** hub is towards the top of the pattern with the front vowels, towards the middle with the mid vowels, and towards the bottom with the back vowels.

Each pair of stop sounds has a different spike fill

1. The **p**–**b** spike fill is darker towards the bottom of the pattern.
2. The **t** **d** spike fill is darker towards the top of the pattern.
3. The **k** -**g** spike fill is darker towards the top of the pattern for the front vowels, towards the middle for the mid vowels, and towards the bottom for the back vowels.

Each pair of stop sounds influences the vowels differently

1. The **p** and **b** influence bars 2 and 3 of the front vowels in the **i** (*eve*), **ɪ** (*it*), **eɪ** (*pay*), **ɛ** (*met*) the most.
2. The **t** and **d** influence bar 2 of the back vowels in the **u** (*boot*), **ʊ** (*put*), **oʊ** (*go*) the most.
3. The **k** and **g** change all the vowels for the three-bar voice by curving bars 2 and 3 together. For the four-bar voice, the **k**-**g** usually curves bars 3 and 4 together in **i** (*eve*), **ɪ** (*it*), and **ɛ** (*met*), but curves bars 2 and 3 together for the other vowels. The **k** and **g** influence the front vowels more noticeably than they do the back vowels.

It is expected that the extent of the review will vary with the degree of interest of the reader in learning to recognize visible speech patterns. Most readers will find it helpful, in associating the principles and patterns of the stop consonants, to review the material in Units 1, 2, and 3 in this chapter, keeping in mind the characteristic clues by which the sounds are recognized and differentiated. It is hoped that many readers will find it interesting to visualize the patterns of sounds and words before looking at them again in the illustrations. In constructing a mental image of a word pattern, think of how the spike fills, the gaps, and the vowel bars of the sounds will appear in combination. A knowledge of the location of the hubs of the adjacent sounds makes it possible to visualize the transitional patterns between the hubs.

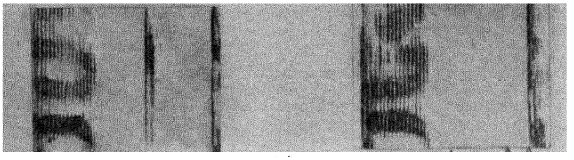
The following suggestions are offered for those readers who are interested in learning to read visible speech within the limitations of this book.

In addition to constructing a mental image of the word patterns it is advisable to practice recording the images on paper. After a little practice it should be possible to visualize and duplicate in graphic form most of the word patterns used in the text. The manual symbols presented in Chapter 6, or actual drawings of the patterns, may be used to represent graphically the word and sentence patterns. Another helpful exercise is to visualize and reproduce graphically the list of additional words included in Units 1, 2, and 3. These word lists provide a means by which the principles, which are explained and illustrated in each Unit, may be applied. The accuracy of the graphic representations may be determined by comparing them with the patterns of the same

sounds and sound combinations as they appear in the illustrated patterns of other words. The use of visualization and of some form of graphic symbolization should be continued through the following chapters by those readers who are interested in learning to read visible speech patterns.

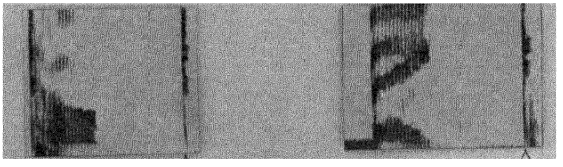
Unit 5 — Reading Combinations of the Stop Sounds

Combinations of the stop sounds occur frequently both within words such as *act*, and between words such as *make good.* Reading the patterns made by these combinations is contingent, of course, upon the recognition of the patterns and influences of the individual sounds. However, there are additional characteristics of the combined sounds that require particular explanation.



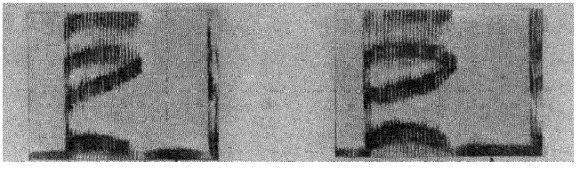
act ***act***

When two stop sounds are combined, as in the word *act*, sometimes both stop sounds are exploded and a spike fill appears on the pattern for each sound. Such a pattern is shown in the illustration at the left above. Frequently the first stop sound is not exploded. In that case, a stop gap is present for the first sound ʌ, but there is no spike fill. The stop gap for the second sound ʌ immediately follows the first, producing a double stop gap. The spike fill for the second stop sound ʌ usually appears on the pattern. A double stop gap is read as two sounds, and the first stop sound may be identified by its influence on the vowel. The second stop sound may be recognized by its pattern, by its influence on the sound that follows it, and in part by the context.



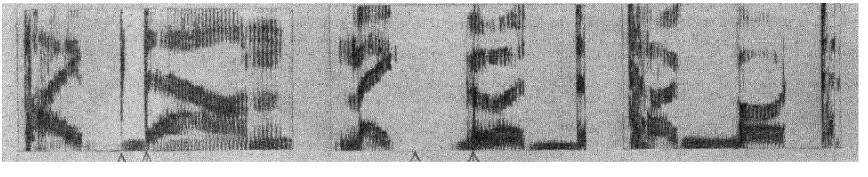
talked ***baked***

If the first stop sound is a voiceless consonant, **p, t,** or **k,** the second stop sound of the combination is almost always a voiceless sound. Even if the second sound of the combination is spelled as a voiced stop like ***d*** or ***ed,*** it is said as a **t.** This is a common occurrence in the past tense of such verbs as *walk, look, pack,* and *ask.* In the patterns of the words *talked* and *baked,* the **k–t** combination is illustrated. Note the double stop gap followed by the spike fill for the **t.** The influence of the imploded **k** is apparent in the characteristic rounding of the vowel bars in each word, and both the nature of the spike fill ∧ and the context indicate that the exploded stop was **t.**



begged ***bagged***

When the first stop sound of the combination is a voiced consonant, **b, d,** or **g,** the second stop is usually said as a voiced sound, too. In the words *begged* and *bagged,* the patterns of the **g–d** combination may be observed. The voice bar ∧ continues through the double stop gap. The influence of the imploded **g** is clearly evident in the rounding together of vowel bars 2 and 3. The voiced spike fill for **d** may be identified in part by its pattern and in part by the context.



pipe down ***make good*** ***tug boat***

In continuous speech, combinations of stop sounds often occur between the final stop sound of one word and the initial stop sound of the next word. Several frequently used combinations are illustrated in the patterns above, but there are numerous other possible combinations. The words *pipe down* show spike fills for both the **p** ∧ and the **d** ∧. The **p** spike is produced when the lips open to make the **d.** The voice bar for **d** begins immediately after the **p** spike. When both spike fills are present, the combination of two words may be read in the same way that each word would be read by itself.

In the words *make good*, the **k** is imploded and the **g** is exploded. The stop gap for the **k** ⩓ precedes the voiced stop gap and voiced spike fill ⩓ of the **g**. The **k** influence is apparent on bars 2 and 3 of the vowel in *make*. The **g** spike fill may be recognized just as it would be if the word *good* were spoken alone.

The pattern made by the **g–b** combination in *tug boat* shows that the **g** was not exploded. The voice bar continues through the double stop gap. The imploded **g** may be identified by its influence on the vowel in *tug*. The **b** is easily recognized both by its pattern and its influence on the vowel in *boat*.

The clues that indicate the presence of a combination of two stop sounds are:

1. The double stop gap followed by one spike fill.
2. The two spike fills present in some combinations.
3. The influence on the vowel of the first sound in the combination.

The clues that identify the particular stop sounds combined are:

1. The voiceless or voiced stop gap.
2. The appearance of the voiceless or voiced spike fills.
3. The influence on the vowels.

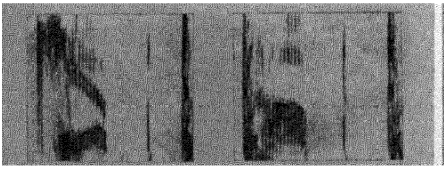
In patterns that contain combinations of stop sounds certain observations should be made:

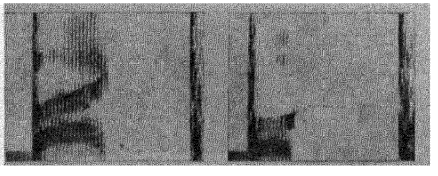
1. Do the stop gaps indicate one, two, or three stop sounds?
2. Are the stop gaps voiced, voiceless, or a combination of voiced and voiceless?
3. How many spike fills are there?
4. Which stop sounds are imploded and which are exploded?
5. Are the spike fills voiced or voiceless?
6. What is indicated by the influences on the vowels?
7. Which stop sounds have been combined?

Review words

baked, talked, act, page 104 ***bagged, begged***, page 105

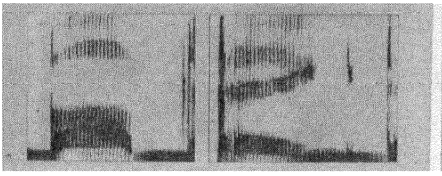
New words to learn



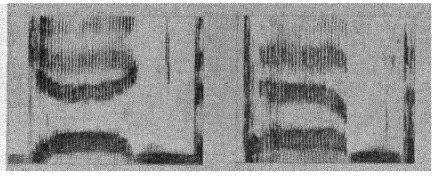


kept ***topped*** ***backed*** ***booked***

Reading combinations of the stop consonants



bobbed *pegged*

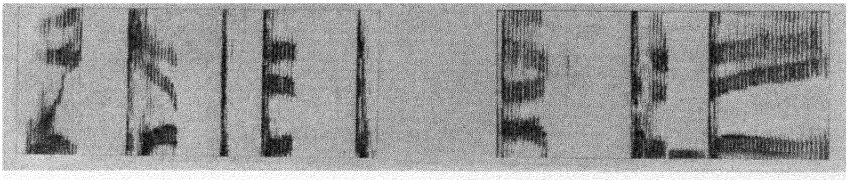


gagged *tabbed*

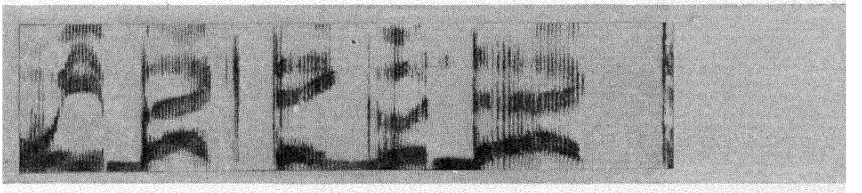
Additional words for practice in visualization

picked *kicked* *packed* *cooked* *piped* *typed*

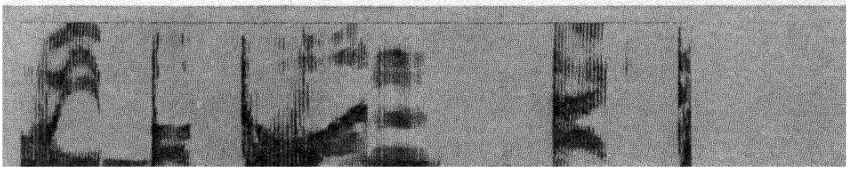
Drill phrases and sentences



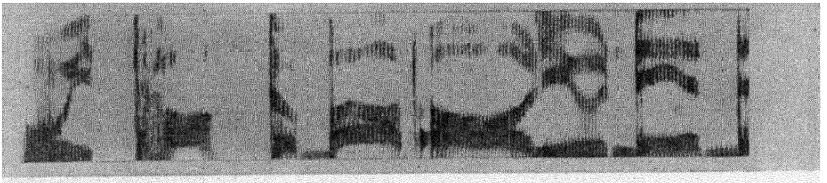
We kept it *Act today*



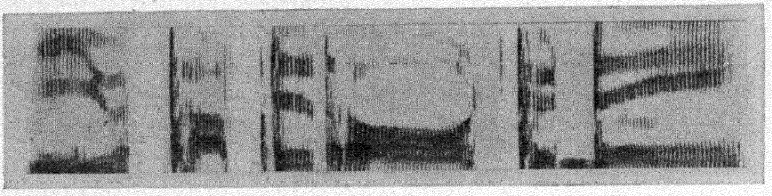
We backed a good act



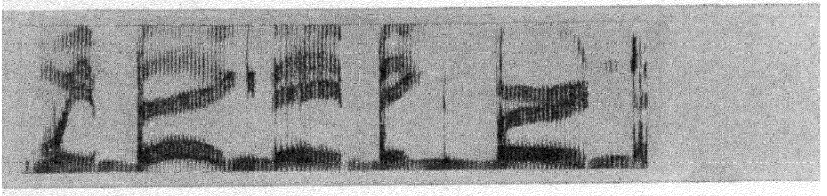
We booked one act



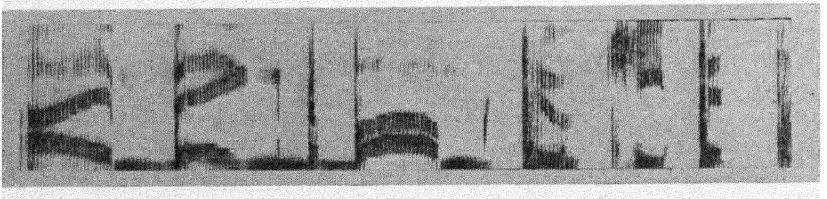
We talked about all you did



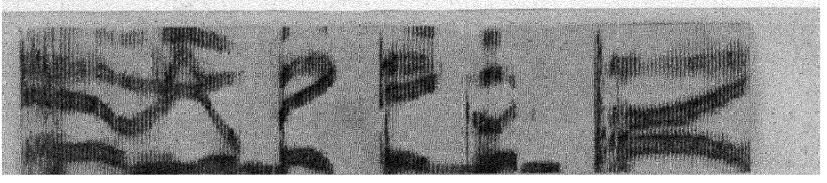
You topped it all today



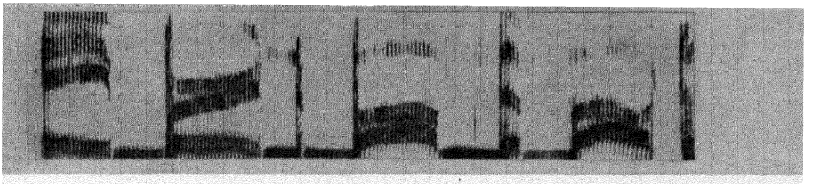
We bagged a big bird



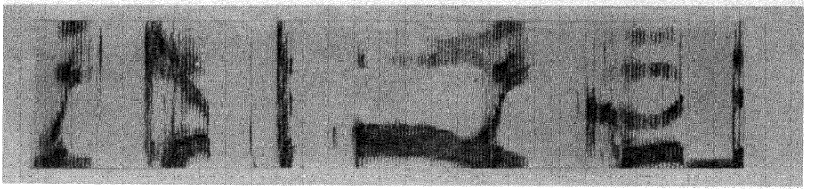
I begged Bob to keep it



Have you baked a good pie?



A bird bobbed about



We kept all we could

Chapter 9

Reading the Fricative Sounds

This chapter presents a discussion of the characteristic patterns by which the fricative sounds are read. There are ten sounds included in the fricative classification: h (*he*), ɦ (*ahead*), f (*for*), v (*vote*), θ (*thin*), ð (*then*), s (*see*), z (*zoo*), ʃ (*she*), ʒ (*azure*). Patterns of this general sound group are recognized by two primary clues:

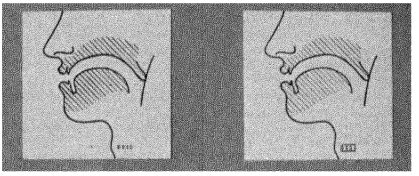
1. Voiceless fills produced by the continuous emission of breath through a small or restricted opening which may occur at one of several places in the articulatory or vocal tract.
2. Voiced fills or combinations of resonance bars and fills produced by the continuous emission of voiced breath through an opening as described above.

The voiceless fricatives are produced primarily by frictional modulation; the voiced fricatives by combined frictional, vocal cord and cavity modulation. Unlike the stop sounds, the breath for the fricatives may be emitted slowly and continuously, and, therefore, the fricative fills usually are wider (longer in duration), than the spike fills for the stop sounds. Frequently, there is some resonation of the voiceless breath in cavities that are formed in the front part of the mouth during the production of the fricatives. When such resonation occurs, weak upper resonance bars may be visible in the voiceless fills. Strong voicing of the voiced fricatives occasionally increases the voiced element in proportion to the fricative element to such an extent, that the resonance bars become much more prominent than the fills and, therefore, the patterns are somewhat vowel-like in appearance.

The voiced and voiceless member of each pair of fricative sounds is considered in greater detail in the following units of this chapter. As in the stop chapter, the transitional influence of each of the fricatives is illustrated by words containing a front vowel, a back vowel, and a vowel towards the middle of the series. The combined text and illustrations should enable the reader to visualize and to understand the basic principles underlying the influences of the various fricatives on all of the vowels and diphthongs. In addition, Unit 6 contains a summary and comparison of the patterns, hubs, and influences by which the voiceless and voiced fricatives may be recognized and differentiated.

Unit 1 — Reading the h (he) and ɦ (ahead)

How the h and ɦ are made



***Profile of* h (*he*)** ***Profile of* ɦ (*ahead*)**

The h and ɦ are produced by:

1. Forming a small or restricted opening either between the vocal cords or at some point in the region of the glottis.
2. Closing the opening between the throat (pharyngeal) and the nasal cavities.
3. Emitting a continuous voiceless or voiced breath stream through the restricted opening.

In producing the h, the articulators generally assume the position for the sound which precedes or follows it, and therefore there is no special articulatory position for the h. As is indicated by the dash lines in the profile above, the h is believed to be produced through a fricative position of the vocal folds or vocal cords, as they pass from the open or breathing position to the closed or vibratory position. The vocal cords are considered to be approximated sufficiently to produce frictional but not vibratory modulation of the breath stream. The h which precedes a vowel is usually a voiceless sound, and it is thought by some phoneticians to be a way of initiating or terminating a vowel rather than a sound entity. Phonetically, the h is described as a glottal fricative. Since the meaning of words is affected by the presence of h, as in *ate* and *hate*, *it* and *hit*, *eat* and *heat;* and since, in visible speech, the h has a definite pattern and a hub; there does not appear to be any question of its inclusion as a separate pattern unit.

The voiced h or ɦ presents a somewhat different problem. When the h occurs between two voiced sounds in continuous speech, and especially in rapid speech, usually the h is said as a voiced sound. Although the ɦ visible speech pattern indicates that it is as definitely the voiced cognate of the voiceless h as the b is of the p, or the z is of the s, in American speech the voicing of h does not affect the meaning of the word or phrase in which it occurs. Nor does the ɦ occur invariably in certain words or phrases, but rather the voicing is dependent upon stressing, rate of speech, and the nature of the adjacent sounds. The classification of sounds that is used in this volume was developed primarily upon the basis of visible sound patterns, and therefore the

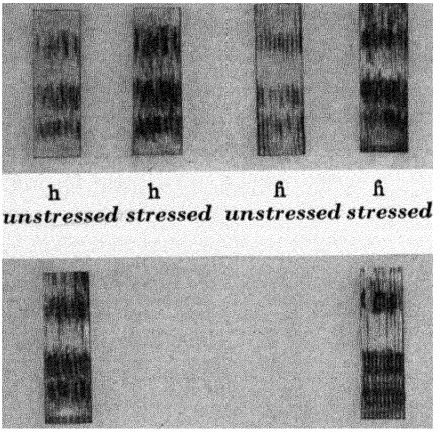
voiced **h,** symbolized by the **ɦ**, is included as a separate unit in the visible speech alphabet. However there is not sufficient evidence concerning the manner of articulation of the **ɦ** to warrant a more complete statement at this time. The **ɦ** profile indicates frictional modulation by the dotted lines, and vibratory modulation by the solid lines.

Frequently a sound is terminated by voiceless or voiced breath that may or may not resemble the **h** in form, but since such a termination does not change the meaning it is of concern in reading visible speech only as a non-information-bearing pattern. Presumably it is produced when the breath is not stopped as the vocal folds pass through a fricative position while moving from the vibratory to the open or breathing position.

The h and ɦ are spelled in the following ways

h (he) *wh (who)* *h (ahead)*

Visible patterns of the h and ɦ



ɦ
frictional modulation predominant

ɦ
vocal cord-cavity modulation predominant

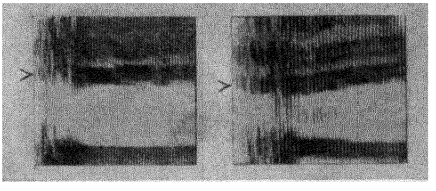
The glottal frictional modulation of the breath stream for the **h** produces the irregular, vertical striations characteristic of the **h** fill. Since the fricatives, unlike the stops, may be continued, they may be said and their patterns produced in isolation. The **h** shown above was spoken before the stressed neutral vowel ʌ as in *hut.* The **h** may be varied from one which is just perceptible to one that is considerably overstressed,

depending primarily upon the amount of air-pressure built up below the glottis and the speed with which the vocal folds move through the fricative to the vibratory position. The pattern at the left illustrates a weak or light **h** fill with the striations spaced rather widely, not very dark, and confined to the areas of the vowel resonances which follow. The next pattern presents a stronger or more stressed **h**. The vertical striations of the fill are spaced more closely, they are darker, and they are spread over a larger vertical area, with the resonance areas not as clearly defined as in the lighter pattern.

When the **h** is voiced, as it frequently is in the word *ahead*, it may be stressed or unstressed. The patterns above illustrate the decrease in darkness of the bars characteristic of unstressing, as well as the complete or partial absence of some of the resonances. Vocal-cord modulation produces the voice bar along the baseline in both the stressed and unstressed sound patterns. In the **ɦ** the voiced element may be weak or strong in relation to the fricative element. As the voicing becomes stronger, the vocal resonances become more apparent and the fricative fill less evident. The right hand pattern illustrates the resonance areas of a rather strongly voiced **ɦ**.

Although the **h** and the **ɦ** have the same hub in the same combinations, neither has a single hub position. One of the chief characteristics of these two variable sounds is that the hub of the **h** and **ɦ** appears to be in approximately the positions of the hubs of the sounds with which they are said. Also, the dark areas of the voiceless **h** fill and the resonances of the voiced **ɦ** fill are usually in the positions of the vocal resonances of the sound with which the **h** or **ɦ** is spoken. When the **h** occurs between two voiced sounds, the dark fill areas or the resonances portray the transitional movement between the two sounds. Since the **ɦ** does not occur initially in American words, and occurs medially in so few common words, and furthermore since the **h** and **ɦ** influence the vowels in the same way, the illustrations of influence will be shown with the **h**. The section of drill phrases and sentences presents word patterns containing both the **h** and the **ɦ**.

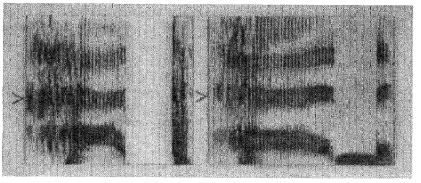
How the h influences vowel patterns



he ***hay, hey***

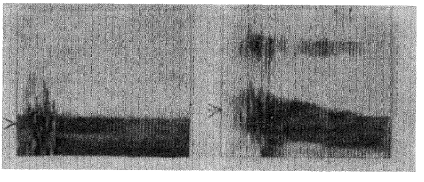
The patterns of *he* and *hay* show how little the **h** influences the vowels. Bar 2 of the vowel sometimes curves downwards very slightly from the **h** hub, but generally the

h hub ≻ takes the same position as the vowel hub. Horizontal separations between the dark fill or resonance areas frequently are present in the **h**. In general, bar 1 of the vowel following the **h** begins first, then bar 2, and finally, bar 3. The pattern of resonation thus portrayed is of considerable interest in phonetic analyses, but in reading it serves merely as an additional clue to the recognition of the **h**.



hat *had*

The fill areas of the **h** and the **h** hub ≻ take the position of the vowel **æ** (*at*) in both *hat* and *had*. Although the **h** does not have an influence on the position of the vowel bars, it does appear to take a bite out of bars 2 and 3. Notice that the vowel is shorter before the voiceless than before the voiced stop.



who *how*

The **h** hub in *who* ≻ takes the position of the hub of the vowel **u** (*boot*) since the word *who* is said and is transcribed phonetically as **hu**. Either the **ɑ** (*father*) or **a** (*ask*) may be used to begin the diphthong **ɑʊ** (*out*) or **aʊ** (*out*). Both sounds have hubs higher than the **u** hub, and in the word *how* the hub of the **h** ≻ spoken before this diphthong is raised correspondingly.

A comparison of the **h** spoken with the **i** and the **u** indicates the range of the **h** hub and also the change in the appearance of the **h** fill. The **h** and its voiced equivalent, the **ɦ**, are variable sounds and have a variable hub. Generally, the **h**–**ɦ** hub approximates the hub of the vowel with which it is combined. When the **h** or **ɦ** occurs between two sounds, in continuous speech, its hub and pattern are influenced by both the preceding and the following sounds.

The visible patterns of the ɦ *differ from those of the* h *chiefly in that:*

1. The ɦ is voiced and its fill has a voice bar along the baseline.
2. The ɦ fill usually contains several resonance bars.
3. The ɦ bars are joined to the vowel bars.

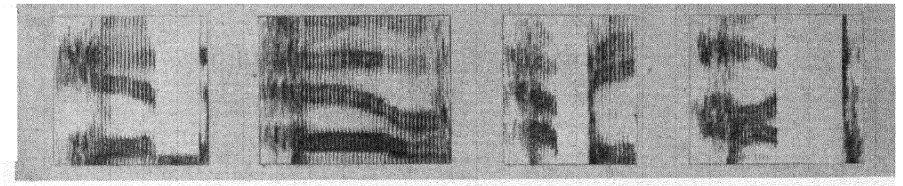
Read the h *and* ɦ *by:*

1. The horizontal separations between the dark fill areas.
2. The progressive initiation of voicing after the h.
3. The variable nature of the h-ɦ hub and pattern.

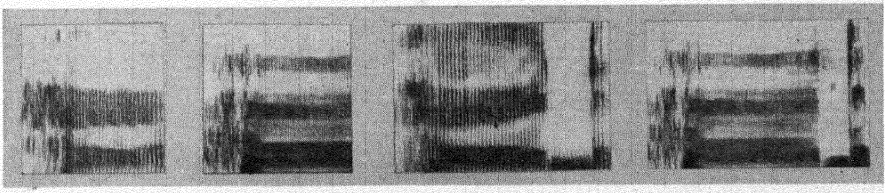
Review words

he, page 113	***who***, page 114	***hay***, ***hey***, page 113
hat, page 114	***how***, page 114	***had***, page 114

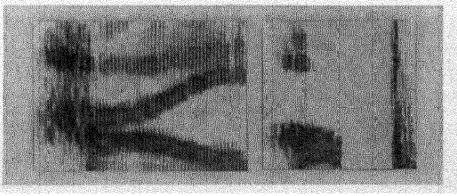
New words to learn



head ***have*** ***happy*** ***hot***



her ***(General American)*** ***her*** ***(Eastern)*** ***heard*** ***(General American)*** ***heard*** ***(Eastern)***

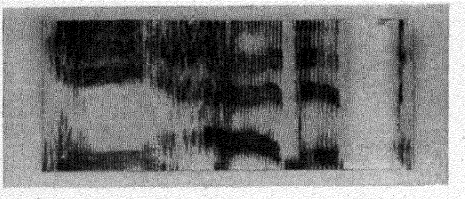


high ***hope***

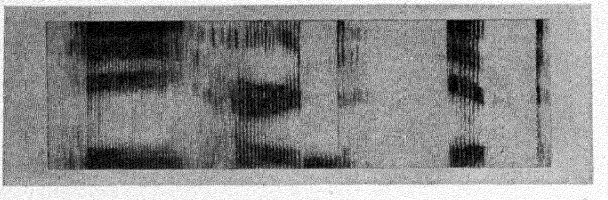
Additional words for practice in visualization

heat ***hurt*** ***height*** ***hide***

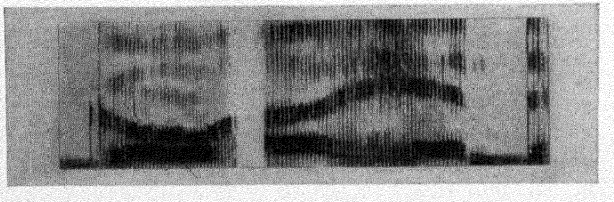
Drill phrases and sentences



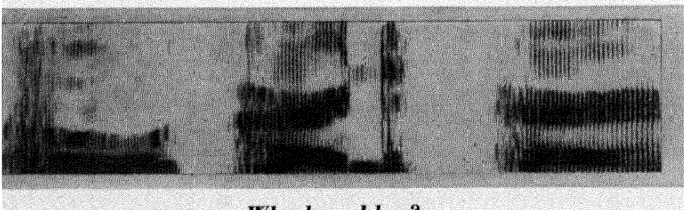
He had it



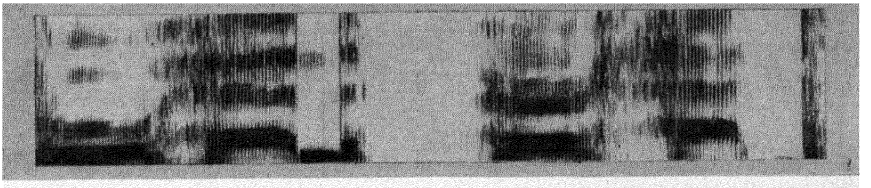
He heard it



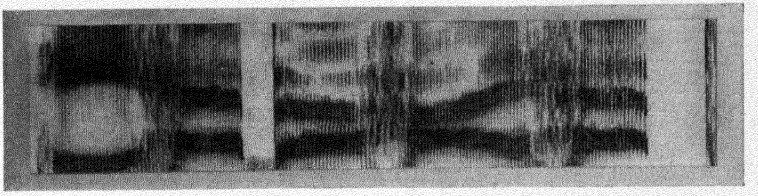
Go ahead



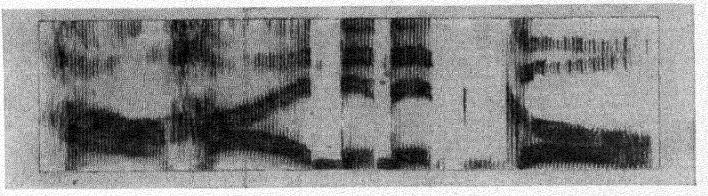
Who heard her?



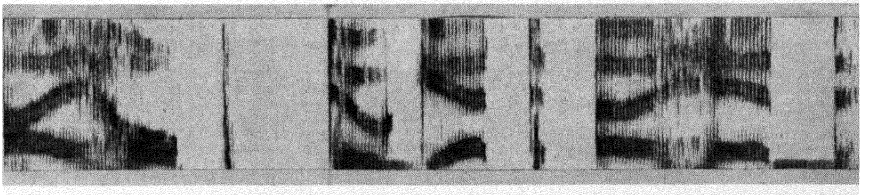
Who had her hat?



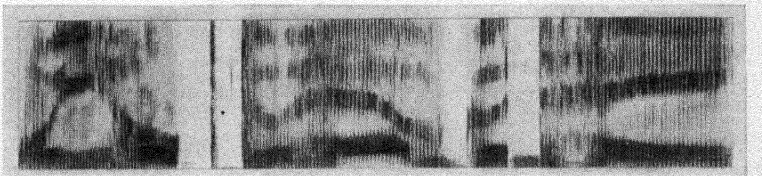
He had a high hat



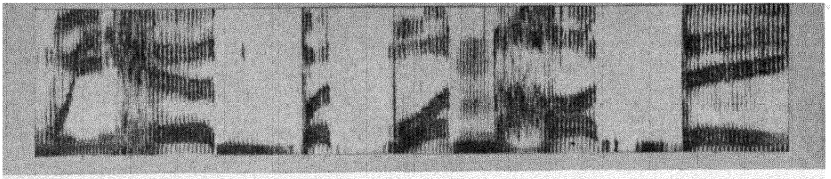
How high did it go?



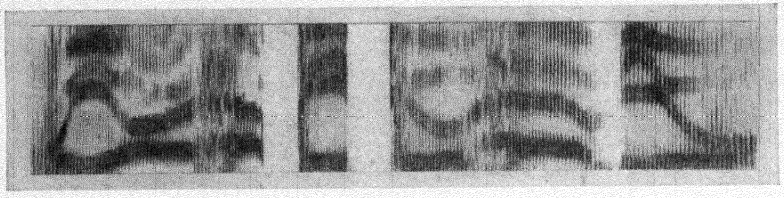
I hope to get ahead



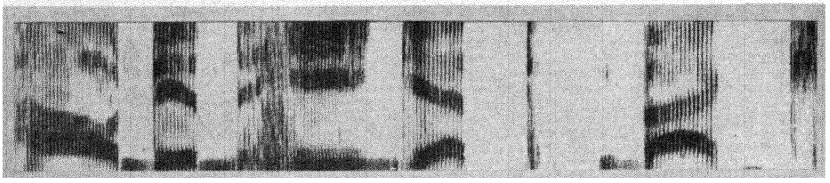
We hope to have good hay



We had but one hot day



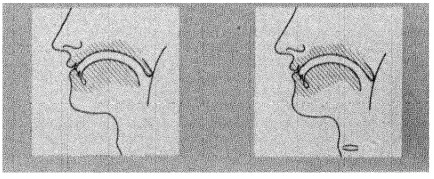
We are happy to have you



How did he get back?

Unit 2 — Reading the f (for) and v (vote)

How the f and v are made



Profile of f (*for*) *Profile of* v (*vote*)

The f and v are produced by:

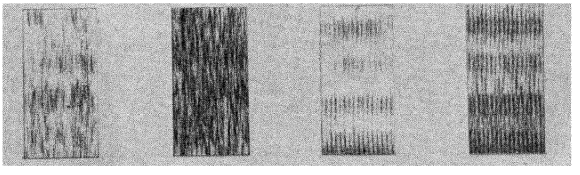
1. Forming a small or restricted opening between the lower lip and the upper teeth.
2. Closing the opening between the throat (pharyngeal) and the nasal cavities.
3. Emitting a continuous stream of voiceless or voiced breath through the restricted opening.

As indicated in the profiles above, the articulators are in similar positions for the f and v, but the sounds are modulated differently. Although both sounds are produced with frictional modulation, the v combines frictional with vocal cord and cavity modulation. Since the f and v are cognates, the f is referred to as a voiceless fricative continuant and the v as its voiced equivalent. In physiologic phonetics, the f is described as a voiceless labio-dental fricative, and the v as a voiced labio-dental fricative.

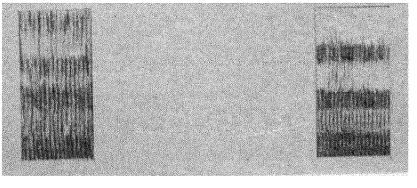
The f and v may be spelled in different ways

f (*feel*)	*ff* (*off*)	*fe* (*knife*)	*v* (*van*)	*ve* (*wave*)
ft (*often*)	*gh* (*laugh*)	*ph* (*triumph*)	*f* (*of*)	*lve* (*halves*)
lf (*half*)				

Visible patterns of the f and v



f *unstressed* | f *stressed* | v *unstressed* | v *stressed*

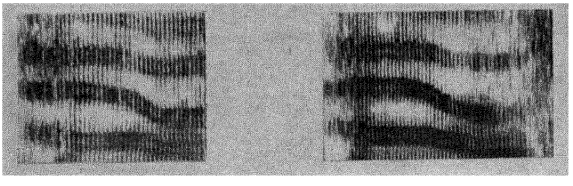


v *frictional modulation predominant* | v *vocal cord-cavity modulation predominant*

As the breath stream is emitted through the lip-teeth opening for the f, the frictional modulation of the breath produces the irregular vertical striations of the fill. The fricatives may be continued at will, and the horizontal length of the fill portrays

the duration of the sound. When the **f** is weak or unstressed the striations are rather widely spaced, vertically shorter, and light, but they are more numerous and darker when the **f** is stressed. Usually there are small dark areas randomly scattered in the **f** fill, and the fill frequently appears darkest towards the baseline.

The vocal cord and partially closed cavity modulation for the **v** produces a voice bar along the baseline of the fricative fill. As the voicing for the **v** becomes stronger in relation to the fricative element, the vocal resonances of the **v** become more prominent. There may be one, two, or three **v** bars present in the pattern, dependent upon the proportion of the voicing and fricative elements.



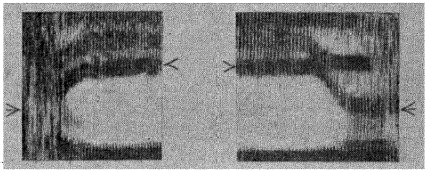
"have" —v voiced throughout ***"have" —v unvoiced at end***

During continuous speech, the voicing for the final **v** frequently is not maintained at a constant level throughout the sound. The pattern of the word *have* at the left, illustrates a constant proportion of vocal cord, cavity, and frictional modulation in the **v**. The **v** resonances are dark and prominent from the transition to the abrupt termination of the sound. In the right hand pattern of the same word, the beginning of *unvoicing* of the **v** is apparent. The **v** resonances are strong at the transition, but they become less evident, and the fricative fill becomes relatively prominent, as the sound continues towards a gradual termination of first cavity and vocal cord, and then frictional modulation.

Because the cognates **f** and **v** are produced from similar articulatory positions, they have the same hub. The hub is not often visible in the voiceless fricative, but it usually is evident in the second resonance bar position of the voiced fricative. The **f** and **v** are considered to be relatively fixed sounds, since the individual variation in the particular lip-teeth position used in different syllabic combinations is slight and does not produce sufficient modification of the hub or of the transitional patterns to be significant for reading.

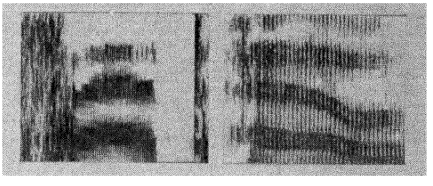
How the f and v influence vowel patterns

The patterns of *fee* and *eve* illustrate the influence of the **f** and **v** on the vowel **i** (*eve*). The hub of the **i** $<$, $>$, is towards the top of the pattern; and the hub of the cognates **f** $\geq$ and **v** $\leq$ is towards the bottom of the pattern. Since the cognates **f** and **v** have



fee *eve*

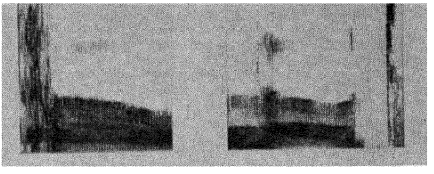
the same hub, the direction of their influence on any vowel is similar although the nature of their transitional patterns is different. Because vocal cord-cavity modulation continues throughout the pattern of *eve*, the transition as the resonance bars of the vowel are joined to the **v** bars is apparent. Bar 1 of the **i** (*eve*) is narrow and is on the baseline, and therefore the shift from open cavity modulation for the vowel to partially closed cavity modulation for the **v** is not very evident (only in the narrowing of the first bar during the transition). The **f** hub is not visible in the fill, but if the curve of bar 2 from the **f** to the **i** is compared with that of the **i** to the **v**, the similarity in hub may be observed. It is interesting to note that bar 3 of the **i** also is curved by the **f–v** hub although not as markedly as bar 2. The influence of the **f–v** hub on bar 3 of the front vowels is helpful in differentiating the **f** and **v** from the θ, ð, **s**, and **z**.



fat *have*

The hub of the vowel æ (*at*) is near the center of the pattern, but the **f–v** hub is towards the baseline. Bar 2 of the vowel curves upwards from the **f** hub in *fat*, and curves downwards towards the **v** hub in *have*. The shift from open cavity to partially closed cavity modulation may be observed in the transition between bar 1 of the vowel and bar 1 of the **v** in *have*. Bar 3 of this vowel is not influenced as much by the **f–v** hub as it was in the **i** in *eve* and *fee*. The relative proportion of friction and voicing in the **v** patterns in these words and in the other illustrations of the **v**, may be compared. Occasionally, the voiceless fricatives are made with sufficient cavity modulation to produce shadows of resonance bars on the fill pattern. If the fills for the various **f** illustrations are compared, both the evidence of cavity modulation and the relative intensity of the fricative element as manifested in the appearance of the striations,

may be observed. The difference between the variable **h** hub and the relatively fixed **v** hub is apparent in the direction of the transitions between the vowel **æ** and the two fricatives in *have.* In addition, the change in type of modulation, from voiceless fricative to vowel to voiced fricative in *have,* should be associated with the appearance of the transitional patterns.



foe ***vote***

Since the **f–v** hub is slightly below the initial hub of the **ou** as in *foe* and *vote,* it does not have much influence on the vowel bars. However, some rounding or curving upwards of bar 2 from the **f–v** hub is apparent in the transitions. The change in type of modulation during the transition from the voiced fricative **v** to the vowel is most evident in the disappearance of vertical fill striations and in the widening of bar 1 in the vowel pattern. The voiceless-fricative-**f**-to-vowel transition illustrates, especially in bar 1, the beginning of vocal cord modulation during the change from the partially closed cavity modulation for **f** to the open cavity modulation for the vowel. Bars 3 and 4 of the back vowels may or may not be present in these patterns and therefore we do not attach much "reading-significance" to their transitions.

The **f** and **v** have a relatively fixed hub which is located towards the baseline, near the hub of the vowel **u** (*foot*). Because the relative distance between hubs is more marked, the **f–v** influences bar 2 of the front vowels more than the back vowels, with a gradual decrease in the extent of influence as the vowels are said in series from **i** (*eve*) to **u** (*foot*). With all the vowels except the **u** (*food*), the direction of the transition is the same, upwards from the **f–v** to the vowel or downwards from the vowel to the **f–v**; it is the degree of curvature or the extent of transition that decreases when the hubs are closer together or increases when they are farther apart. In combination with the **u** (*food*) the transition of bar 2 is downwards from the **f–v** hub towards the **u**, or upwards to the **f–v** hub from the **u**, since the **u** hub is slightly lower than the **f–v** hub. Bar 3 of the vowels is influenced like bar 2 in the front vowels, but in the back vowels the influence is so slight that for reading it is considered non-significant. The influence of all of the fricatives on bar 1 is similar, since it is determined primarily by the distance between the voice bar of the voiced fricatives and bar 1 of the vowel, and by the speed of initiation of voicing during the transition between the voiceless fricatives and the vowel.

The visible patterns of the **v** *differ from those of the* **f** *chiefly in that:*

1. The **v** is voiced and its fill has a voice bar along the baseline.
2. The **v** fill usually contains several resonance bars.
3. The **v** bars are continuous with the vowel bars.

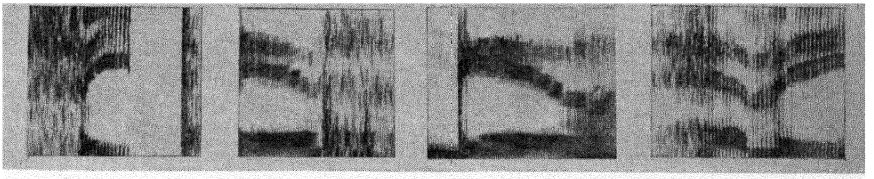
Read the **f** *and* **v** *by:*

1. The appearance of the striations of the fill.
2. The appearance of the voiced fill.
3. The influence on the vowels.

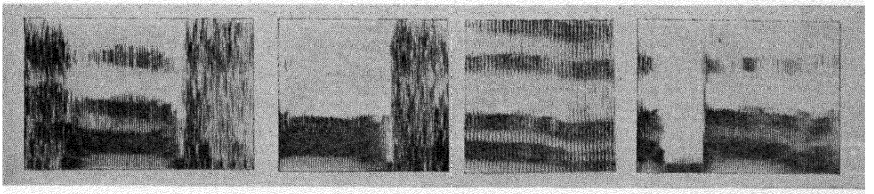
Review words

fee, page 121 *foe*, page 122 *eve*, page 121 *vote*, page 122
fat, page 121 *have*, page 121

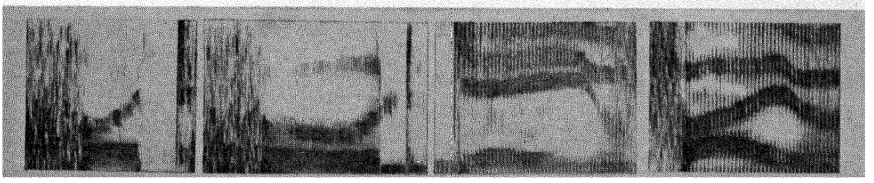
New words to learn



feet *if* *give* *heavy*



half *off* *of* *above*

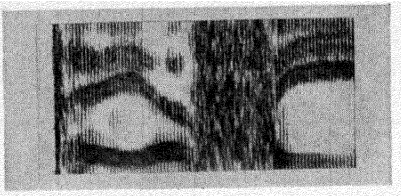


foot *food* *gave* *five*

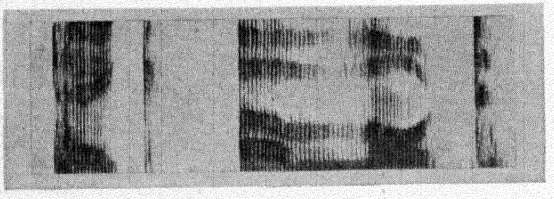
Additional words for practice in visualization

feed *perfect* *fight* *fit* *fact* *divide* *vow* *avoid*

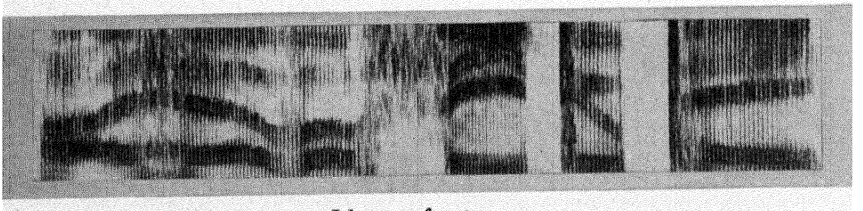
Drill phrases and sentences



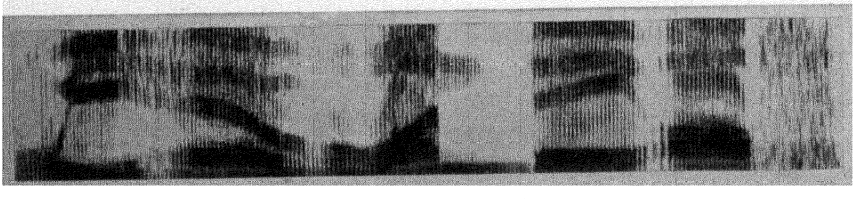
Pay a fee



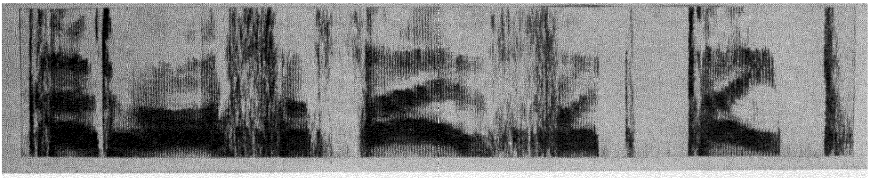
Take a vote



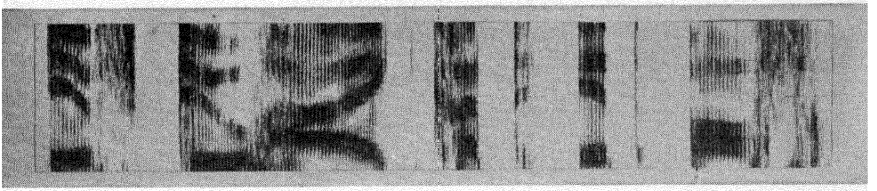
I have a fee to pay



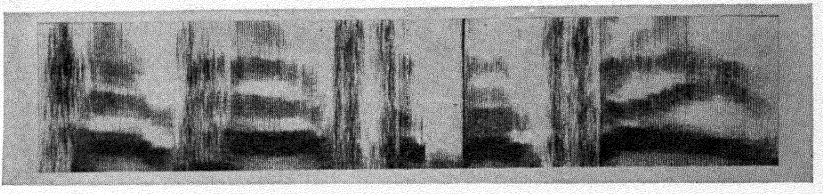
We have one day off



Cut off a five-foot pipe



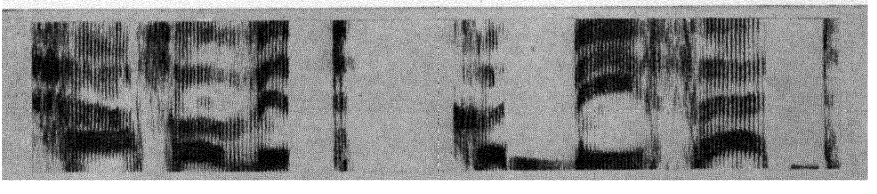
If too high, cut it off



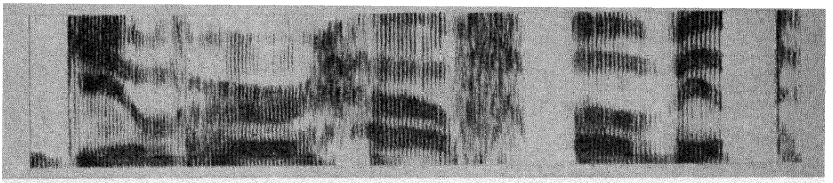
Have half above five



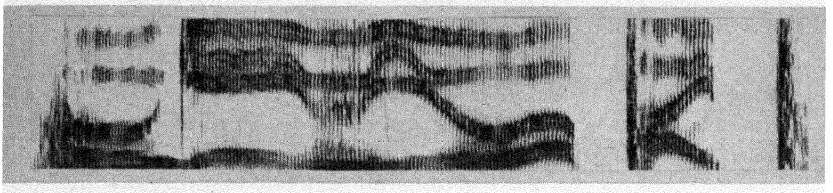
Have a heavy bag of food



Half of it could be fat



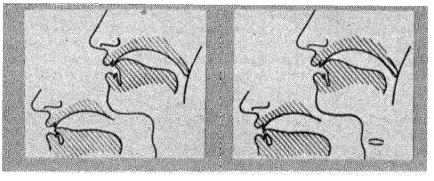
Give her half of it



Who gave you a pipe?

Unit 3 — Reading the θ (thin) and ð (then)

How the θ and ð are made



***Profile of* θ (*thin*)** ***Profile of* ð (*then*)**

The profiles above indicate the positions of the articulators used by many speakers in producing the θ and ð sounds. As illustrated, speakers may place the tip of the tongue either in light contact with the inner surface of the teeth or between the teeth. Regardless of the position used, the θ and ð are produced by:

1. Forming a small or restricted opening between the tip of the tongue and the upper teeth.
2. Closing the opening between the throat (pharyngeal) and nasal cavities.
3. Emitting the voiceless or voiced breath stream through the restricted opening.

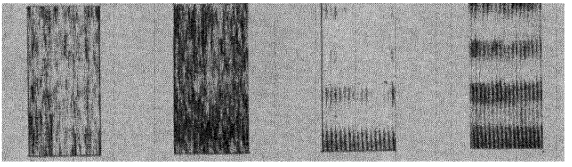
The frictional modulation thus produced is characteristic of both the θ (*thin*) and the ð (*then*) sounds. The ð differs from θ in that it combines vocal cord and cavity modulation with frictional modulation. Because of the difference in the modulation of the

sounds, θ (*thin*) is described phonetically as a voiceless lingua-dental fricative continuant, ð (*then*) as a voiced lingua-dental fricative continuant.

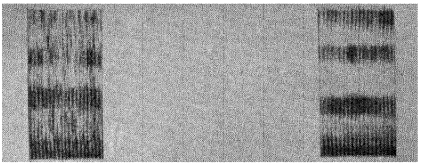
The θ and ð may be spelled in the following ways

th (*thin*) *th* (*then*) *the* (*breathe*)

Visible patterns of the θ and ð



θ ***unstressed*** θ ***stressed*** ð ***unstressed*** ð ***stressed***



ð ***frictional modulation predominant*** ð ***vocal cord-cavity modulation predominant***

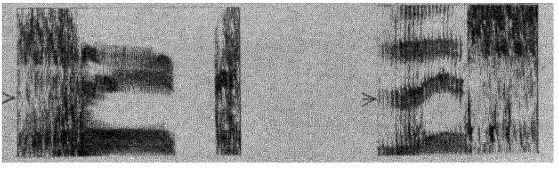
Since the θ and the ð, like the **f** and **v**, are continuant sounds, they can be spoken by themselves and can be continued at will. The voiceless breath stream that flows through the restricted tongue-teeth opening for the θ produces frictional modulation portrayed by a fill as shown in the stressed and unstressed patterns above. If the flow of voiceless breath is weak, as it is when the sound is unstressed, the vertical striations of the θ fill will be light, less numerous, shorter vertically, or missing, as is apparent in a comparison of the stressed and unstressed patterns above. When the breath flow is strong in the stressed sound, the fill striations are darker and for the θ they usually appear darkest towards the bottom of the pattern. Occasionally some partially-closed cavity modulation is combined with the frictional and results in shadows of resonance bars visible through the fill.

The continuous voiced breath stream flowing through the restricted tongue-teeth opening for the ð results in combined frictional-vocal cord-cavity modulation portrayed

by a voiced fill. The bars in the ð pattern are produced by the vocal cord-partially closed cavity modulation; the vertical striations of the fill by the frictional modulation. In the unstressed ð sound, the bars and fill are lighter in appearance than in the stressed, and some of the bars may not be visible. The number of bars is related to the amount of stressing and also to the manner in which the different kinds of modulation are proportioned. When vocal cord-cavity modulation predominates, the resonance bars are prominent and the fill is not very evident. When frictional modulation predominates, the voicing bar may be the only resonance present and the fill striations are usually quite dark and numerous. There are an infinite number of possible combinations of these types of modulation; and frequently a particular voiced fricative will begin with one combination and shift into another, depending upon the sounds which precede and follow it.

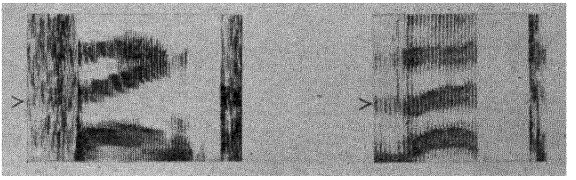
The θ and ð are produced from similar articulatory positions and are cognates; therefore, their hubs are similar. Bar 2 of the ð indicates the position of the hub for the θ as well as the ð. Although there is some variation in the tongue-teeth position used for the θ and ð, it is not sufficient to modify the hub or the transitional patterns significantly for legibility in reading.

How the θ and ð influence vowel patterns



thick *this*

The influence of θ and ð on the vowel ɪ (*it*) is illustrated in the patterns of the words *thick* and *this*. Both consonants have the same influence on the vowel as is shown by the direction of the transitional movement. Differences in the nature of the transitional pattern are related to the differences in the types of modulation combined. Since the θ fill has no bars in it, its hub position is not apparent in the pattern of the θ sound itself, but the beginning of bar 2 of the vowel indicates the position of the θ hub to be the same as that of the ð hub. The ð influence is more easily observed and recognized because vocal cord-cavity modulation is continued; and, therefore, the bars in the ð fill are joined with the bars of the vowel. Bar 2 of the vowel begins at the hub of the θ > and ð > and curves upwards, because the vowel hub is considerably higher than the θ–ð hub. The ð voice bar in *this* curves upwards and joins bar 1 of the vowel. Bar 3 of ɪ is not influenced very much by θ and ð.



thank *that*

The patterns of the words *thank* and *that* show the influence of the θ and ð on the vowel æ. Since the θ–ð hub is comparatively fixed, the reader who is able to visualize its position and the positions of the æ (*at*) bars will know how the sounds influence each other before looking at the illustrations. The beginning of bar 2 of the vowel curves up from the θ–ð hub >, but the curve is not as great as it is with ɪ (*it*) since the æ (*at*) hub is lower. The voice bar of ð in *that* joins with bar 1 of æ, and the joining curve is greater than it is in *this* since bar 1 of æ is farther from the baseline than is bar 1 of ɪ. Bar 3 of æ, like bar 3 of ɪ, is not influenced very much by the θ and ð sounds. The lack of influence of the θ–ð on bar 3 is helpful sometimes in distinguishing the θ–ð from the **f–v.**

both *though*

The θ and ð also influence the back vowels in a predictable manner. In the word *both* the ʊ (*foot*), which is the final sound of the diphthong oʊ (*go*), is adjacent to the θ. Because the hub of ʊ is lower than the θ hub <, bar 2 curves upwards from the ʊ towards the θ hub. Bars 1 and 3 of the vowel are influenced only slightly by the θ. In *though* the ð influence is on the vowel o (*obey*), the first element in the diphthong oʊ (*go*). Bar 2 of the o curves downwards from the ð hub to the o hub and then to the ʊ hub. The voice bar of the ð is joined with bar 1 of the vowel, and the transitional curve of bar 3 is evident also.

The θ–ð influence on all of the front vowels can be visualized and predicted from the information given concerning the ɪ (*it*) and æ (*at*) vowels. If the ð is combined with the i (*eve*) as in the word *the*, the ð voice bar joins with bar 1 of the i. Since both the voice bar and the first bar of the i are on the baseline, the transitional pattern shows only a slight shift in the width of the two bars due to a change in type of cavity-vocal

cord modulation. After the **i**, bar 1 of each of the front vowels said in series is farther from the baseline; therefore, the transitional curve increases progressively between the ð voice bar and vowel bar 1. Bar 2 of each of the front vowels also is joined with bar 2 of the ð, with the transitional curve most marked between the **i** (*eve*) and ð hubs. With each of the front vowels said in series from **i** (*eve*) to **a**(*ask*), the curve is decreased by the distance the second bar is lowered. Bar 2 or the hub of the vowel **a** (*ask*) coincides with the ð hub. Bar 3 of the front vowels usually is not influenced by the ð sufficiently to be significant for reading, except as the absence of transitional influence is helpful in differentiating the ð and **v**. The θ influence is similar to that of the ð in the direction of the transitional curve, but differs in the nature of the transitional pattern and in the amount of transitional curve apparent. The types of modulation combined determine the general appearance of the transitional pattern; the positions of the hubs combined determine the direction of the transitional pattern.

To predict the θ–ð influence on the back vowels, one should visualize the position of the vowel bars in relation to the θ–ð hub and the ð voice bar. Bar 1 for ɑ (*father*) is farther from the baseline than bar 1 of any other vowel. As the back vowels are said in series from ɑ (*father*) to **u** (*boot*), bar 1 is lowered progressively until in **u** (*boot*) it is on the baseline. Therefore the transitional curve decreases between the ð voice bar and bar 1 of each of the back vowels from ɑ to **u**. The transitional curve between the θ–ð hub and bar 2 of each of the back vowels varies with the distance between the two hubs, just as it did in the case of the front vowels. As mentioned previously, bar 3 for the back vowels is frequently absent. When bar 3 is present, its transitional patterns are of little significance in reading the θ and ð back vowel combinations.

The visible patterns of the ð *and* θ *differ chiefly in that:*

1. The ð is voiced; therefore, its pattern has a voice bar on the baseline and it may have one or more resonance bars.
2. The ð pattern, when the sound is voiced strongly, may have few or no vertical striations.
3. The ð bars are usually joined to the vowel bars.

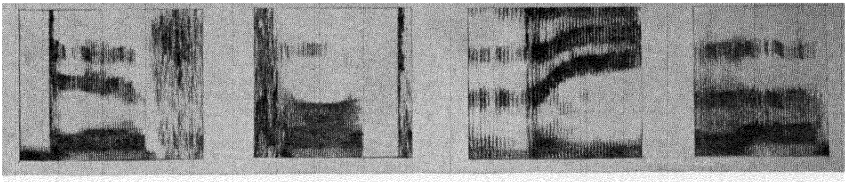
Read the θ *and* ð *by:*

1. The appearance of the voiced and voiceless fill.
2. The position of the hub.
3. The influence on the vowels.

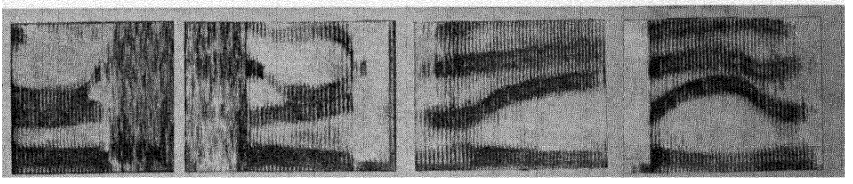
Review words

thick, page 128 ***thank***, page 129 ***this***, page 128 ***that***, page 129
both, page 129 ***though***, page 129

New words to learn



death *thought* *the, thee* *the*

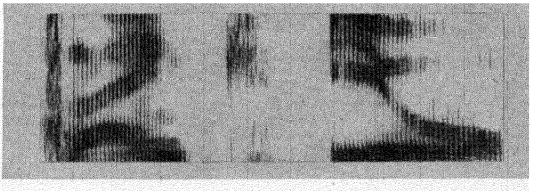


earth *third* ***they*** ***bathe***

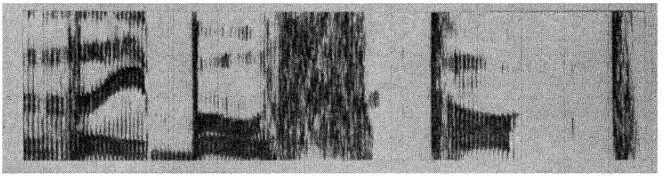
Additional words for practice in visualization

thirty *path* ***thy*** *thus*

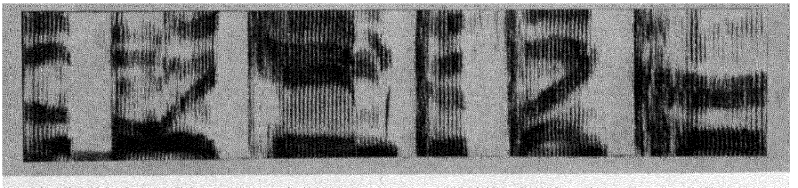
Drill phrases and sentences



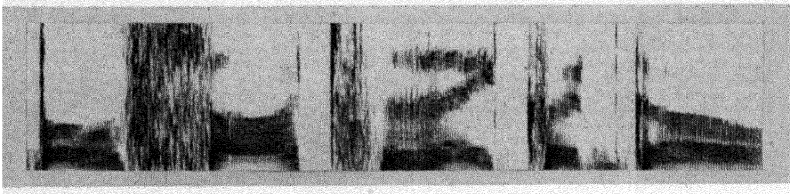
Thank you



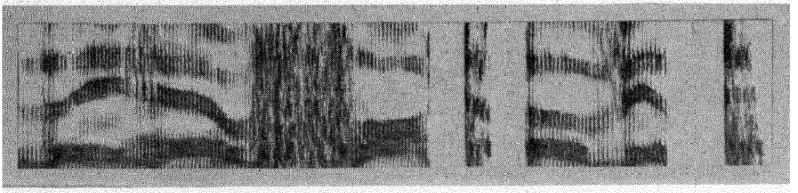
They both talked



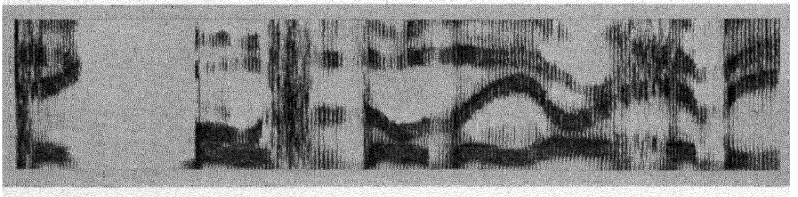
The boy came to thank her



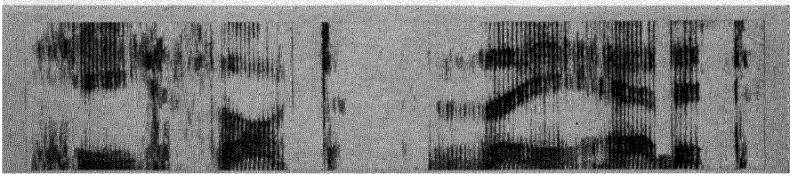
Both thought they could go



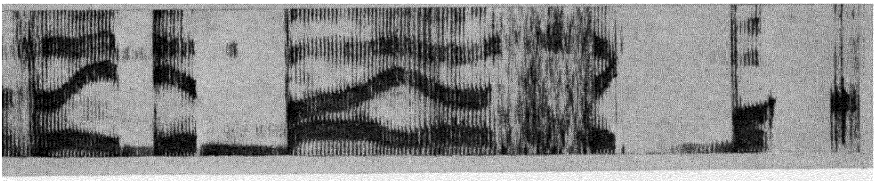
They have thought of it



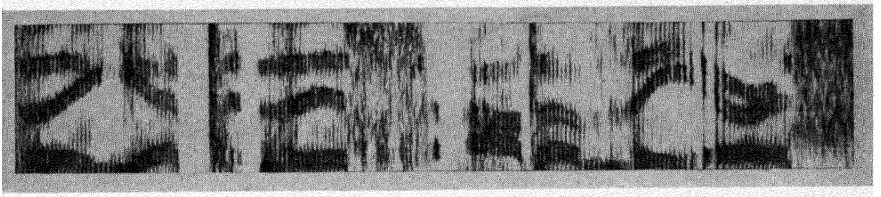
Take both, though they are heavy



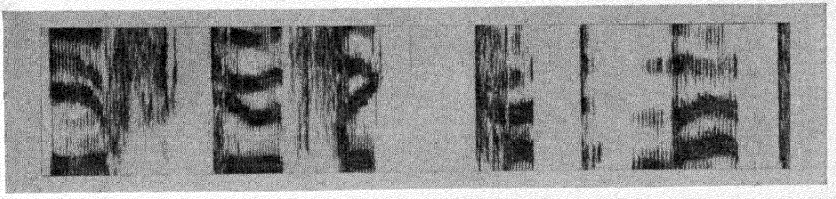
He thought they had it



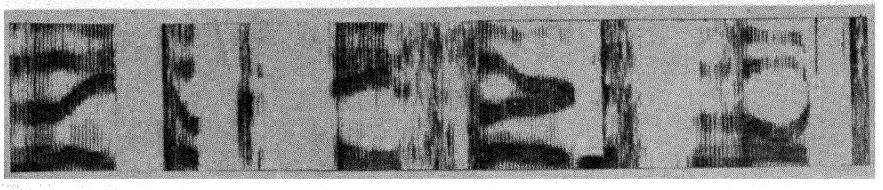
They did buy a thick book



They get this out of the earth



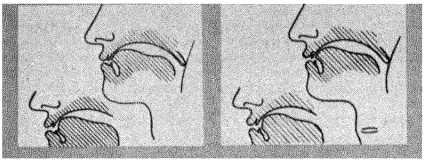
If too thick, cut that



They took a third vote

Unit 4 — Reading the s (see) and z (zoo)

How the s and z are made



Profile of s (*see*) *Profile of* z (*zoo*)

The s and z are made by most speakers from one or both of two positions which are usually described as the tongue-point alveolar ridge, and the tongue-blade alveolar ridge positions. The profiles above illustrate both these positions. Important considerations in making the s and z are:

1. Forming a narrow groove along the midline of the tongue tip (or blade) as it is held against the alveolar ridge.
2. Placing the sides of the tongue against the teeth and teeth ridge to direct the breath stream through the narrow groove.
3. Having a narrow opening between the upper and lower front teeth.
4. Opening the lips and spreading them a little.
5. Closing the opening between the throat (pharyngeal) and nasal cavities.
6. Emitting a continuous voiceless or voiced breath stream through the narrow tongue groove and against the cutting edge of the teeth.

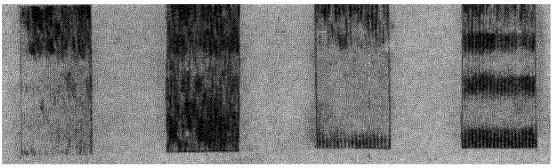
The frictional modulation produced by sending a thin stream of breath against the cutting edge of the teeth is characteristic of both the s and the z sounds. The z differs from the s in that it is voiced, since vocal cord and cavity modulation are combined with frictional modulation in producing the z. This difference in the modulation of the sounds has led to the phonetic description of the s as a voiceless fricative continuant, and the z as a voiced fricative continuant.

The s and z may be spelled in many ways

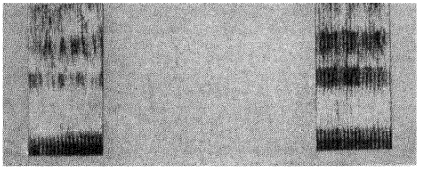
s (*sew*)	*ss* (*hiss*)	*se* (*nurse*)	*z* (*zone*)	*zz* (*buzz*)	*sc* (*discern*)
sc (*scene*	*sch* (*schism*)	*st* (*listen*)	*ze* (*breeze*)	*cz* (*czar*)	*ts* (*tsar*)
sw (*answer*)	*c* (*cell*)	*ce* (*force*)	*s* (*is*)	*se* (*wise*)	*es* (*hives*)
ps (*psalm*)	*z* (*quartz*)				*x* (*xylophone*)

Visible patterns of the s and z

The pattern for the s is characterized by fills, as shown in the illustrations on the facing page. The appearance of the fill varies mainly with the amount of stress used in



s *unstressed* s *stressed* z *unstressed* z *stressed*



z *frictional modulation predominant* z *vocal cord-cavity modulation predominant*

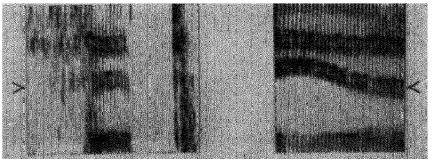
making the sound. However, both the stressed and unstressed s fills usually appear darkest towards the top of the pattern. Unstressing decreases the number, vertical length, and darkness of the striations. Stressing increases the number, length, and darkness of the vertical striations, and frequently produces light resonance bars in the upper part of the fill. The resonance bars in the s fill are due primarily to modulation that occurs in a cavity between the teeth and lips. The positions of these bars can be varied by moving the lips from the i (*eve*) to the u (*boot*) positions as the s is emitted. In general, it may be hypothecated that cavity modulation of the voiceless fricatives occurs in a cavity formed between the frictional source and the point of emission of the breath from the vocal tract.

The pattern of the z may be thought of as the s fill with a voice bar and one or more resonance bars in it. The vertical striations, characteristic of the s fill, are produced by frictional modulation, and the z bars are added by the vocal cord and cavity modulation. Like the s, the appearance of the z pattern is modified considerably by the amount of stress used in making the sound. Comparison of the illustrations of stressing and unstressing presented above, indicates the way in which unstressing decreases the darkness and the number of the resonance bars. If frictional modulation is dominant, the z pattern will have strong vertical striations and weak bars. When vocal cord and cavity modulation are dominant, the pattern will have strong bars and the vertical striations will be weak or absent. Frequently, the z spoken in continuous speech is unvoiced partially, especially when it precedes an unvoiced sound as in

has to. The pattern in such a case starts with z voicing and bars, shifts into the predominant frictional modulation type pattern, and then into an unvoiced fill or frictional modulation pattern.

Corresponding to other cognate pairs of sounds discussed in the preceding units the s and z have the same hub because they are made from approximately the same articulatory positions. The location of the hub for both sounds is the bar 2 position of z. The z hub is used as a reference because it is usually visible, while the s hub is seldom apparent in the s fill. Variation in the part of the tongue used in forming the restricted opening for the s and z changes the hub position and the transitional pattern slightly. The s and z are considered to be relatively fixed sounds; and, although the modifications are of interest to the phonetician and the speech correctionist, they do not affect the reading legibility of the patterns and therefore are not included in this explanation.

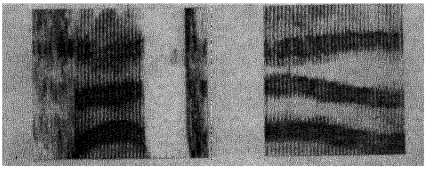
How the s and z influence vowel patterns



sit ***is***

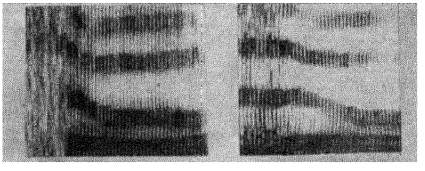
The patterns of the words *sit* and *is* illustrate the influence of the s and z on the vowel ɪ (*it*). Since the s and z have similar hubs, their influence on any vowel is similar. The position of the s–z hub, > and <, is just below the center of the pattern while the hub of ɪ is considerably higher. When s precedes ɪ as in *sit*, the beginning of bar 2 of the vowel curves upwards. Part of the transitional pattern in the s fill can be observed immediately before the beginning of the vowel bars. In *is*, the z follows the ɪ, and the curve of bar 2 of the vowel is downwards from the ɪ hub to the z hub. Because of the continuous voicing from ɪ through z, the vowel bars are joined with the z bars, thus making the transitional pattern very apparent. As bar 1 of ɪ joins the voice bar of z, sometimes, there is a slight curve in the bar. Bar 3 slopes downwards from its ɪ position to the bar 3 position of z.

The influence of s and z on æ (*at*) can be seen in the patterns of the words *sat* and *as.* An inspection of the patterns will reveal that the s–z hub is only slightly lower than the æ hub; therefore, the s–z influence is less marked on bar 2 of æ (*at*) than of ɪ (*it*). The degree of influence on bar 1 is determined by its distance from the baseline, so



sat *as*

that bar 1 of æ is curved more by the s and z than is bar 1 of ɪ. Observe that bar 1 of æ begins on the baseline in *sat* due to the shift in type of modulation, and returns to the baseline in *as* for the same reason. Bar 3 of æ is a little lower than bar 3 of ɪ, which results in a corresponding decrease in the s–z influence.



sue *zoo*

Bar 2 of all of the back vowels is lower than the s–z hub. The extent of the difference between the hub positions of s–z and the vowel increases with each of the back vowels as said in series from ɑ (*father*) to u (*boot*). In u, as the patterns of *sue* and *zoo* show, bar 2 curves downwards from the s–z hub to the u hub. Since bar 1 of u is on the baseline, it is influenced less by the s and z than is bar 1 of the other back vowels. It is evident in the pattern of *zoo* that the voice bar of the z, which is always on the baseline, joins with bar 1 of u with only a slight change in the width of the bar due to shift in modulation. Bar 3 in *zoo* curves downwards from its z position to its u position. That the s and z have the same transitional influence on the u can be seen by comparison of the patterns of the words *sue* and *zoo*. The differences in the appearance of the transitional patterns result from differences in type of modulation, as has been discussed previously.

By extending the information presented in the above illustration, the s and z influence on the other vowels may be predicted and visualized. The s–z hub is relatively fixed and is approximately the same as the æ (*at*) hub; but it may be slightly above or below the æ hub in certain combinations. Because the s–z hub is just below the middle of the pattern, its greatest influence is evident on the i (*eve*) hub (the highest vowel hub) and on the u (*boot*) hub (the lowest vowel hub). Therefore, the influence

of the s and z on bar 2 of the front vowels decreases with each of the vowels as said in series from i (*eve*) to æ (*at*), and the s–z influence on bar 2 of the back vowels increases with each of the vowels as said in series from ɑ (*father*) to u (*boot*). Bar 1 of both the i (*eve*) and u (*boot*) are on the baseline, and are consequently uninfluenced by the s–z. In the front vowel series, the s–z influence on bar 1 increases with each vowel from i (*eve*) to a (*ask*), while in the back vowel series their influence decreases with each of the vowels from ɑ (*father*) to u (*boot*). Bar 3 of the front vowels is influenced by the s and z in the same way as bar 2 except that the extent of the influence is not so great. Bar 3 is frequently absent in the patterns of the back vowels, but when it is present the s–z influence is not very evident except for u (*boot*), in which case the influence is similar to that on bar 2.

The visible patterns of the s *and* z *differ chiefly in that:*

1. The z is voiced; therefore, its pattern has a voice bar on the baseline and usually has one or more resonance bars.
2. The z pattern, when the sound is voiced strongly, may have few or no vertical striations.
3. The z bars are usually joined to the vowel bars.

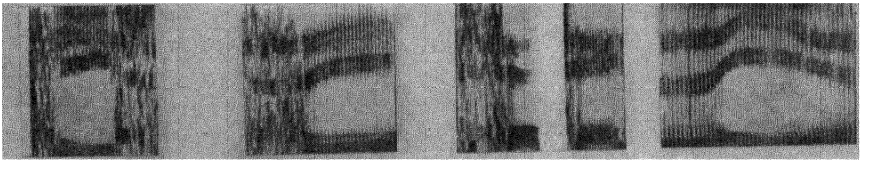
Read the s *and* z *by:*

1. The appearance of the voiced and voiceless fill.
2. The position of the hub.
3. The influence on the vowels.

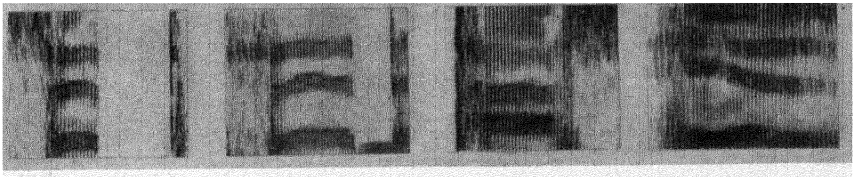
Review words

sit, page 136	***this***, page 128	***is***, page 136	***as***, page 137
sat, page 137	***sue***, page 137	***zoo***, page 137	

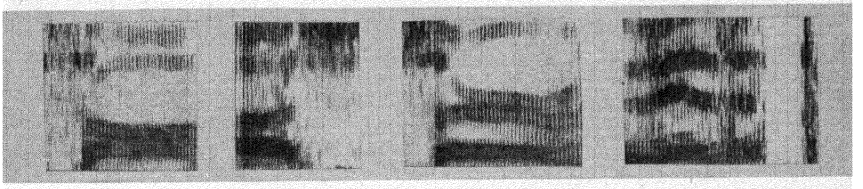
New words to learn



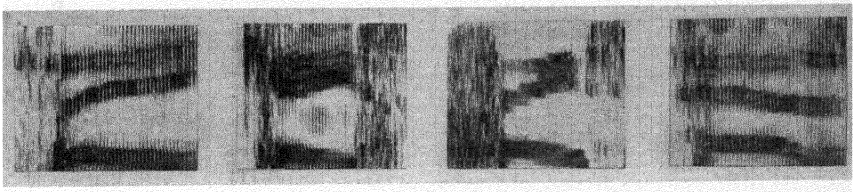
peace, piece ***see, sea*** ***city*** ***these***



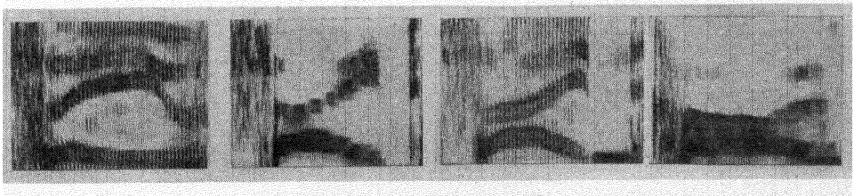
set *said* *pass* *his*



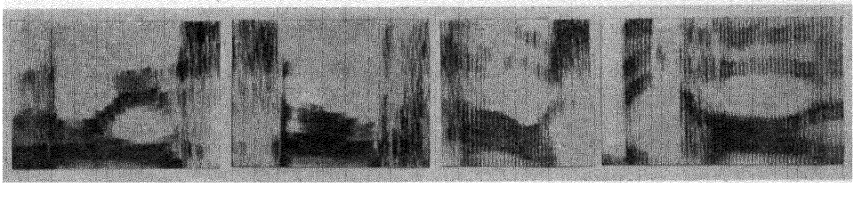
saw *us* *serve* *visit*



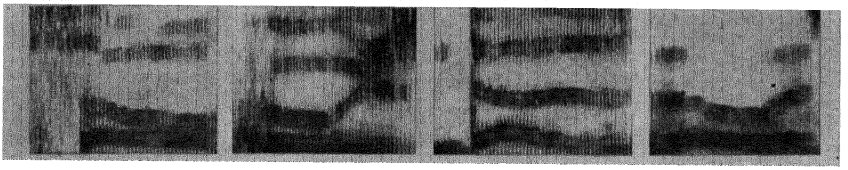
say *case* *face* *has*



save *sight* *side* *cause*



voice *south* *house* *because*

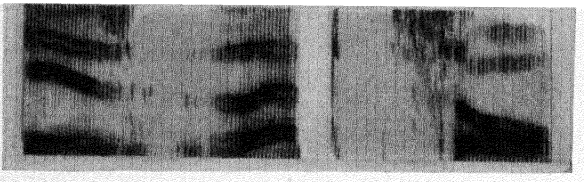


so, sew *whose* *does* *those*

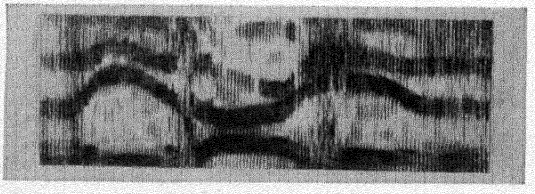
Additional words for practice in visualization

seat	*basket*	*seed*	*sat*	*seek*	*suit*	*sieve*
seize	*thus*	*sick*	*service*	*kiss*	*safe*	*busy*
office	*decide*	*guess*	*outside*	*escape*	*easy*	*size*

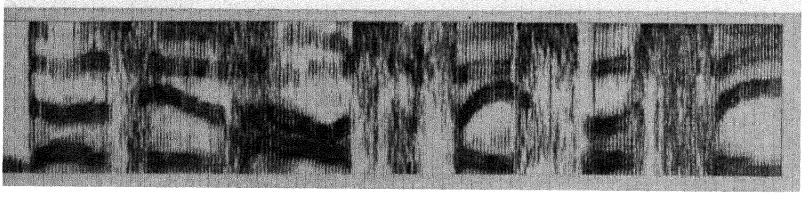
Drill phrases and sentences



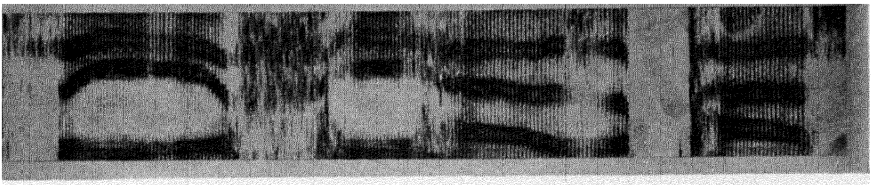
Is that so?



These are his



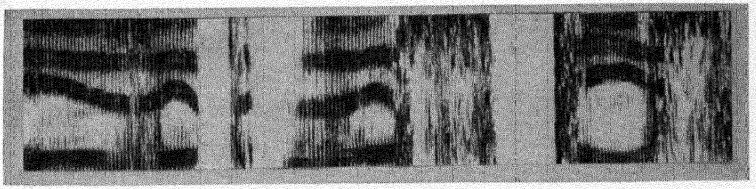
Does his house face the sea?



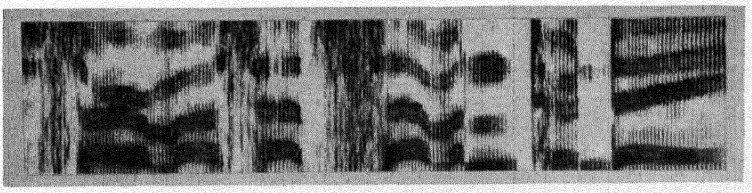
See if he has a pass



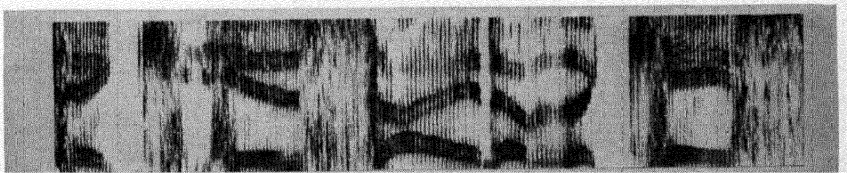
I said so, because I saw her



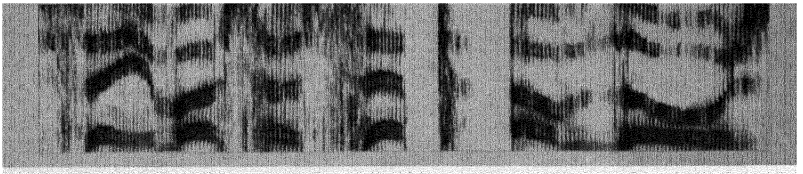
Is it this piece?



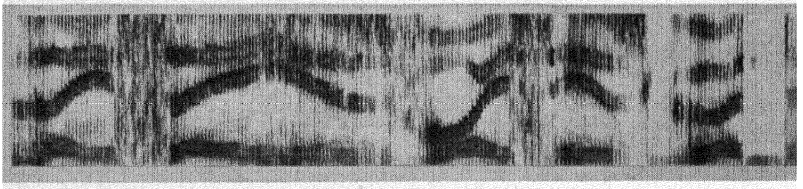
Serve us at seven today



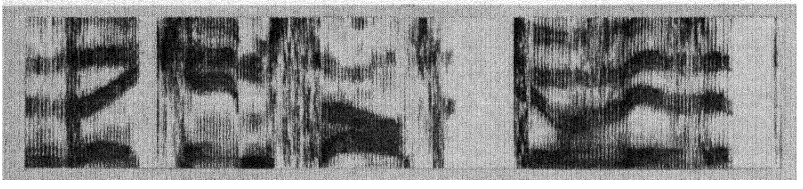
Take his side of the case



Save us a set of those



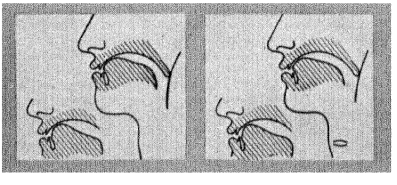
They say his voice is good



They came south to visit

Unit 5 — Reading the ʃ (she) and ʒ (azure)

How the ʃ and ʒ are made



Profile of ʃ (*she*) *Profile* ʒ (*azure*)

The ʃ and ʒ are produced generally from the positions diagrammed in the above profiles. In general, the tip of the tongue is raised beneath the anterior portion of the

palate; but some speakers make these sounds, especially in combination with certain back vowels, by elevating the blade of the tongue instead of the tongue tip. The physiological description of the sounds, as lingua-palatal, does not differentiate the parts of the tongue used in their production. The more important factors in making the ʃ and ʒ include:

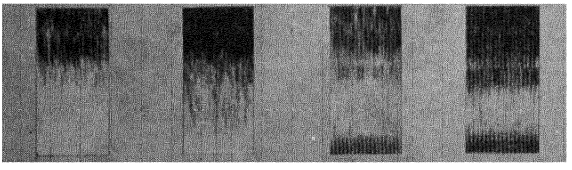
1. Forming a wide, shallow opening between the flattened tongue tip (or blade) and the anterior palate.
2. Placing the sides of the tongue against the teeth and teeth ridge to direct the breath stream through the wide shallow opening.
3. Having a narrow opening between the upper and lower front teeth.
4. Opening the lips and rounding them a little.
5. Closing the opening between the throat (pharyngeal) and nasal cavities.
6. Emitting a continuous voiceless or voiced breath stream through the wide shallow groove against the cutting edge of the teeth and through the cavity formed by the lips.

Since the particular type of frictional modulation produced as described above is common to both the ʃ and the ʒ, the sounds are differentiated by the vocal cord-partially closed cavity modulation which is present in ʒ and absent in ʃ. This difference in the modulation of the sounds is the basis for the phonetic description of the ʃ as a voiceless fricative continuant, and the ʒ as a voiced fricative continuant. In physiologic phonetics the ʃ is called a voiceless lingua-palatal continuant and the ʒ is designated a voiced lingua-palatal continuant. There are not many English words in which the ʒ sound is used by itself. It occurs most frequently in combination with the d as dʒ in such words as *edge* and *judge*. The dʒ is discussed in Unit 7 of this chapter.

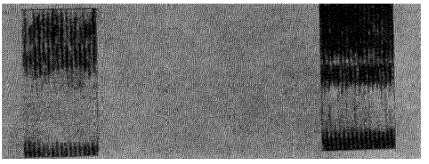
The ʃ and ʒ may be spelled in many ways

sh (*she*)	***s*** (*sure*)	***ss*** (*issue*)	***z*** (*azure*)	***zi*** (*glazier*)
se (*nauseous*)	***sci*** (*conscious*)	***chs*** (*fuchsia*)	***si*** (*vision*)	***ssi*** (*abscission*)
si (*pension*)	***ssi*** (*mission*)	***ci*** (*social*)	***g*** (*regime*)	***ge*** (*rouge*)
ce (*ocean*)	***ch*** (*machine*)	***che*** (*moustache*)	***s*** (*pleasure*)	
ti (*nation*)	***x*** (*luxury*)			

Visible patterns of the ʃ and ʒ



ʃ	ʃ	ʒ	ʒ
unstressed	***stressed***	***unstressed***	***stressed***



ʒ
frictional modulation predominant

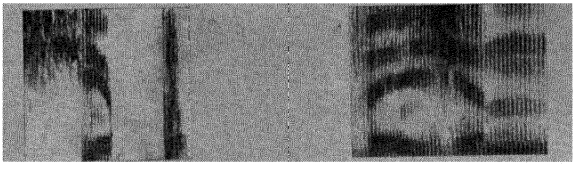
ʒ
vocal cord-cavity modulation predominant

Since the ʃ is produced primarily by frictional modulation, its pattern is characterized by a fill as illustrated above. Certain aspects of the ʃ fill, like the s fill, vary with the amount of stress used in producing the sound. As in the other voiceless fricatives, unstressing decreases the number and darkness of the vertical striations. Unlike the other fricatives, the length of the ʃ vertical striations is not affected very much by stressing or unstressing, and the position of the fill in the upper part of the pattern is constant. Sometimes stressing produces light resonance bars in the ʃ fill. Such bars also occur in other voiceless fricative fills, and are believed to be related to cavity modulation in the forward part of the vocal tract.

The ʒ is made by frictional modulation plus vocal cord and partially closed cavity modulation. Therefore, the ʒ pattern is characterized by a combination of vertical striations made by the frictional modulation, and of bars that are produced by the vocal cord-cavity modulation. The voice bar is always on the baseline because of the restricted opening of the vocal cavity. Bar 2 is located in about the same position as bar 2 of the vowel i (*eve*); and bar 3, when apparent, is near the top of the pattern. The positions of bars 2 and 3 may vary slightly when the ʒ is produced from different articulatory positions. Stressing and unstressing modify the number and darkness of the ʒ bars and fill. The appearance of the sound also is determined in part by the relative proportions of the fill and bar components. If frictional modulation predominates, the ʒ pattern will have a voice bar but the upper bars may be hidden in the fill. Conversely, if vocal cord-partially closed cavity modulation predominates, the bars will be stronger and the fill less apparent. Usually the fill is more apparent in the ʒ pattern than it is in the case of the other voiced fricatives. To some extent, this is due to the high intensity level of the ʃ and ʒ. They are among the strongest sounds produced, and are considerably higher in level than any of the other fricatives. Also, the ʃ–ʒ fill area coincides with the hub and the positions of the upper resonance bars. The voiced and voiceless fills for these sounds are generally so dark or intense, especially in comparison with other sound patterns, that it is difficult to differentiate the component fill and bar elements.

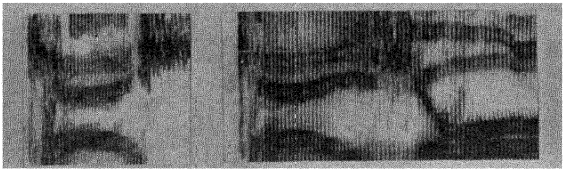
The ʃ and ʒ have similar hubs because they are made from similar articulatory positions. The location of the hub for both sounds is the bar 2 position of ʒ, which nearly coincides with the i (*eve*) hub. Although the ʒ hub is used as the reference, the hub is also frequently visible in the ʃ pattern.

How the ʃ and ʒ influence vowel patterns



ship ***vision***

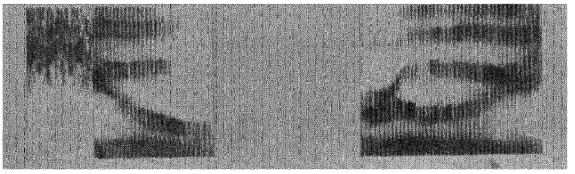
The influence of the ʃ and ʒ on the vowel ɪ (*it*) is illustrated in the words *ship* and *vision*. Since the ʃ and ʒ have similar hubs, their primary influence on any vowel is similar. The position of the ʃ–ʒ hub is towards the top of the pattern, and the hub of the vowel ɪ (*it*) is slightly lower. When ʃ precedes the vowel ɪ as in *ship*, there is only a small shift in the beginning of bar 2, as well as the beginnings of bars 1 and 3. The sharp downward curves of the ends of bars 2 and 3 are due to the influence of **p**. In the pattern of *vision*, the ʒ follows the vowel ɪ and there is little noticeable influence on either bar 2 or bar 3. The ʃ and ʒ influence the vowels ɪ (*it*) and i (*eve*) less than any other vowels because the hubs of these vowels are so close to the ʃ–ʒ hub.



cash ***casual***

When the ʃ and ʒ are combined with the vowel æ as in *cash* and *casual*, all of the vowel bars are influenced. The degree and direction of the primary influence may be explained by the difference in the hub positions. The ʃ–ʒ hub remains relatively fixed towards the top of the pattern, while the æ hub is near the center of the pattern. In *cash*, the ʃ follows the æ, so that the ending of bar 2 curves upwards from the æ hub to the ʃ hub. The ending of bar 3 also curves upwards but not as much as the ending of bar 2. Bar 1 curves downwards from its æ position towards the baseline. In *casual*,

the ʒ follows the æ, and the degree and direction of the ʒ influence on the vowel is approximately the same as the ʃ influence as illustrated in the pattern of *cash*. The continuous voicing of the vowel and the ʒ makes the transitional influence and pattern more apparent than that between the voiceless ʃ and the æ. Differences may be observed in the appearance of the transitional patterns due to the types of modulation combined.



shoe ***rouge***

As shown in the patterns of *shoe* and *rouge*, the ʃ and ʒ influence is most marked on the vowel **u**. This is because the ʃ–ʒ hub is farther from the **u** hub than it is from the hub of any other vowel. In *shoe*, bar 2 curves downwards from the ʃ hub to the **u** hub, bar 1 remains on the baseline uninfluenced by the ʃ, the beginning of bar 3 curves downwards from the ʃ to its **u** position, and bar 4 remains straight and uninfluenced. The ʒ, as shown in *rouge*, has the same transitional influence as the ʃ, on the vowel **u**. In *rouge*, the ʒ follows the vowel; therefore, bar 2 curves upwards from the hub of **u** to the hub of ʒ. Bar 3 also curves upwards slightly, but bars 1 and 4 remain straight. Since the ʒ is a voiced sound, its resonance bars are joined with the vowel bars. Because the ʃ and ʒ are made from relatively fixed and high positions, a larger portion of the transitional pattern is visible in the movement of the vowel bars than usually is apparent in the fricative-vowel combinations.

The ʃ–ʒ influence on the ɪ (*it*), æ (*at*), and **u** (*boot*) shows that the ʃ–ʒ hub remains relatively fixed towards the top of the pattern. Thus the direction of the primary or bar 2 influence of these sounds on any vowel can be predicted when the vowel hub is known. Since the hubs of all vowels, excepting **i** (*eve*) and ɪ (*it*), are lower than the ʃ–ʒ hub, the direction of the transitional pattern may be predicted easily. The range of the transitional patterns is determined by the distance between the hubs of the combined sounds. The appearance of the transitional patterns is determined by the types of modulation used in the combined sounds.

The visible patterns of the ʃ *and* ʒ *differ chiefly in that:*

1. The ʒ is voiced; therefore its pattern has a voice bar on the baseline and may have one or more resonance bars.
2. The ʒ pattern, when the sound is voiced strongly, may have few vertical striations.
3. The ʒ bars are usually joined to the vowel bars.

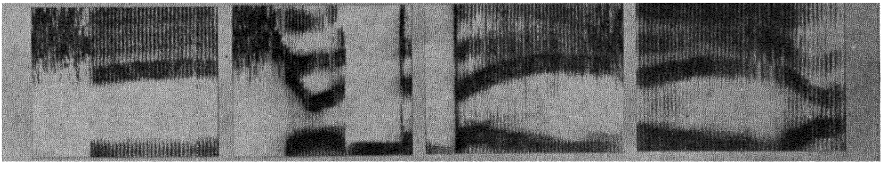
Read the ʃ and ʒ by:

1. The appearance of the voiced and voiceless fill.
2. The position of the hub.
3. The influence on the vowels.

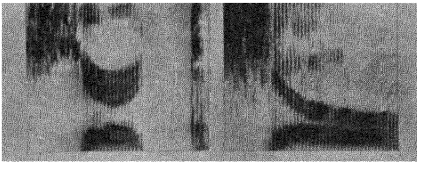
Review words

ship, page 145 *cash*, page 145 *casual*, page 145 *vision*, page 145
shoe, page 146 *rouge*, page 146

New words to learn



she *should* *beige* *Asia*

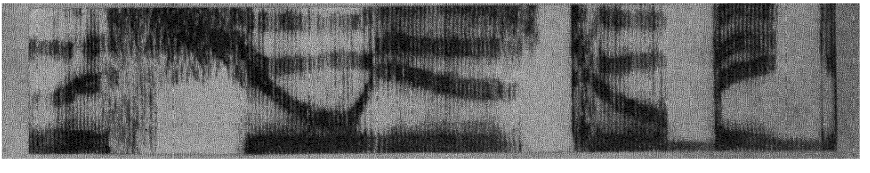


shirt *show*

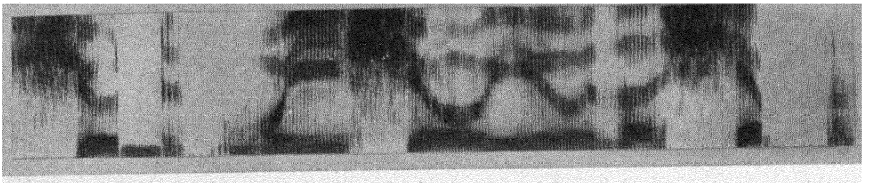
Additional words for practice in visualization

sheep	*fish*	*shop*	*shut*	*lesion*
shade	*shake*	*shape*	*shout*	*collision*

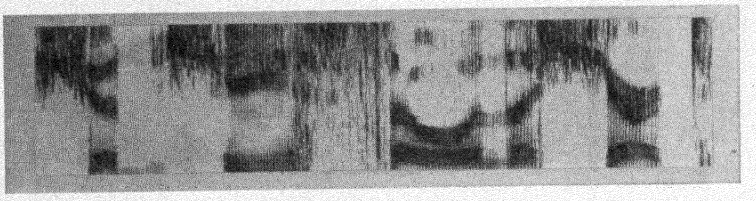
Drill phrases and sentences



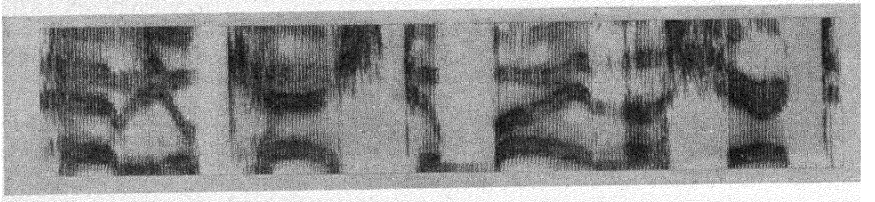
This shoe is too big



Should we show you the ship?



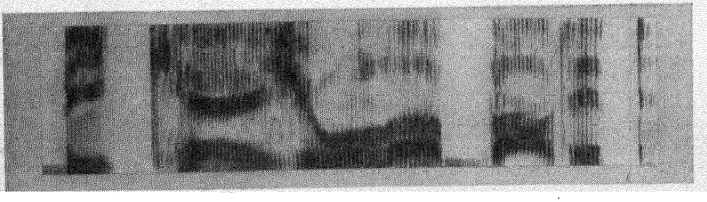
Should she sew the shirt?



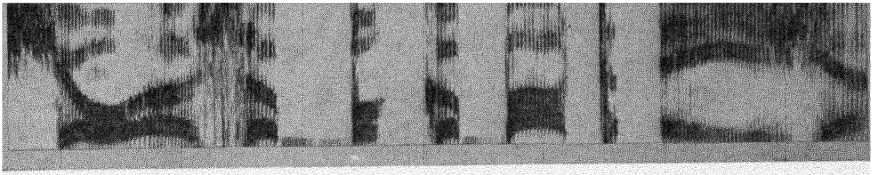
Have you cash to buy the shirt?



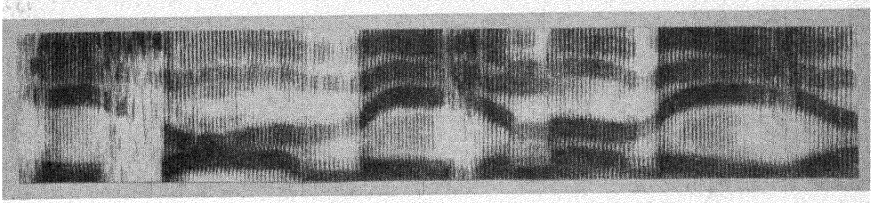
We should pay cash to save



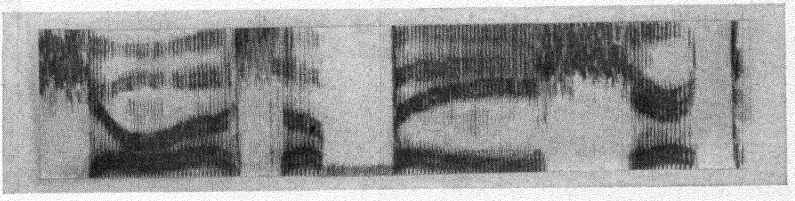
Be casual about it



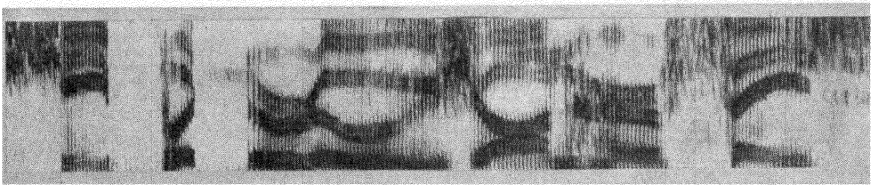
Show us a book about Asia



He saw a vision of Asia



Show us that beige shirt



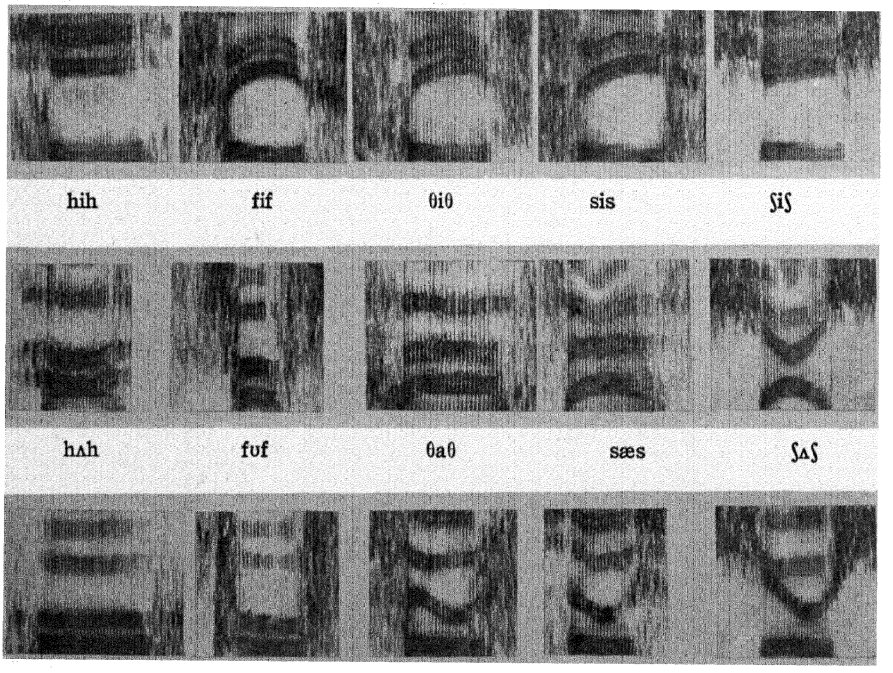
She put rouge on her face

Unit 6 — Comparative Review of the Fricative Sounds

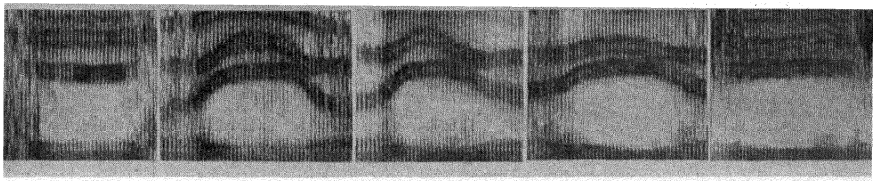
This comparative review of the fricative sounds is presented in order to facilitate a comparison of the similarities and differences in their visible speech patterns. Like the stop sounds, the fricative sounds have been discussed in pairs, h–ɦ, f–v, θ–ð, s–z, ʃ–ʒ,

because the two consonants in each pair are made from similar articulatory positions. The voiceless sound in each pair, **h–f–θ–s–ʃ**, is produced primarily with frictional modulation, which is sometimes supplemented with cavity modulation. The voiced sound of each pair, **ɦ–v–ð–z–ʒ**, is made with frictional modulation plus vocal cord and partially closed cavity modulation. The presence or absence of vocal cord-partially closed cavity modulation results in variations in the appearance of the patterns by which **h** can be differentiated from **ɦ**, **f** from **v**, **θ** from **ð**, **s** from **z**, and **ʃ** from **ʒ**. Since each pair of fricative sounds is made from different articulatory positions, their patterns provide certain clues which are usually characteristic of the sounds and by which each pair may be differentiated from the other pairs. The most important of these clues are illustrated and summarized below for comparison.

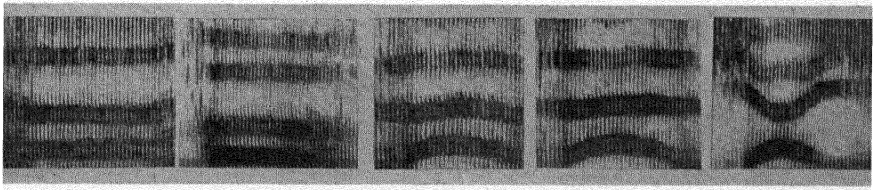
In the figure that follows there are patterns of each of the fricative sounds as it is said before and after three different vowels. Two vowels, the **i** (*eve*) and **u** (*boot*), are used with all the fricatives. Between the patterns that contain the **i** and the **u**, there is a pattern which includes a vowel with a hub that most nearly coincides with the particular fricative hub under consideration. In case the **i** or the **u** hub most nearly



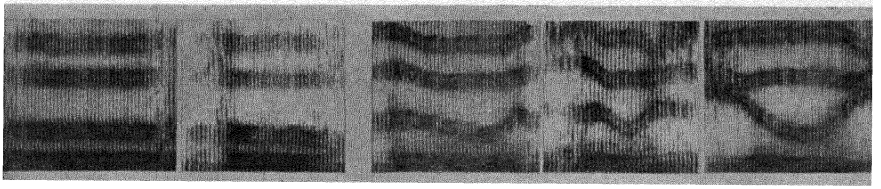
coincides with a fricative hub, the neutral vowel ʌ (*up*) is used as the second vowel with that fricative. This occurs in the ʃ–ʒ series because the i hub is as near to the ʃ–ʒ hub as is the hub of any vowel. Therefore, to avoid duplication of the ʃiʃ and ʒiʒ patterns, the neutral vowel is substituted for the i in this series. The neutral vowel is used also in the h–ɦ series but for a different reason. Since the h–ɦ hub is that of the vowel with which it is combined, it is impossible to select one vowel to illustrate it. The first three patterns in any vertical column portray a voiceless fricative before and after three different vowels. The last three patterns in any vertical column show a voiced fricative before and after the same vowels used with its voiceless cognate. Thus a comparison of the patterns within a single vertical column will reveal the similarities and differences of cognate sounds. A comparison of the patterns of different vertical columns will show the similarities and differences among different fricative sounds when combined with the vowels i (*eve*) and u (*boot*). The patterns between the i and u patterns for both the voiceless and voiced series, except for the ʃ–ʒ patterns as explained above, permit a comparison of the hubs of the different fricative sounds.



ɦiɦi viv ðið ziz ʒiʒ



ɦʌɦi vʊv ðað zæz ʒʌʒ



ɦuɦi vuv ðuð zuz ʒuʒ

Each pair of fricative sounds has a different hub position

1. The **h**–**ɦ** hub is in the same position as the vowel with which the **h** or **ɦ** is made.
2. The **f**–**v** hub is towards the bottom of the pattern near the **ʊ** (*foot*) hub.
3. The **θ**–**ð** hub is below the middle of the pattern near the **ɑ** (*ask*) hub.
4. The **s**–**z** hub is approximately at the middle of the pattern near the **æ** (*at*) hub.
5. The **ʃ**–**ʒ** hub is towards the top of the pattern near the **i** (*eve*) hub.

Each voiceless fricative has a different fill

1. The **h** is darkest in the positions of the vowel bars with which the **h** is made.
2. The **f** usually has small dark areas randomly scattered, with the darkest areas towards the baseline.
3. The **θ** is darker towards the bottom of the pattern.
4. The **s** is darker towards the top of the pattern and the spikes frequently have small gaps between them.
5. The **ʃ** is very dark at the top of the pattern and the fill seldom extends below the center of the pattern.

Each voiced fricative has a different voiced fill

The voiced fill for any voiced fricative is a combination of the fill of the voiceless cognate plus resonance bars; therefore, each voiced fill is differentiated by the unique characteristics of the fill component plus the position of bar 2 or the hub of the sound. Since the unique characteristics of the fills and the hub positions, for the various fricatives, are summarized above, a restatement of them here would be unnecessarily repetitious.

Each pair of fricative sounds influences the vowels differently

1. The **h** and **ɦ** usually do not influence the vowel.
2. The **f** and **v** influence the front vowels more than they do the back vowels. Their greatest influence is on bars 2 and 3 of **i** (*eve*) and they have the least influence on the vowel **ʊ** (*foot*).
3. The **θ** and **ð** influence the front vowels more than they do the back vowels. Like the **f** and **v**, their greatest influence is on bars 2 and 3 of **i** (*eve*); however, the **θ** and **ð** do not influence bar 3 of the other front vowels as much as do the **f** and **v**. The vowel **ɑ** (*ask*) is influenced the least by the **θ** and **ð**.
4. The **s** and **z** influence bar 2 in **u** (*boot*) and in **i** (*eve*) the most. They influence bar 3 more in the front vowels than in the back vowels.
5. The **ʃ** and **ʒ** influence bar 2 in **u** (*boot*) the most. They influence the back vowels more than the front vowels.

The characteristic clues by which the fricative sounds are recognized and differentiated are discussed in considerable detail in Units 1 to 5. A review of these units will benefit those interested in learning to read visible speech. As mentioned in the conclusion of the summary for the stop sounds, the nature of the review will vary with the interest of the reader. As an aid to learning, we suggest that the fills and voiced

fills, and their combination patterns in words, be visualized and reproduced in some graphic form. By using the additional words provided for visualization, the accuracy of the visualized and graphed pattern may be verified by comparison with the visible speech patterns of the words provided in each unit.

Unit 7 — Reading Combinations of the Fricatives and Stops

Combinations of the fricatives, and combinations of the fricatives and stops, occur frequently within words and also between words and syllables in continuous speech. The characteristics essential to the recognition of the individual sound patterns, hubs, and influences, have been discussed in previous units, but there are additional characteristics which aid in the recognition of the combined sounds and which arise from the types of modulation and the articulatory positions involved.

This unit is designed to explain and illustrate the principles or clues underlying the recognition of various types of fricative and stop combinations that occur frequently in speech and therefore in visible speech reading.

Reading the tʃ (church) and dʒ (judge)

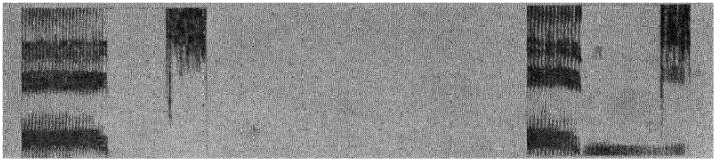
Both the tʃ (*church*) and the dʒ (*judge*) often are designated as sound units, primarily because they have been considered as individual units in spelling and also because they are the only sound combinations that occur initially, medially, and finally, in English words. We have included the tʃ and dʒ in the Group 6 or Combination Sound classification since, first, each is produced by combining two sounds, and second, each is portrayed by a combination of two distinct sound patterns.

The tʃ (*church*) is a combination of the stop t (*to*) and the fricative ʃ (*she*), and frequently is described as an *affricative* sound. The tongue takes the t position to stop or implode the breath; then the t is exploded quickly as the tongue moves into the ʃ position. The dʒ (*judge*) consists of the stop d (*day*) and the fricative ʒ (*azure*); and, like its voiceless equivalent, the tʃ, it is called an affricate. The tongue moves from the d position to the ʒ position as the d is exploded. Since the t, d, ʃ, and ʒ, are relatively fixed sounds, their articulation usually is altered slightly so that the transitional movement between the combined sounds may be made quickly and easily.

The tʃ and dʒ are spelled in the following ways

ch (*chew*)	*tch* (*etch*)	*g* (*gem*)	*dge* (*edge*)
ti (*question*)	*te* (*righteous*)	*j* (*just*)	*ge* (*page*)
t (*picture*)		*dg* (*lodger*)	*di* (*soldier*)
		gg (*exaggerate*)	*dj* (*adjust*)
		d (*verdure*)	

Visible patterns of the tʃ and dʒ

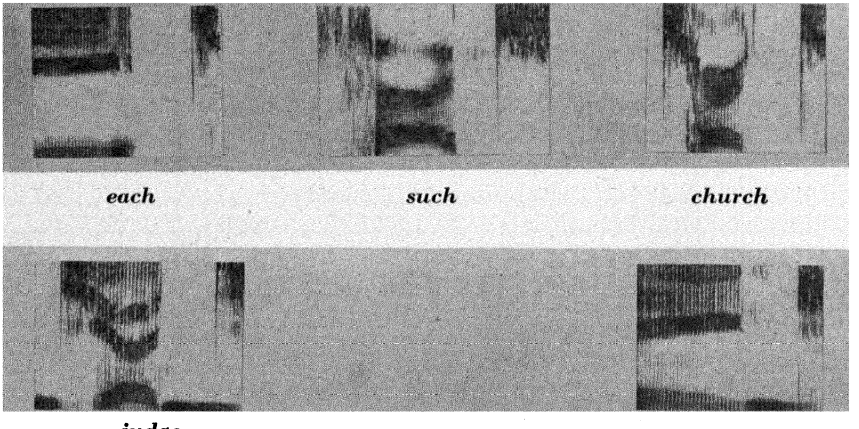


etch *edge*

Since the tʃ is a combination of a voiceless stop and a voiceless fricative, its pattern has a stop gap, a spike, and a fill. The spike fill of the t is not usually distinct from the ʃ fill because of the rapidity of transition. The dʒ, as the voiced cognate of the tʃ, is portrayed by a voiced stop gap, voiced spike, and voiced fill. The voiced spike and voiced fill form a continuous pattern which tends to be joined to the pattern of the sound that follows. The dark area of the fill striation may be observed towards the top of the pattern in the ʃ–ʒ position. The hub of the tʃ–dʒ approximates the t–d hub for sounds that precede, and the ʃ–ʒ hub for sounds that follow. In general, the tʃ–dʒ combined hub appears to be slightly lower than the ʃ–ʒ hub; and, therefore, the tʃ and dʒ influence the front vowels a little more and the back vowels a little less than do the ʃ and ʒ.

Review words *etch*, page 154 *edge*, page 154

New words to learn



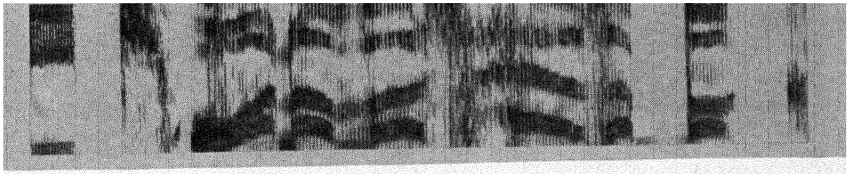
each *such* *church*

judge *age*

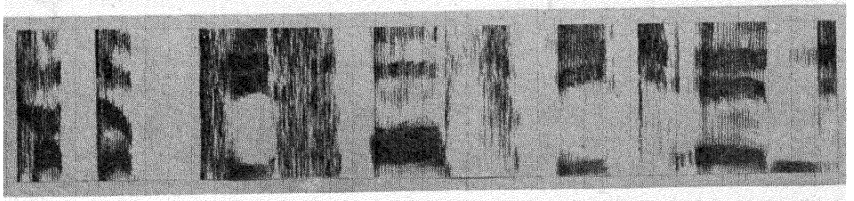
Additional words for practice in visualization

teach	*catch*	*object*	*subject*
choose	*touch*	*page*	*joy*

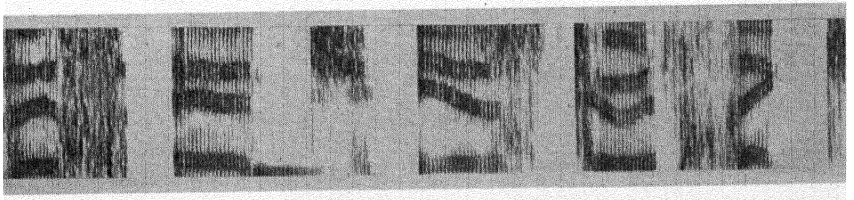
Drill phrases and sentences



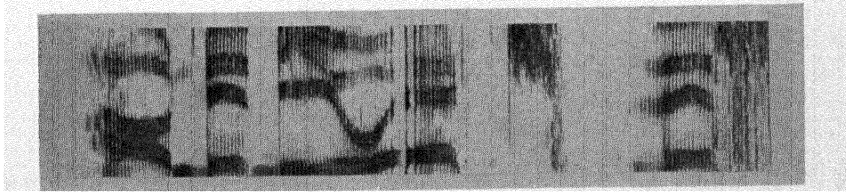
Each one of us has a book



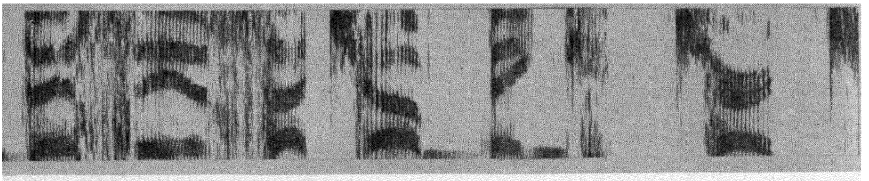
Cut a piece off each edge



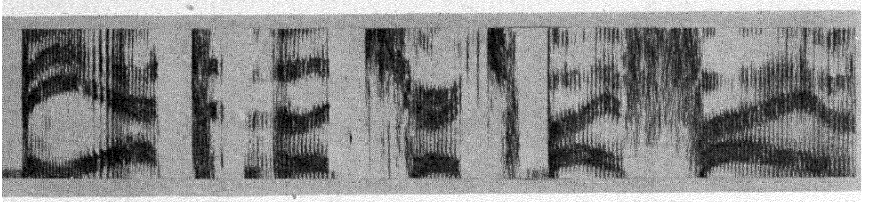
This edge is too thick



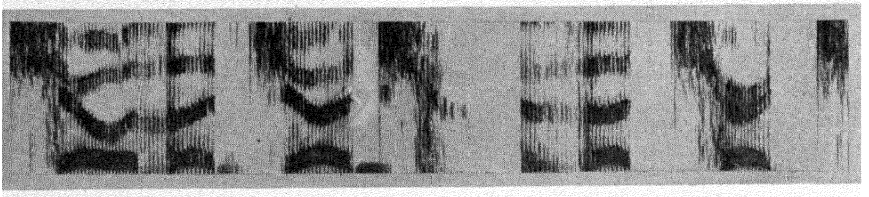
How did you etch this?



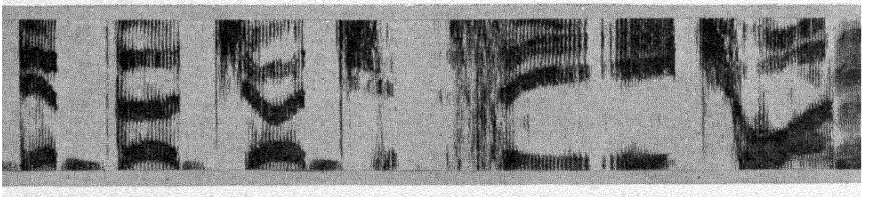
This is such a big church



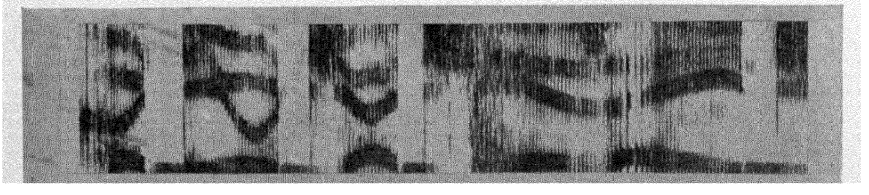
Be at the church by five



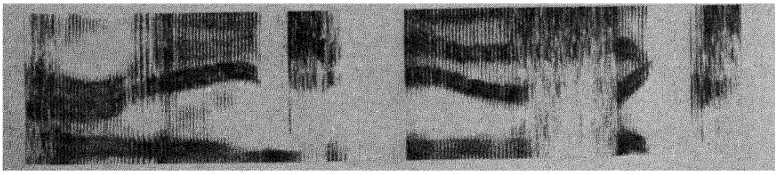
Show the judge the church



Did a judge see each one?

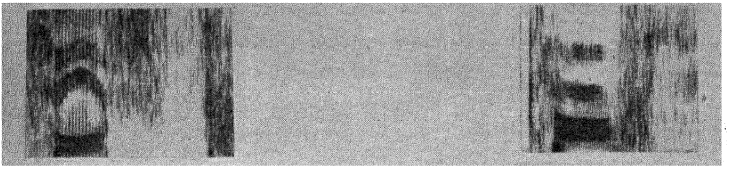


Could you judge his age?



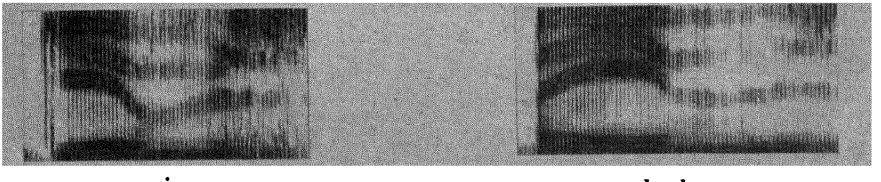
Her age is six

Reading combinations of the fricative sounds



fifth *cuffs*

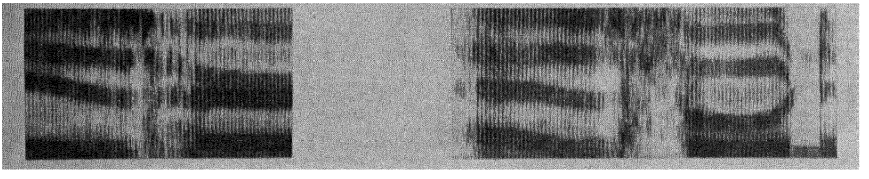
Frequently two voiceless fricative continuants are combined, as in the words *fifth* and *cuffs*. The characteristic pattern of each fricative is apparent in the combined sound although the point of transition is not always clear. If, as is customary, the breath stream is continued during the transitional movement, the termination of the first continuant is quite difficult to detect. The combination **fθ** in the word pattern at the left, *fifth*, may be compared with the **f** and **θ** spoken in isolation in the **f** and **θ** units, as may the **fs** in *cuffs* with the individual patterns of the **f** and **s**. It is possible to identify the first voiceless continuant in such a combination by its influence on the preceding sound and by its fill pattern. The second voiceless continuant may be recognized by its fill pattern and by its influence on the succeeding sound. To some extent, the speed of talking may modify the appearance of the fill patterns since, in rapid articulation, the position for a particular sound may be approximated rather than



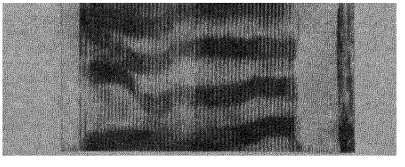
gives *bathes*

attained. Although the mutual influence of combined sounds has been discussed previously, this concept should be kept in mind by the reader throughout the explanations and discussions of individual and combined sounds.

When two voiced fricative continuants are combined, as in the words *gives* and *bathes*, the characteristic pattern of each sound is represented. The patterns of the **zv** combination in *gives* and the **ðz** in *bathes* may be compared with the patterns made by the sounds spoken in isolation and in combination with other sounds, as they have been illustrated in previous units. Unlike the voiceless fricative combinations, the voiced fricatives portray the transitional movements between the combined sounds, and thus identification of the sounds is facilitated. As in other combination sounds, the first sound may be identified in part by its influence on the preceding sound, the second combined sound by its influence on the sound which follows. Although rapid articulation may alter the individual patterns of the combined sounds, in general, the characteristic influences of the fricatives on the sounds which precede or follow them are not modified sufficiently to make identification and differentiation difficult. Movement of the resonance bars down to the **v** or **ð** and then up to the **z** position is apparent in both *gives* and *bathes*.



is her ***has food***

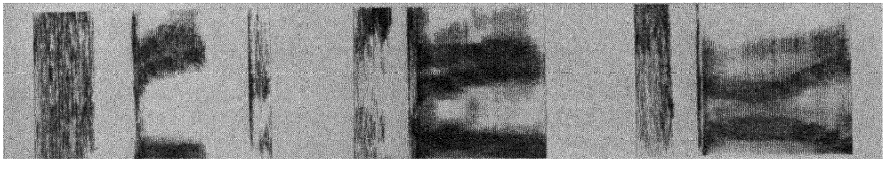


give that

There are many combinations of fricative sounds that may occur between syllables and between words in continuous speech. In the patterns above, representative voiced and voiceless fricative combinations are illustrated. The voiceless fricative **h** in *her* may be wholly or partially voiced, or may be voiceless, depending upon the speed at which vocal cord modulation is terminated and reinitiated during the transition from

z to h to the vowel. The z–f combination in *has food* portrays the shift as vocal cord-cavity modulation in the voiced fricative is terminated during the transition to the voiceless fricative. Part of the movement between the z and f articulatory positions is shown in the voiced portion of the transition. Since vocal cord modulation is continued during the transition from v to ð in *give that*, the resonance bars are continuous throughout both sound patterns. The direction of the transitional movement pattern from the v hub to the ð hub indicates the movement from one articulatory position to another.

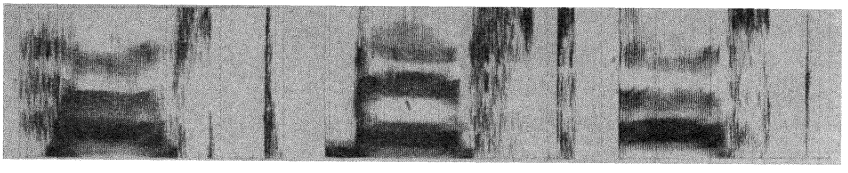
The patterns produced by the fricative combinations between words may be compared with the patterns of the individual sounds and with the patterns presented in the two preceding illustrations. Variations resulting from rapid articulation usually are evidenced in shorter patterns, in hubs tending towards the hub of the combined sound, and in reduction in the extent of transitional movement. Precise aural and visual differentiation of the types of modulation characteristic of an individual sound becomes more difficult in rapid speech, since the modulation changes are made continuously rather than abruptly. Therefore, especially in sounds of short duration, a characteristic modulation combination may not be reached completely or may be produced only for a very brief interval of time. For example, in a rapidly articulated voiced fricative-voiceless fricative combination, the transitional period may be short and vocal cord modulation may not be terminated completely until the voiceless fricative has been begun. Frequently when a vowel follows the voiceless fricative h, there is not sufficient time for some of the vowel resonances to build up, and consequently the third and fourth vowel resonance bars may be partially or completely absent.



speak ***stay*** ***sky***

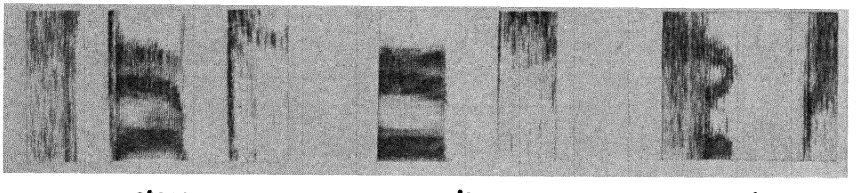
In a voiceless fricative-voiceless stop combination, as in the illustrations above, the fricative fill is separated from the stop spike by the stop gap. When the combination occurs initially or is not elided with the preceding sound, the fricative is identified primarily by its pattern. If the combination is spoken in continuous speech, the fricative influence may be apparent in the preceding sound pattern. The stop may be recognized both by its pattern and by its influence on the sound that follows. Usually,

in a voiceless fricative-stop blend, there is less aspiration following the stop than there is in the case of a stop-vowel combination. In rapid speech, the stop gap tends to be shortened, and it also varies somewhat in duration with the particular sounds combined. Generally the stop gap for the **s–t** blend is shorter than those for the **s–p** or **s–k** blends, due to the closer proximity of the articulatory positions of **s** and **t.**



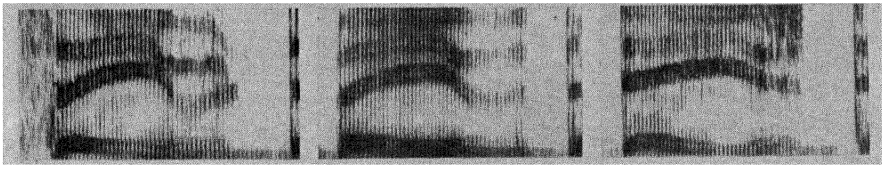
hasp ***best*** ***ask***

When the voiceless fricative-stop combination occurs in a final position, the fricative influence and pattern serve to identify the sound. The duration of the stop gap is similar to that in the same combinations in the initial position. Unless the stop sound is followed by another sound, it is identified primarily by the pattern of its spike.



steps ***its*** ***six***

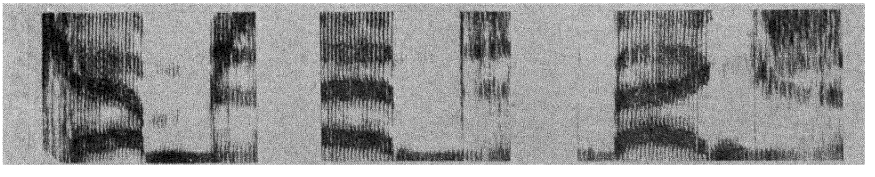
The stop-fricative combination usually occurs in the final position, or between two words as in *up stairs.* There is an essential difference in the timing and stress between blends occurring initially or finally in a syllable, and blends occurring between syllables or words. In the final combinations illustrated above, the stop spike-fricative fill patterns resemble those previously described for the **tʃ** (*church*) combination. Although the stop spike is quite evident at the beginning of the fill, the dark area of the spike may seem to disappear in the striations of the fill. Generally, the **t** spike is more difficult to distinguish from the **s** fill than are the **p** or **k** spikes. It should be noted that due to the articulatory positions involved, the **p** spike and **f** fills are related in a way similar to the **t** and **s**, but the **k** pattern usually may be distinguished from the fill of the fricative with which it is combined. In stop-fricative combinations occurring between syllables or words, the stop sound frequently is imploded and thus is identified by the nature of its stop gap and by its influence.



saved *bathed* *dazed*

The voiced fricative-voiced stop combination occurs frequently in the final position of words or syllables but, in English, is not used at the beginning of syllables. Such combinations also occur between syllables in continuous speech, and illustrations may be found in some of the sentence patterns in this and subsequent units. Differences in timing and stress may be observed between the monosyllabic blend patterns above and the between-syllable blend patterns contained in the sentences. Duration of the individual sound pattern, darkness of the pattern, position of the hub, and extent of transitional movement, portray the variations in speed and precision of articulation, and in speech rhythm and stress. Because this volume is confined to illustration and explanation of the elementary and basic principles of visible speech, it is not possible to discuss many of the modifications and variations that have been observed in continuous speech. Further explanation and illustration must be reserved for future texts.

In the words illustrated above, vocal cord and cavity modulation continue throughout the **v-d**, **ð-d**, and **z-d** patterns. Although the stop gap terminates the fricative resonances rather abruptly, a portion of the transition from the fricative hub towards the stop hub is visible in the pattern. Individual sounds in the voiced fricative-voiced stop blends are somewhat easier to identify than their voiceless equivalents. The voiced fricative may be recognized by its influence on the preceding sound and by the way in which it is influenced by the stop. The stop may be identified by its influence on the fricative and, in continuous speech, by its influence on the following sound. The **v**, **ð**, and **z** in the patterns above, each influences the vowel **ɪ** in the diphthong **eɪ** in characteristic fashion according to its hub. In addition, each of the fricatives is influenced by the **d**.



cabs *adds* *bags*

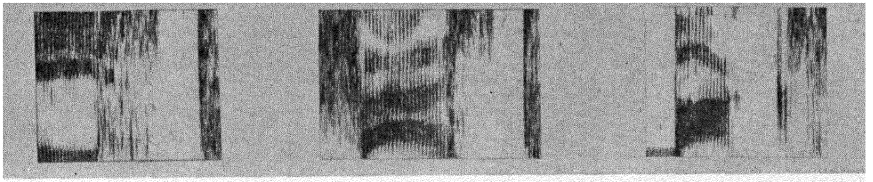
Voiced stop-voiced fricative combinations occur most frequently in the plural and possessive forms of words. As in the equivalent voiceless stop-fricative blends, the stop appears to be exploded into the fricative, and also the same transitional relationships may be observed between the hubs or the articulatory positions of the combined sounds. The continuation of vocal cord modulation is evident in the voice bars for the stop and fricative, but the change in type of cavity modulation may be seen in the width of the voice bars and in the resonance positions as the closed-cavity stop modulation shifts into the partially-closed-cavity frictional modulation. Identification of the stop sound generally is based on its influence both on the preceding sound and on the fricative. The pattern of the stop usually is difficult to differentiate from that of the fricative except in the case of the **g**. The fricative may be recognized by its pattern, by its transition from the stop hub, and by its influence on the sound which follows. The voiced stop–**z** combination is very frequent in speech, and many people partially unvoice the **z**, especially when it occurs before a voiceless sound or in a final position in a sentence. Patterns of differences in the relative proportions of frictional and vocal cord modulation have been presented in the previous units.

The particular combinations of stops and fricatives illustrated are among the common blends, but there are many other combinations that occur between syllables or words in continuous speech. More complex combinations of stops and fricatives often are found in words like *sixth*, *mists*, *midst*, and *adjusts*, or in the inter-word combinations in speech. The general principles stated for fricative-fricative, stop-fricative, and fricative-stop combinations may be applied to the identification of all such combinations. Differences in rate, precision of articulation, rhythm, and stress, modify both the articulatory positions and the transitional movements. These modifications are apparent in the patterns; and, if sufficiently extreme, they may make identification of the patterns difficult. Such articulatory variations are manifested in both the aural and visual patterns, but some part of our recognition of the aural variations may be ascribed to the extent of our auditory associational background. Whether greater familiarity with the visible speech patterns and improvements in the portrayal of the transitional material and of the separate sounds will make identification of the visible pattern variations equally possible, awaits further experimentation. For easy recognition, at the present time, sound combinations must be articulated carefully and at a rate no faster than a speaking rate of about 130 words a minute.

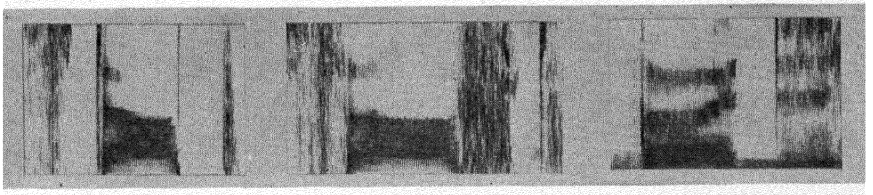
Review words

fifth, page 157
speak, page 159
sky, page 159
best, page 160
steps, page 160
cuffs, page 157
stay, page 159
hasp, page 160
ask, page 160
its, page 160
gives, page 157
saved, page 161
dazed, page 161
adds, page 161
six, page 160
bathes, page 157
bathed, page 161
cabs, page 161
bags, page 161

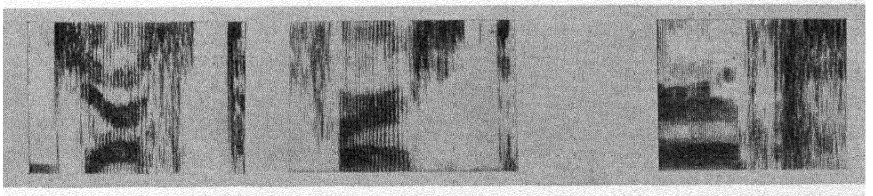
New words to learn



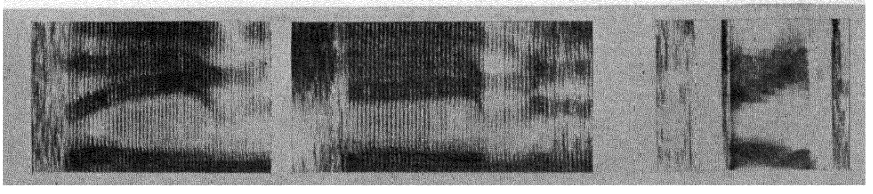
east *fast* *box*



stop *soft* *goods*

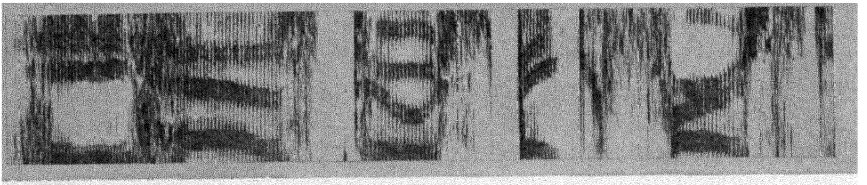


just *first* *earth's*

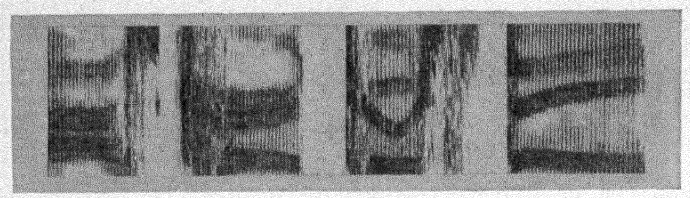


saves *shaves* *state*

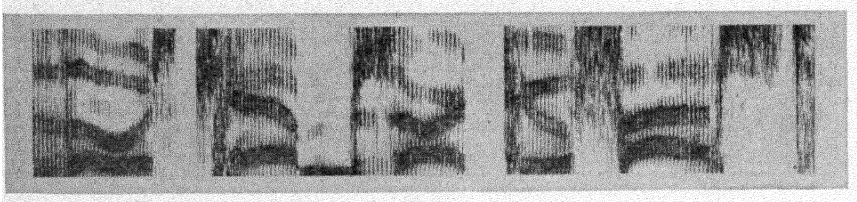
Drill phrases and sentences



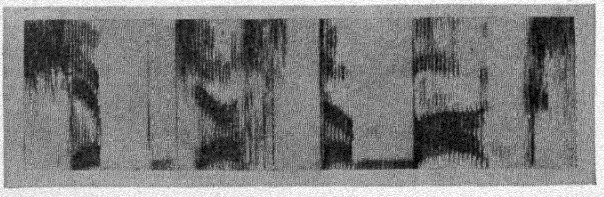
He has to speak first



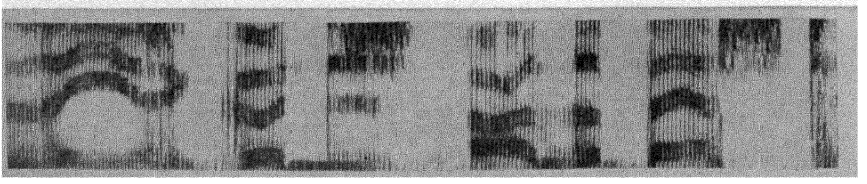
Ask her to stay



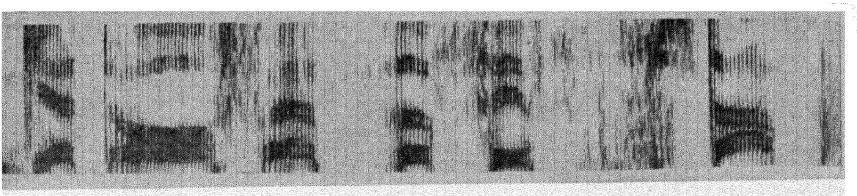
Those cabs are too fast



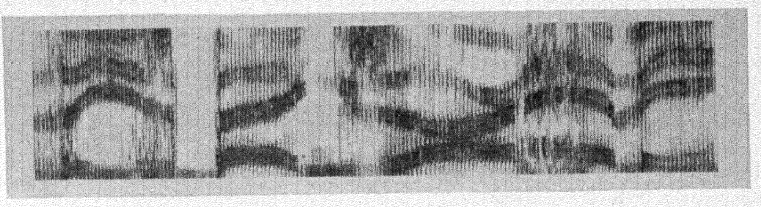
Ship just a box



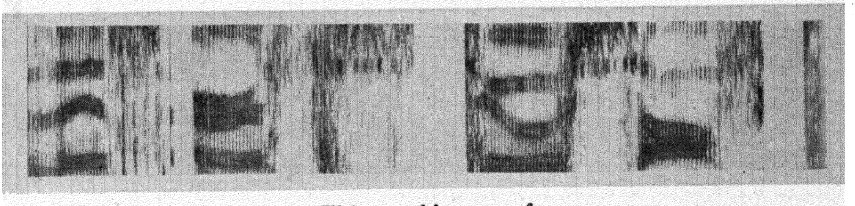
These goods are the best



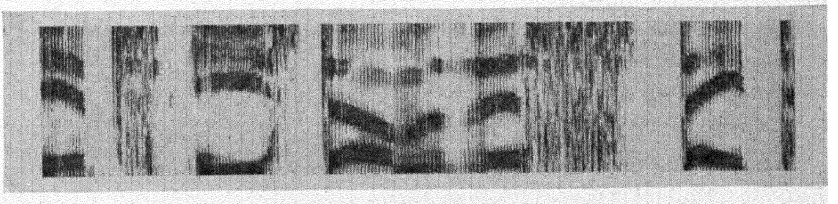
Get off at the fifth stop



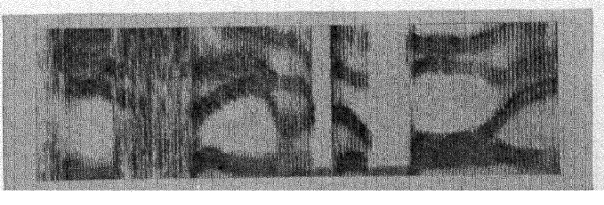
These bags are heavy



This earth's too soft



It's east of this state



He saved a boy

Chapter 10

Reading the Vowel-Like Sounds—The Nasal Sounds

There are only three sounds in English that are characterized primarily by nasal resonance. These sounds, m (*me*), n (*no*), and ŋ (*sing*), have been classified as a part of the larger unit of vowel and vowel-like sounds. However, they are included in a separate lesson chapter because their aural and visible patterns are differentiated from those of the other vowel-like sounds by the results of combined oral and nasal cavity resonances. Patterns of this nasal group are recognized by two primary clues:

1. Resonance bars produced by the continuous modulation of voiced breath in the various vocal cavities.
2. Voice bars on the baseline resulting from combined open and closed cavity modulation.

Unlike the stop and the fricative sounds, all of the vowel-like sounds are produced with vibratory or vocal cord modulation; and, therefore, there are no voiceless cognate sounds occurring within this classification.

The nasal sounds are continuant sounds produced in the combination of closed mouth and open nasal cavities. Vocal cord-cavity modulation results in the vowel-like appearance of the resonance bars, but the combined open and closed cavity modulation produces a voice bar along the baseline which differentiates the nasal from the vowels and other vowel-like sounds. The number, darkness, and resolution of the bars depends upon the degree of stressing and upon the relative proportions of closed and open cavity modulation. An unstressed nasal may be portrayed by a voice bar alone, while an overstressed nasal may produce a vertical pattern of overlaid bars which have a diffuse or smeared appearance. That particular resonance area which occurs in the region of the hub of a nasal sound may or may not be present in the pattern. It appears to be associated with the balance or proportion of closed and open cavity modulation.

When a nasal is said in combination with a non-nasal sound, the opening or closing of the nasal port may be made abruptly or gradually. If the soft palate is moved abruptly to close or open the nasal port, the pattern displays a single spike between the nasal and the combined sound. The spike resembles a stop spike and especially a glottal stop spike. Sudden termination of both cavity and vocal cord modulation, or rapid shift from an articulatory position, also may produce a stop spike. A more gradual transition between nasal and non-nasal sounds is portrayed by curves of the resonance bars similar to those occurring between voiced fricative and vowel combinations. Frequently, in rapid speech, the non-nasal sound occurring before or after

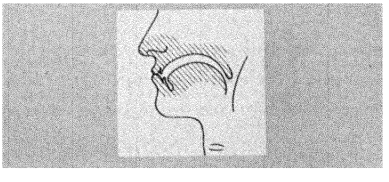
a nasal may be partially nasalized due to the gradual initiation and termination of the nasal resonances.

There are some minor differences observable between the patterns made by the two experimental spectrographs used in preparing the illustrations for this and the following lesson chapter. These differences, resulting from variations in circuit design and in filter characteristics, are apparent in the juncture of vowel bar 1 with the combined sound. In one spectrograph, that juncture and the influence on bar 1 appears to be curved; while in the other, the juncture and influence are square in appearance. Since the influence on bar 1 appears to be due primarily to change in type of modulation, all of the voiced sounds influence bar 1 of any given vowel in the same way and the extent of influence of the voiceless sounds also is dependent upon the relative position of bar 1 of the vowel, rather than upon the particular influence of the combined sound. Therefore this particular difference in the bar 1 patterns produced by the two spectrographs is not significant for reading, although it may be of interest in phonetic studies and in speech correction.

Each of the nasal sounds is discussed in some detail in the following units of this chapter. As in previous lesson chapters, the transitional influence is illustrated by words containing representative vowels. Unit 4 contains a summary and comparison of the patterns, hubs, and influences by which the m, n, and ŋ may be recognized and differentiated.

Unit 1 — Reading the m (me)

How the m is made



Profile of m (*me*)

The m is produced by:

1. Opening the port between the nasal and throat (pharyngeal) cavities.
2. Maintaining a closure at the lips as in the stop phase of the b (*be*).
3. Emitting a continuous stream of voiced breath through the nasal cavity.

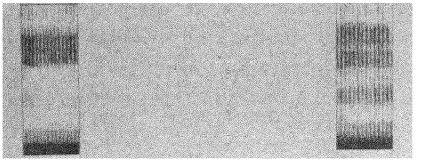
In producing the m, the articulators take positions resembling those that they assume for the preliminary voicing of the b (*be*), with the exception of the soft palate. Since the nasal port is opened by the lowering of the soft palate, the voiced breath

for **m** is resonated in both the closed mouth cavity and the open nasal cavity in addition to the other cavities of the vocal tract. Therefore, the **m** is characterized by a combination of vocal cord, closed cavity, and open cavity modulation. Phonetically, the **m** has been described as a bi-labial nasal continuant. The similarity in articulatory positions for the **p**, **b**, and **m** results in somewhat similar influences and hub positions, as will become evident in this Unit. However, due to the difference in cavity modulation between the **b** and the **m**, and to the greater variability of the **m** which is introduced by the added resonances, the **m** pattern, transitions, and hub are not fixed as markedly as those of the **p** and **b**.

The m may be spelled as:

m (*may*)	*mm* (*hammer*)	*mb* (*comb*)	*gm* (*diaphragm*)
me (*same*)	*mn* (*hymn*)	*lm* (*calm*)	*chm* (*drachm*)

Visible patterns of the m

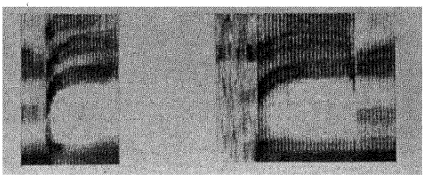


m **m**
unstressed ***stressed***

Although the **m** is a continuant and may be said alone, it has been spoken, for the above illustration, before the neutral vowel ʌ (*up*) in order that the patterns of the various nasals may be compared. Combined vocal cord, closed mouth, and open nasal cavity modulation for the **m** produces horizontal bars that resemble in form the bars produced by the vowels. The bar on the baseline, like the voice bars for the voiced stops and fricatives, is believed to result primarily from vocal cord-closed cavity modulation. The voice bar for **m** is quite dark along the baseline, but there usually is a lighter resonance area just above it. This area appears to be characteristic of nasal resonance, and serves to widen the voice bar and to help in differentiating the nasals from the other voiced sounds. A stressed **m** may have one or more additional resonance bars, dependent upon the relative proportions of oral and nasal resonance. If the **m** is overstressed, the general appearance of the patterns may be diffuse without distinct resolution of the resonances. Generally, the stressed **m** has a resonance area directly above the voice bar, another resonance which varies somewhat in position in the lower central part of the pattern, and a resonance area that resembles a double

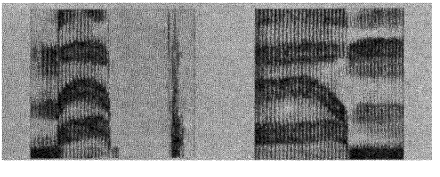
bar at about the same position as bar 3 of the vowel æ (*at*). Unstressing of the m may be portrayed either by the voice bar alone, or by the voice bar plus one or more weakly defined resonances. The resonance occurring in the hub area may or may not be visible, since it appears to be dependent upon the relative proportions of the various cavity resonances combined for the m.

How the m influences vowel patterns



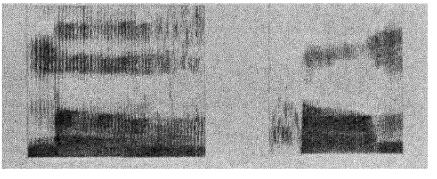
me ***seem, seam***

In comparing the m bars with the vowel bars, the positions of the bars and their appearance should be observed carefully. Since the m hub is towards the baseline and the i (*eve*) hub is near the top of the pattern, the direction of the transition in *me* and *seem* may be predicted. As in the p–b influence, bars 2, 3, and 4 of the vowels i(*eve*) and ɪ (*it*) generally show similar transitional curves, although bar 2 displays the most marked transitional movement. The transition between the m and the vowel i (*eve*), as has been discussed previously, may be made more or less abruptly, depending upon the speed and precision of articulation. The two word pattern above illustrate a difference in the speed of articulation. In *me*, the transitional period is more gradual and thus more apparent than it is in *seem*. Partial nasalization of the vowel would have been portrayed by nasal resonance superimposed upon the vowel bars. Although the nasal port was closed and opened rather rapidly in these words, as is evidenced by the sharp initiation and termination of the vowel resonances and by the brief spike between the nasal and vowel, there is an apparent widening of vowel bar 1 which is observed frequently in nasal-front vowel combinations. The drawn-out appearance, or "tail", at the ends of the vowel bars is characteristic of the m and also of the p–b.



map ***am***

The distance between the m hub and the hub of the vowel æ (*at*) is not as great as that between the m and the i (*eve*); therefore, the extent of transitional movement is less. In *map*, the similarity between the m and p hubs and influences is evident. Bar 3 of the æ is not influenced noticeably by the m as the positions of the third bars of both sounds are similar. Differences in rate of articulation and in degree of stressing may be observed in the patterns above. The slope of bar 2 of the vowel in *am* is more gradual than in *map*, indicating slower articulatory movement; and also, the nasal resonances are not connected as closely with the vowel resonances as they are in *map*. Four resonance areas, including a rather wide third resonance, indicate that the final m was stressed more strongly than the initial m. Comparison of the positions of the m resonances in the different illustrations demonstrates the fixed nature of the m, and the m hub and influence.



move ***whom***

There is not much influence exerted by the m on the bars of the vowel u (*boot*). The m and u hubs both are located towards the baseline, and, therefore, the transitional movement between the sounds is slight. Although these words were spoken by individuals with voices of very different fundamental pitch, the primary m resonance areas are similar in frequency position. Thetransition between vowel and nasal resonance is more gradual in *whom* than in *move*. The same speaker on different occasions may produce a considerable range of variations in the speed and nature of nasal — non-nasal transition. Many of these variations are non-significant in regard to aural intelligibility or to the aural recognition of a particular syllabic combination, and for reading visible speech the differences in the m patterns and transitions in the preceding and following illustrations are regarded as variations within the visible m phoneme. For the purposes of the phonetician and speech specialist, more specific interpretation of the significance of the various resonances must await further experimental investigation.

The m hub is located towards the baseline; and, therefore, the m influence is most marked on the vowel i (*eve*) and is least noticeable on the vowels ʊ (*foot*) and u (*boot*). There is a gradual decrease in extent of influence, although the direction of the transition is the same, as the vowels are said in series from i (*eve*) to u (*boot*). There usually is a resonance area at about the position of the third bar of the vowel æ (*at*), with

which the third bar of the vowels seems to join. Bar 1 of the vowels is joined to the voice bar for the **m** in similar fashion to the junction of vowel bar 1 and the voice bar of other voiced sounds. There are many variations possible in the nature of the transitions to and from the **m** produced by differences in the speed and precision of articulation and by changes in the relative proportions of the various resonances present.

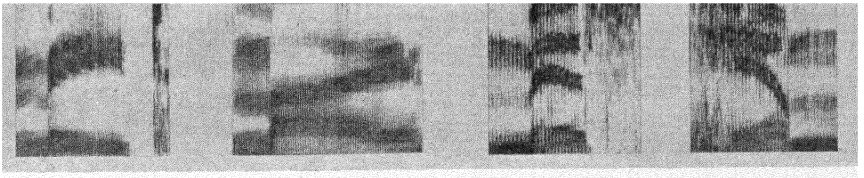
Read the **m** *by:*

1. The position and appearance of the **m** bars.
2. The influence on the vowels.

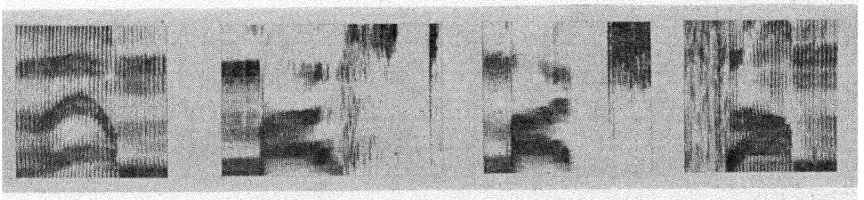
Review words

me, page 169	***move***, page 170	***am***, page 169
map, page 169	***seem, seam***, page 169	***whom***, page 170

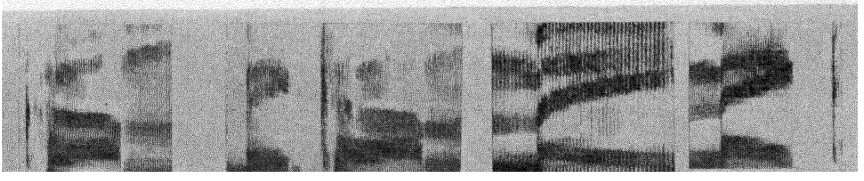
New words to learn



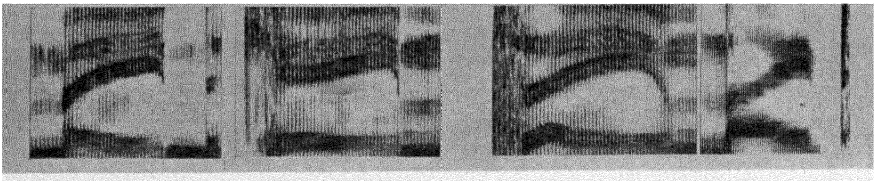
meet, meat ***my*** ***miss*** ***him***



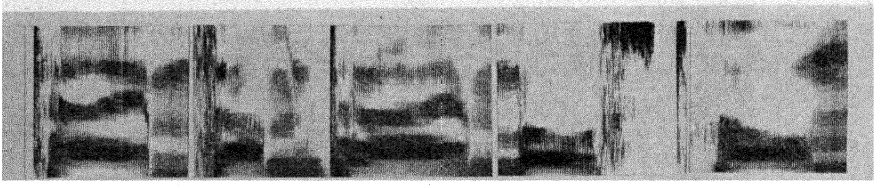
them ***must*** ***much*** ***some***



come ***become*** ***may*** ***make***



made *came* *same* *might*

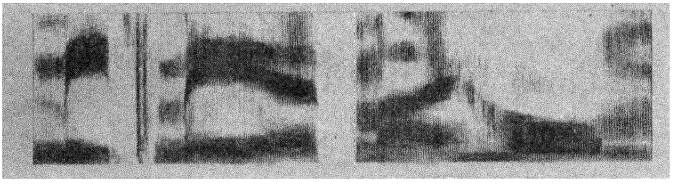


time *sometime* *most* *home*

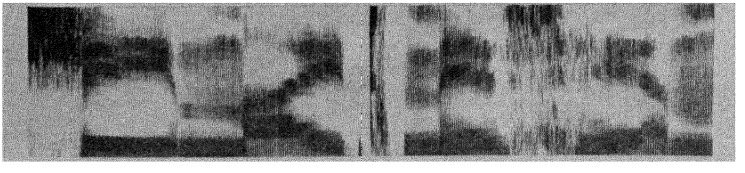
Additional words for practice in visualization

bottom *mouth* *famous* *firm* *game*

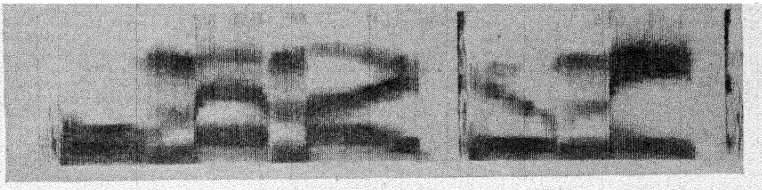
Drill phrases and sentences



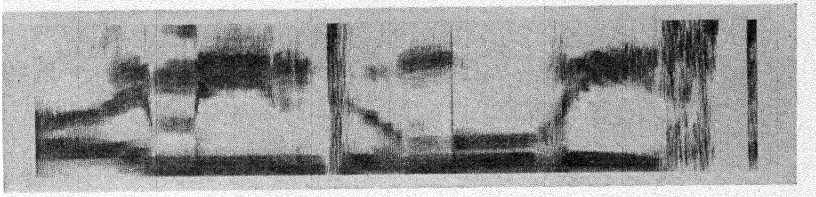
Meet me at my home



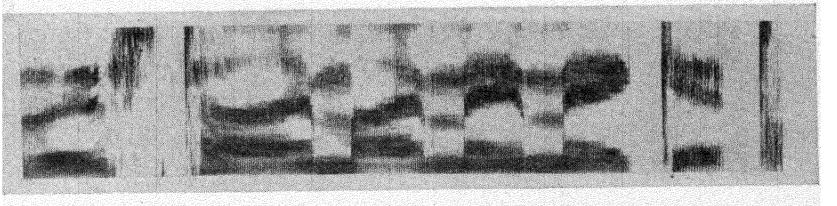
She might miss him



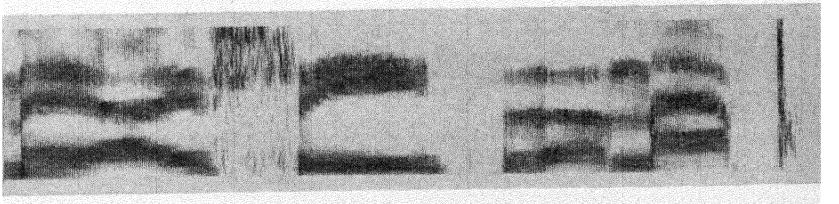
Whom am I to meet?



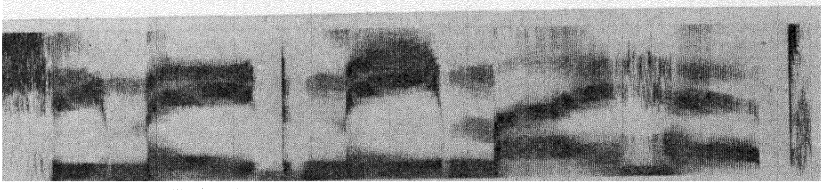
I mean to move east



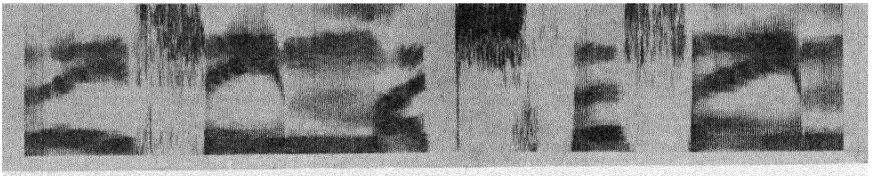
This time I may make it



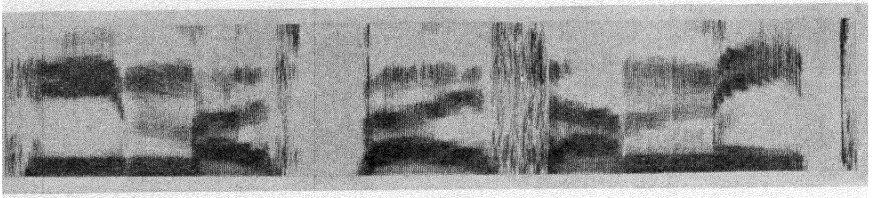
May I see the map?



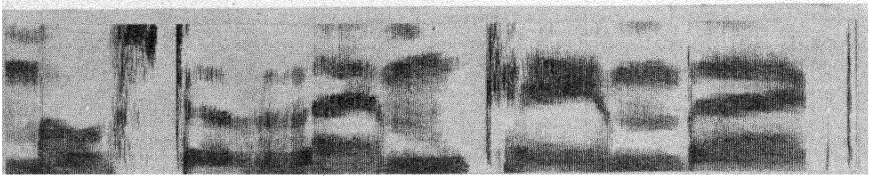
She made me my hat



They seem much the same



He must buy some meat



Most of them came back

Unit 2 — Reading the n (no)

How the n is made

Profile of n (*no*)

The n is produced by:

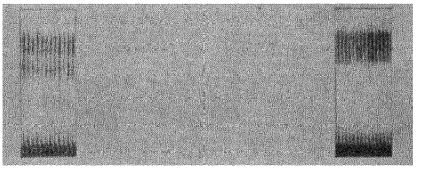
1. Opening the port between the nasal and throat (pharyngeal) cavities.
2. Maintaining a closure in the oral cavity as in the stop phase of the d (*day*).
3. Emitting a continuous stream of voiced breath through the nasal cavity.

As indicated in the profile above, the articulators used to produce the n take positions resembling those that they assume for the preliminary voicing in the stop phase of the d (*day*), with the exception of the soft palate. The opening of the nasal port permits resonation of the voiced breath in both the closed mouth cavity and the nasal cavity in addition to the resonation occurring in the other cavities of the vocal tract. The n is characterized by a combination of vocal cord, closed cavity, and open cavity modulation. Phonetically, it has been described as a lingua-alveolar nasal continuant. Since the articulatory positions for the t, d, and n are somewhat similar, their influences and hub positions are related, as may be observed from the illustrations in this unit. Like the t and d, the n is classified as a relatively fixed sound. However, the differences in cavity modulation between the d and n, especially the introduction of nasal resonance, produce a greater variability of the n, as evidenced in the n pattern, transitions, and hub.

The n may be spelled as:

n (*no*)	*nn* (*inn*)	*gn* (*sign*)	*mn* (*mnemonics*)
ne (*lane*)	*kn* (*know*)	*pn* (*pneumatic*)	

Visible patterns of the n



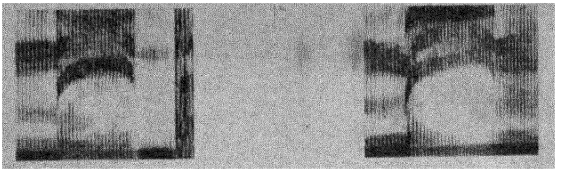
n
unstressed

n
stressed

Horizontal bars, resembling the bars produced by the vowels, result from the combined vocal cord, closed and open cavity modulation for the n. Since the n is a relatively fixed sound, it has been spoken before the neutral vowel ʌ (*up*) for purposes of comparison with the m and the ŋ. The voice bar, like that for the m is generally dark along the baseline with a shadow-like resonance area a little above it. Stressing may produce one or more additional resonance bars, which are believed to be associated with the relative proportions of oral and nasal resonance. Overstressing is portrayed by a diffuse pattern containing many poorly defined resonance areas. The stressed n often has a double resonance in about the same position as the double resonance

for the **m**, and also may have a resonance area a little below the bar 2 position of the vowel **æ** (*at*) similar to the resonance occurring in the **m** pattern. The particular resonance which occurs in the hub area for **n** is not always visible in the stressed **n** pattern and usually is absent in the unstressed **n**. It is believed to be related closely to the relative proportions of the various cavity resonances present in the **n**. Isolated patterns of the **m** and **n** are rather difficult to differentiate, since their second resonance bars do not regularly portray their hub positions. However, the influences of the **m** and **n** hubs, as shown in the direction and extent of the transitional patterns, may be differentiated quite readily.

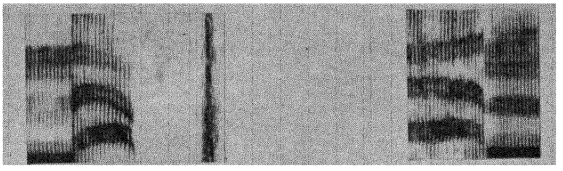
How the n influences vowel patterns



need, knead ***mean, mien***

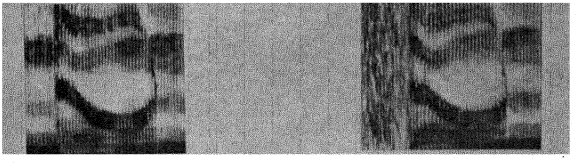
The hub of the **n** is in the area of the second bar of the vowel **æ** (*at*), but may vary slightly above or below this area in relation to the sound with which the **n** is combined. In the words *need* and *mean*, the direction of the transition between the **n** hub and the **i** (*eve*) hub is quite evident and illustrates the similarity between the **n** and the **d** hubs. Bar 3 of the vowel may be observed to repeat in lesser fashion the influence of **n** on bar 2. The pattern of the word *mean* illustrates rather clearly the similarity in the patterns of the **m** and **n**, and also the differences in their hubs and influences. In examining the patterns above, some difference in the type of transition between the vowel and the **n** may be noted. The transitional **n**-vowel movement was more gradual in *need* than in *mean*, and the **n** pattern itself is more clearly defined in *need*. Such differences in pattern and transition should not be assumed to be characteristic of the initial or final syllabic position of the nasal but rather, until further evidence is available, should be considered as variable factors related to the speed and precision of articulation, to stress, and to the nature of the cavity modulation. In both of the word patterns above, although there is a clear distinction between the initiation and termination of the nasal and the vowel, some additional resonance may be observed in the vicinity of vowel bar 1. This is a frequent occurrence in vowel bar 1 of **i** (*eve*), **ɪ** (*it*), **eɪ** (*say*), **u** (*boot*), and **ʊ** (*foot*).

In the patterns of *nap* and *an*, the hub of the **n** is slightly below the hub of the **æ** (*at*). The pattern of the word *nap* may be compared with that of the word *map*, and *an*



nap *an*

with *am*, as presented in Unit 1 of this chapter. Although the directions of the m and n transitions with this vowel are the same, the extent of the m transition is greater than the n since the m hub is further from the vowel hub. Bar 3 of the vowel is not influenced to any great degree by the n, while bar 1 is influenced in much the same way that it was when combined with other voiced sounds. The final n in *an* was spoken with more stressing than the initial n in *nap* as is evident in the definition of the resonance areas.



noon *soon*

Since the n hub is approximately at the middle of the pattern while the u (*boot*) hub is near the baseline, the transitional movement from n to the vowel is like that of the d (*day*) combined with the same vowel as in the word *do*. The s, which is produced from an articulatory position related to the n and d, may be observed in the word *soon* to have an initial influence on the vowel resembling that of the n. In both patterns, the influence of the final n is not as marked as that of the initial n, illustrating the variability which may be observed in the n hub and also a modification of the vowel u which is noted frequently when it is spoken after the n, d, t, and l.

Because the n hub is located towards the middle of the pattern, resembling the t–d hub, it influences bar 2 of the vowel i (*eve*) and the vowel u (*boot*) almost equally. There is a gradual decrease in the extent of the transitional movement as the n is combined with the vowels said in series from i (*eve*) to æ (*at*), and a gradual increase in transitional movement with the vowels in series from ɑ (*father*) to u (*boot*). The direction of the transitional movement may shift either with the vowel æ (*at*) or with the vowel a (*ask*). The n hub appears to be more variable than the t–d hub, but rarely shifts below the hub of the vowel a (*ask*). Bar 1 of the vowel is joined to the n

voice bar in the same way as it is for the other voiced sounds. Variations in the nature of the **n** combined sound transitions are similar to those which have been discussed previously for the **m**.

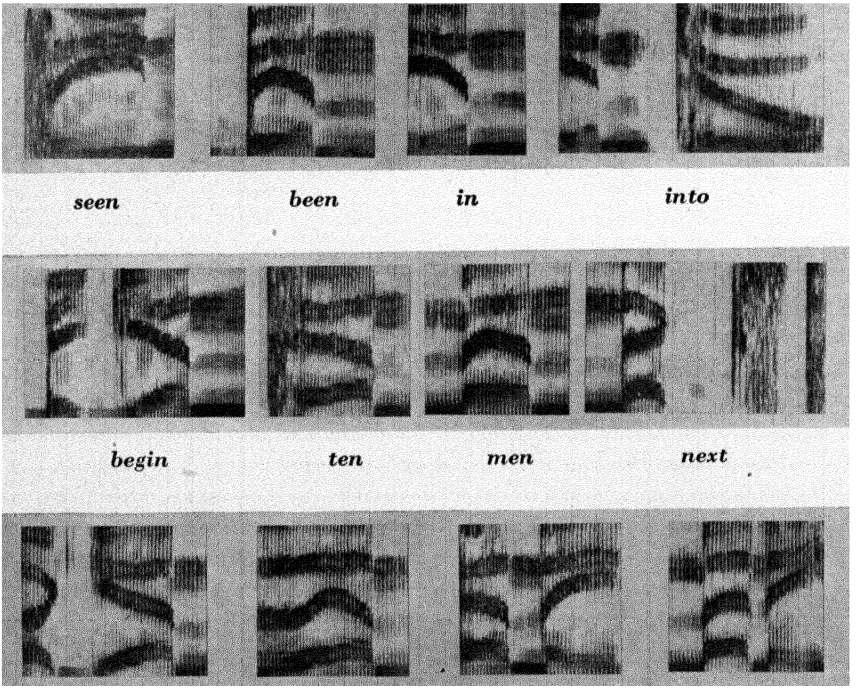
Read the **n** *by:*

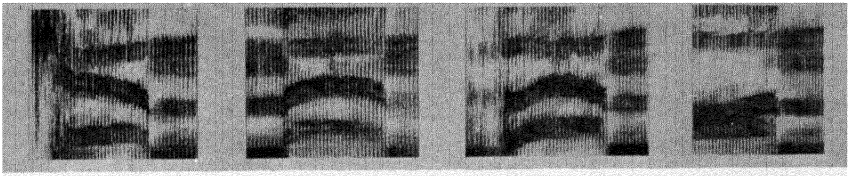
1. The position and appearance of the **n** bars.
2. The influence on the vowel.

Review words

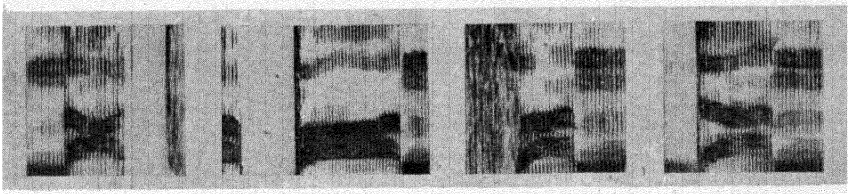
need, knead, page 176	***soon***, page 177	***nap***, page 177
mean, mien, page 176	***noon***, page 177	***an***, page 177

New words to learn

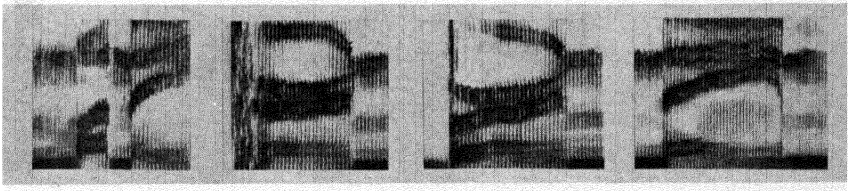




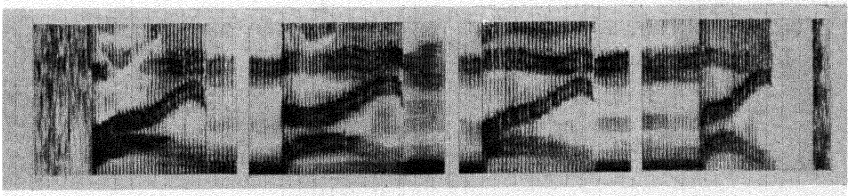
can *man* *than* *on*



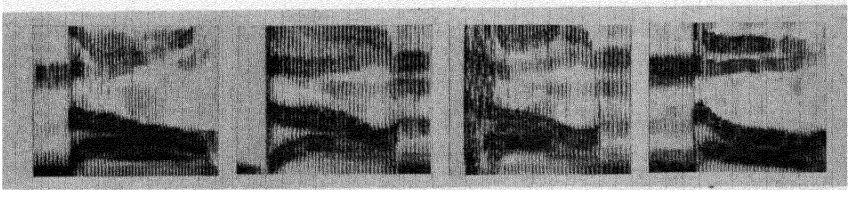
not *upon* *son, sun* *done*



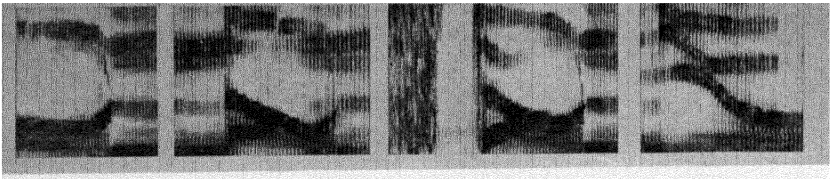
money *turn* *burn* *name*



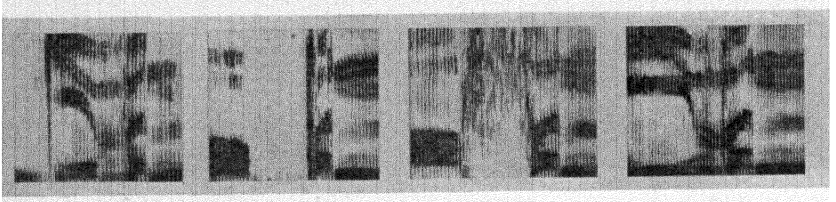
fine *nine* *mine* *night*



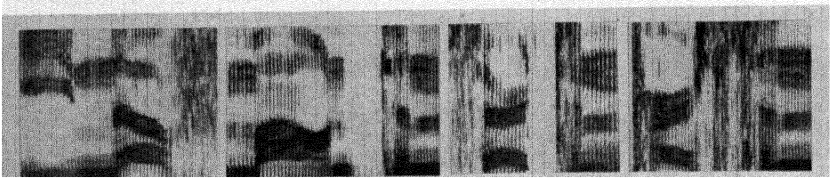
now *down* *town* *no, know*



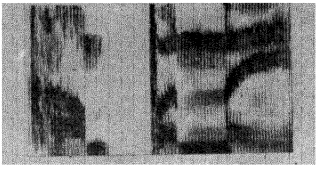
own *known* *stone* *new*



given *open* *often* *even*



enough *mountain* *certain* *person*



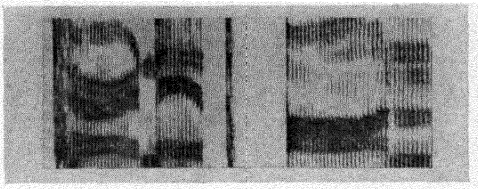
company

Additional words for practice in visualization

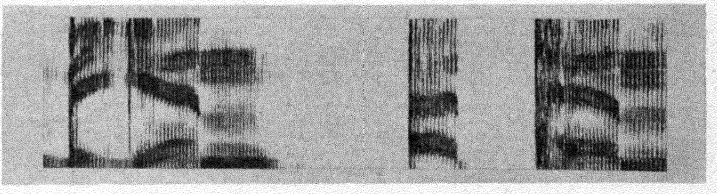
knee	*thin*	*minute*	*finish*	*business*
instead	*indeed*	*captain*	*journey*	*neck*
nest	*enemy*	*pen*	*enjoy*	*began*
cannot	*gone*	*moon*	*none*	*common*

station	*ocean*	*condition*	*pain*	*gain*
chain	*nation*	*contain*	*shine*	*nice*
sign	*join*	*nose*	*notice*	*note*
bone	*happen*	*sudden*	*season*	*heaven*
seven				

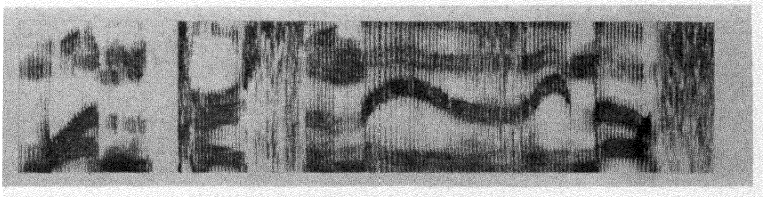
Drill phrases and sentences



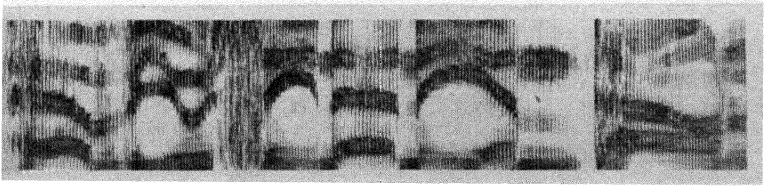
Turn it on



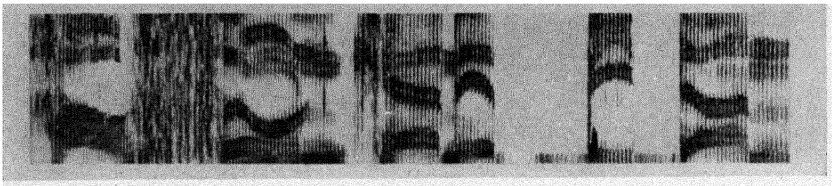
Begin at ten



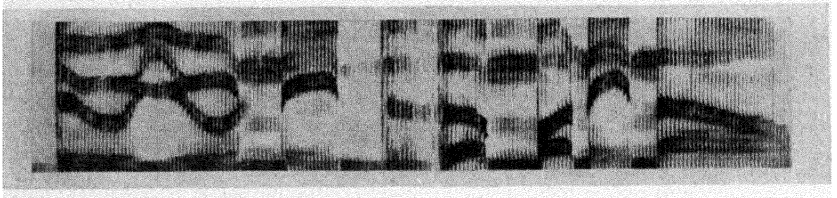
One person is enough



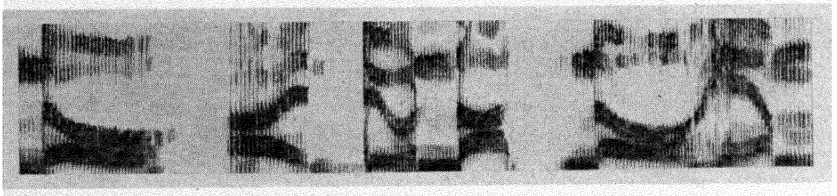
Have you seen any in town?



How soon can it be done?



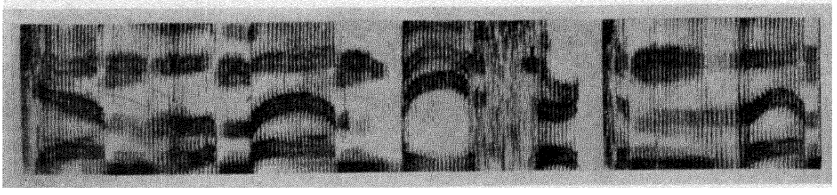
Do you need the money now?



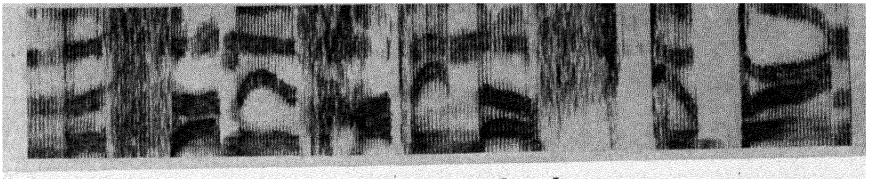
No, I do not know him



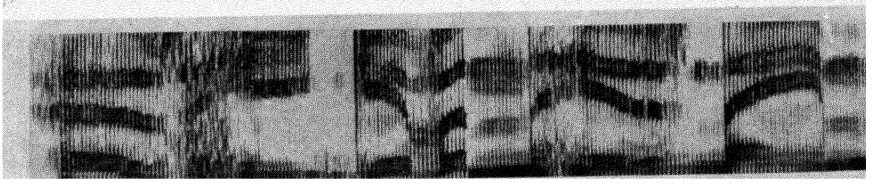
Meet me downtown tonight



Can the man be certain then?



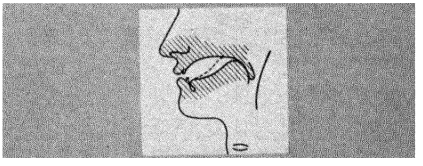
The sun is hot enough to burn



Has he given his name?

Unit 3 — Reading the ŋ (sing)

How the ŋ is made



Profile of ŋ (sing)

The ŋ is produced by:

1. Opening the port between the nasal and throat (pharyngeal) cavities.
2. Maintaining a closure by placing the back of the tongue against the posterior hard palate and part of the soft palate as in the stop phase of the **g** (*go*).
3. Emitting a continuous stream of voiced breath through the nasal cavity.

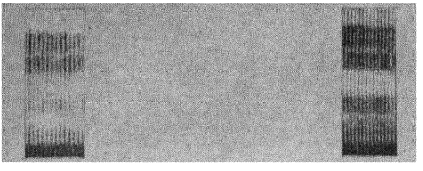
Like the **g** (*go*), the ŋ is classified as a variable sound, since it may be produced from a number of slightly different points of lingua-palatal contact depending upon the sound with which it is combined. The range of articulatory positions for ŋ are indicated in the profile above. Usually, it is produced towards the front with the front vowels, towards the back with the back vowels, and in intermediary positions with other sounds. Characteristic variations in ŋ articulation are illustrated in the patterns in this unit. The ŋ is described phonetically as a nasal lingua-palatal continuant, and

it is produced by a combination of vocal cord, closed cavity, and open cavity modulation. Due to similarity in the articulatory positions for the **k**, **g**, and ŋ, their hub positions and influences are quite similar, although the nature of the modulation used for the ŋ results in a somewhat greater variability of pattern for the ŋ than for the **k**–**g**.

The ŋ may be spelled as:

ng (*sing*) *ngue* (*tongue*) *n* (*ink*; longer, *zinc*)

Visible patterns of the ŋ

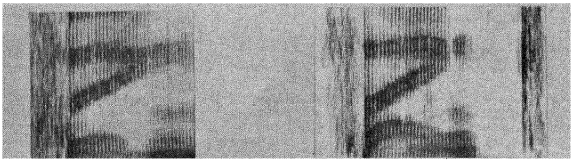


ŋ ŋ
unstressed ***stressed***

The resonance bars portraying the combined vocal cord, closed cavity, and open cavity modulation for the ŋ resemble those produced by the **m** and **n** and, to a lesser degree, those characteristic of the vowels. Both the voice bar and the resonance area directly above it are characteristic of all of the nasals rather than of the ŋ specifically. Nor is there much difference in the position of the wide or double resonance area towards the top of the pattern. With the front vowels, the ŋ top resonance area may be slightly higher than that of the **m** or **n**, but in most cases the position is so similar that differentiation is difficult. The stressed ŋ may be portrayed by the voice bar with resonance above it, a second resonance bar a little below the middle of the pattern, and a double resonance area towards the top in about the position of the second and third ɪ (*it*) bars. An unstressed ŋ may be represented by a voice bar alone, or by a combination of voice bar and one or more weaker resonance areas. As in the other nasals, the hub area is not always portrayed by a resonance bar, and the positions and relative intensities of the resonances may be altered by varying the proportions of the various cavity resonances.

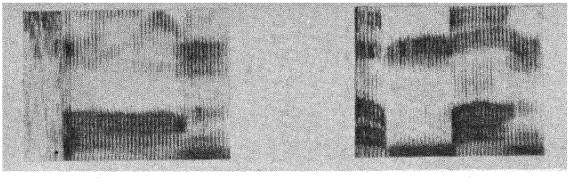
How the ŋ influences vowel patterns

The ŋ is not used as an initial sound in English, and it occurs medially and finally with only a few of the vowel sounds, as will be illustrated in the words presented in this unit. Since the hub of the ŋ is variable like that of the **k**–**g**, it usually is a little higher with the vowel ɪ (*it*), than with the æ (*at*). The influence of the ŋ, especially on the front vowels, curves bars 2 and 3 of the vowel together characteristically, in the way that the **g** does with the same vowels in such words as *big* and *bag*. Bar 1 of the



thing *thank*

vowel joins the voice bar of the ŋ just as it does in the case of the other nasals and voiced sounds. In *thank*, the ŋ is shortened because of the k which follows, and usually it is relatively unstressed. It is interesting to observe the continuation of voicing in the juncture of bar 1 of the vowel and the ŋ, and to see the momentary break between the upper resonances of the vowel and the ŋ. As has been discussed previously, this transition may vary from abrupt to gradual, depending upon the speed and precision of articulation and upon the relationship of the different cavity resonances.



song *among*

In the illustrations above, the hub of the ŋ has shifted down towards the hubs of the back vowel ɒ (*not*) in *song*, and the mid vowel ʌ (*up*) in *among*. Since the ŋ hub is near the vowel hub in each case, it has little influence on the second vowel bar; however, the ŋ hub may be slightly above or below the vowel hub, thus altering the direction of the influence. Observation of vowel bar 3 and of a weak resonance just below it, which is visible in *among*, shows some evidence of the characteristic ŋ influence on the upper vowel resonances. These upper resonances are not consistently visible in the patterns of the back vowels and, therefore, may not be used as reading clues. The difference between the m and ŋ influences on the same vowel is apparent in the curve up from m to bar 2 and then the curve up from vowel bar 2 to the ŋ. Comparison of the rather gradual vowel-nasal transitions in these illustrations with the transitions in *thing* and *thank* present examples of some of the possible variations.

The ŋ hub varies with the hub of the sound with which it is spoken and resembles that of the k–g. The characteristic rounding together of the upper vowel resonances is most evident in the front vowel-ŋ combinations. There is not as striking an influence visible in the back vowel-ŋ combinations, but it is readily differentiated from the influences of the other nasals.

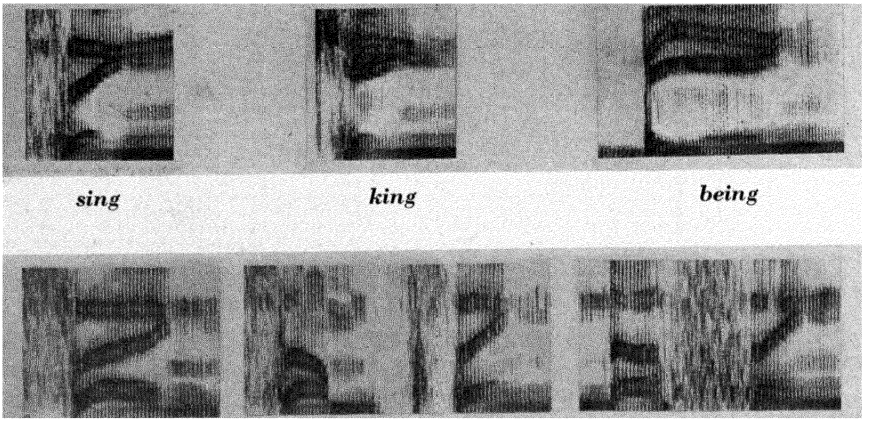
Read the ŋ *by:*

1. The position and appearance of the ŋ bars.
2. The influence on the vowels.

Review words

thing, page 185 *thank*, page 185 *song*, page 185 *among*, page 185

New words to learn



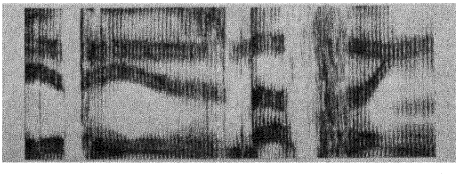
sing *king* *being*

sang *something* *nothing*

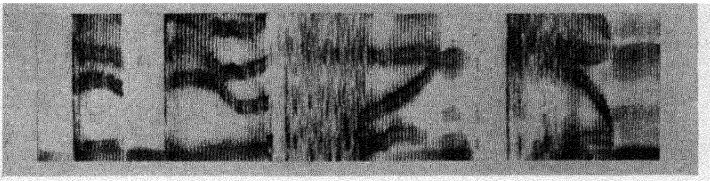
Additional words for practice in visualization

evening *going* *anything* *hang* *coming* *tongue*

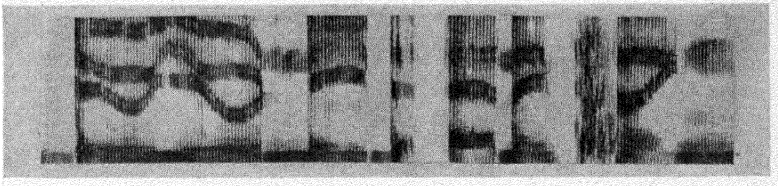
Drill phrases and sentences



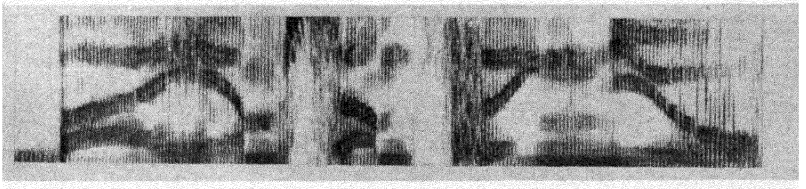
It is nothing



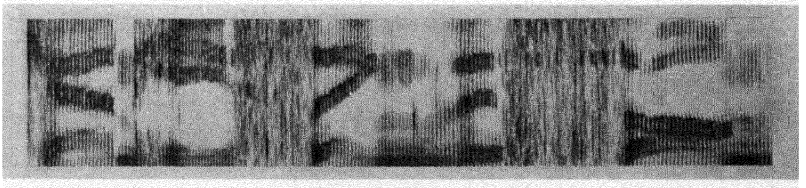
Did you thank him?



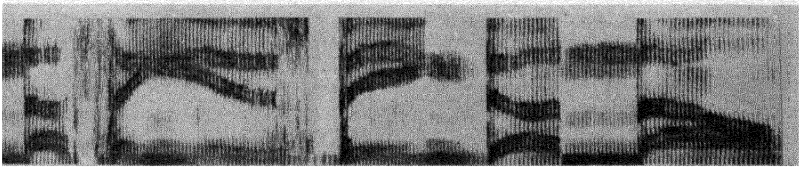
Do you need anything?



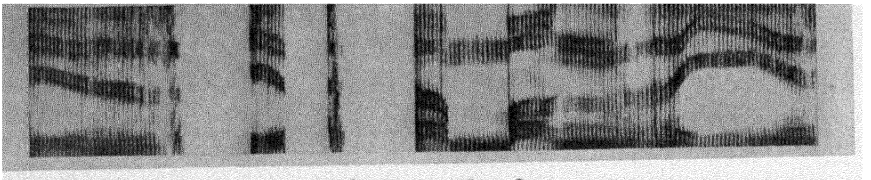
Buy him something new



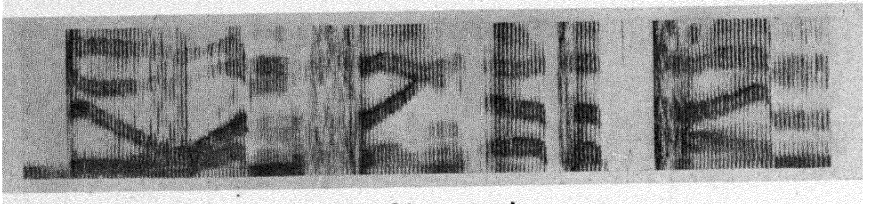
Can he sing this song?



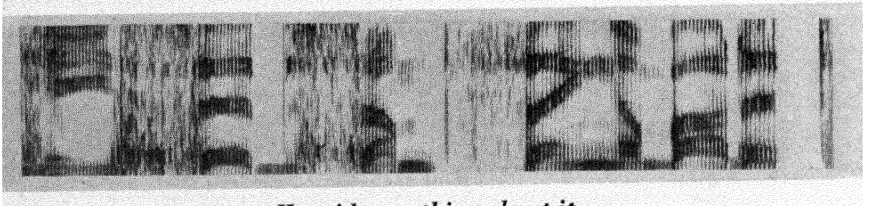
Nothing is being done now



Is it among these?



Do one thing at a time



He said something about it

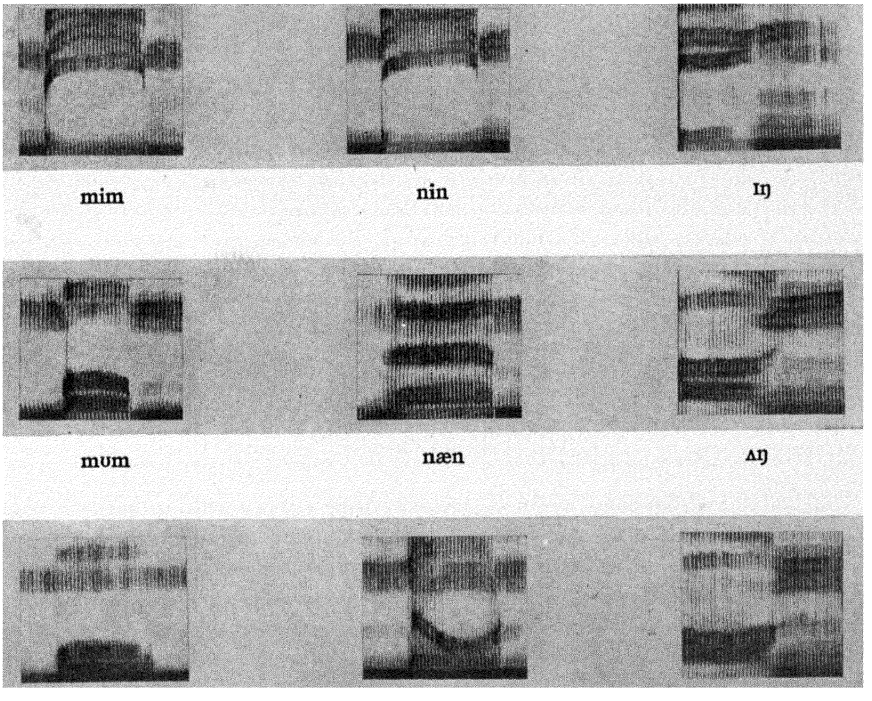


The King may speak soon

Unit 4 — Comparative Review of the Nasal Sounds

In order to compare the visible speech patterns and influences of the nasals, the characteristic clues by which they may be differentiated are presented in this unit. Although the three nasal sounds are produced by similar combinations of vocal cord-cavity modulation, their patterns may be differentiated because of the differences in the oral articulatory positions and the particular combinations of cavity resonances.

The figure below contains patterns of each of the nasals said with three different vowels. The **m** and **n** were spoken before and after these vowels, and the **ŋ** after the vowels, since it occurs normally in that syllabic position. The vowels **ɪ** (*it*), **ʌ** (*up*), and **ɒ** (*not*) were said with the **ŋ** to demonstrate the variability of the **ŋ** hub. Both the **m** and **n** were spoken with the **i** (*eve*) and **u** (*boot*). In addition, the **m** appears with the **ʊ** (*foot*), and the **n** with the **æ** (*at*), since the hub of each of these vowels approximates in position the hub of the nasal with which it was said.



Each nasal has a different hub position

1. The **m** hub is towards the baseline, similar to the **p–b** hub.
2. The **n** hub is near the middle of the pattern, similar to the **t–d** hub.
3. The **ŋ** hub is variable: towards the top with the front vowels, towards the middle with the mid-vowels, towards the baseline with the back vowels.

Each nasal influences the vowels differently

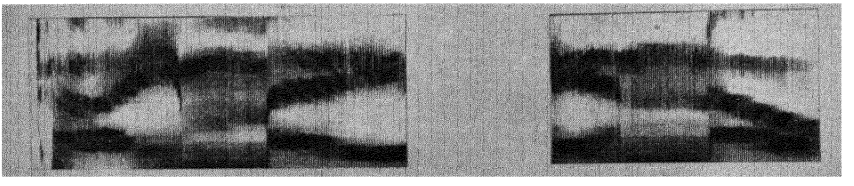
1. The **m** influences the front vowels most, with a gradual decrease in influence as the vowels are said in series from **i** (*eve*) to **u** (*boot*).
2. The **n** influences the most the back vowels **u** (*boot*) and **ʊ** (*foot*), and the front vowels **i** (*eve*) and **ɪ** (*it*). The direction of influence may shift either at the **æ** (*at*) or **a** (*ask*).
3. The **ŋ** influence is most evident on the front vowels and is characterized by a curving together of the upper vowel resonances.

The nasal pattern may be recognized and differentiated from the vowel and the other vowel-like patterns by the voice bar with a weak resonance directly above it, and by one or more additional resonances. Frequently there is a double resonance in about the bar 3 position of the vowel **æ** (*at*). Sometimes, a resonance which may shift slightly in position, appears in the lower half of the pattern. Although the individual patterns of the three nasals are difficult to differentiate in the 300-cycle patterns, their hubs and influences are distinctive and make ready differentiation possible in most cases. There is considerable variation possible in the nature of the transition between nasal and vowel, which may be related to the speed and precision of articulation and to the relative proportions of cavity resonances. As has been suggested in the previous comparative review units, the additional words for visualization may be visualized and then reproduced in graphic form for comparison with the visible speech patterns of the words illustrated in the lesson units. Those interested in learning to read visible speech may profit by extending the visualization process to idiomatic phrases, to sentences, and to paragraphs, and also by emphasizing in their study the kinaesthetic associations suggested in each unit.

Unit 5 — Reading Combinations of the Nasals, Stops, and Fricatives

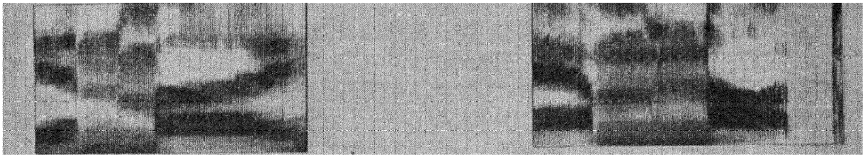
Combinations of the nasals occur frequently either within words, or between words and syllables in continuous speech. There are also many combinations of nasals with stops and fricatives; and, although the characteristics essential to the recognition of the individual sound patterns, hubs, and influences have been discussed in previous units, there are additional clues, basic to the recognition of combined sounds, which are illustrated and described briefly in this unit.

Reading combinations of the nasal sounds

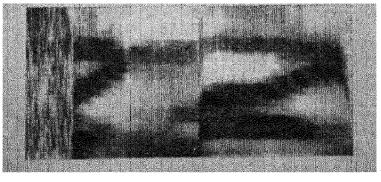


time may ***in no***

Combinations of nasals within a single syllable are infrequent in English, but often occur between syllables and words in continuous speech. When the same nasal sound occurs finally and initially in two adjacent syllables, as in the words *in no* or *time may*, the syllables may be spoken with a break in voicing between them, or they may be said continuously, in which case the nasal usually is spoken as a single prolonged sound. Sometimes there are shifts in cavity resonances or type of cavity modulation, and consequently a change in the number, positions, or relative predominance of the bars. However, the hub and influence of the nasal follow their usual pattern and thus provide a predictable clue to the recognition of the sound.



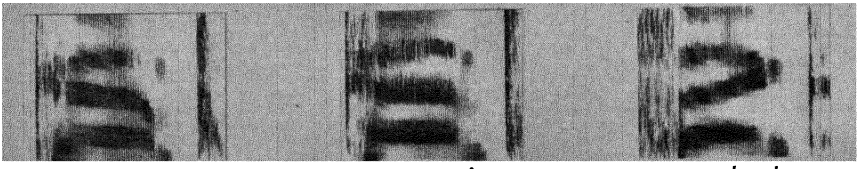
in my ***am not***



sing my

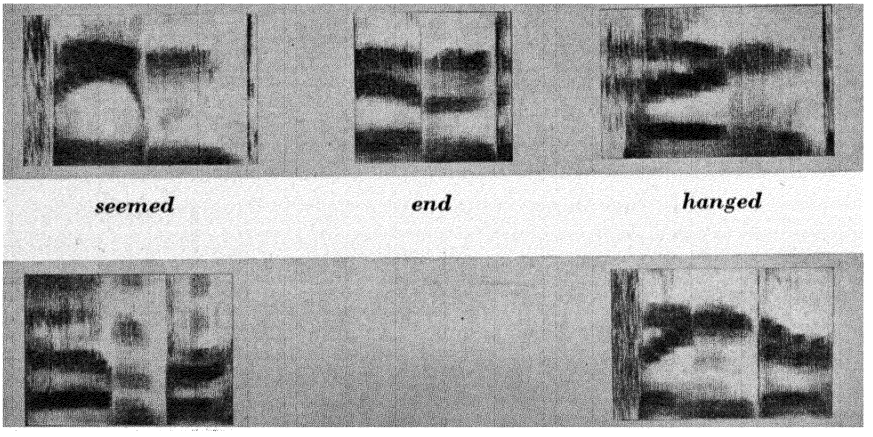
Since the characteristic pattern of each sound is represented in combinations of two nasals, it may be compared with the patterns made by the sounds spoken both in isolation and in combination with other sounds as illustrated in previous units. Like combinations of voiced fricatives, the nasal combinations portray the transitional movements between the combined sounds, thereby facilitating identification of the sounds. The patterns above show examples in which both nasals in a combination were spoken with approximately equal stress. Variations in stress as well as shifts in the nature of cavity resonance or type of modulation produce changes in the relative predominance of the resonance areas for each of the combined sounds, but these variations do not alter the hub or direction of influence significantly.

In terminal combinations of voiceless stops and nasals, the nasal may be identified both by its pattern and influence; but the stop usually is identified primarily by its pattern unless it is followed by another sound in continuous speech. The nasal sound



camp *can't* *thank*

is usually short and often is unstressed. It is, therefore, rather difficult to perceive the stop influence on the nasal. As might be expected, the most common combinations within syllables occur between the nasals and voiceless stops that are produced from similar articulatory positions; **m** with **p**, **n** with **t**, **ŋ** with **k**. Other combinations are frequent between a terminal nasal sound in one syllable and an initial voiceless stop in the succeeding syllable, but since these combined sounds present clues resembling the monosyllabic combinations, they are not illustrated except as they occur in the illustrations of sentences. There is some difference found in the duration of the stop gap which tends to be shorter in combinations of sounds made from similar articulatory positions than it is in sound combinations produced from dissimilar positions. It is of interest to note the customary pronunciation of the **n** preceding **k** or **g** as though it were spelled **ng**.

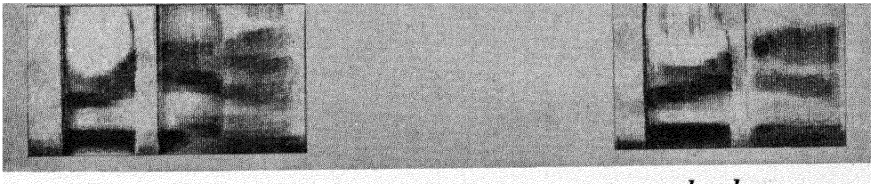


seemed *end* *hanged*

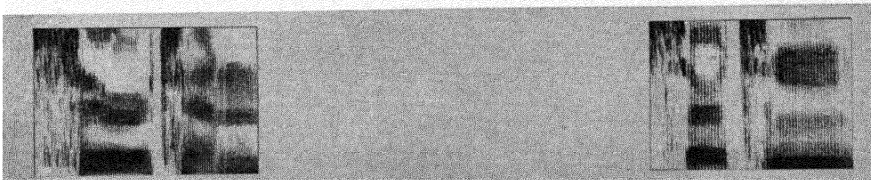
amber *finger*

In monosyllabic combinations of nasal sounds with voiced stops, the **m–b** and **ŋ–g** combinations do not occur frequently, although the **m–d** (*seemed*) and **ŋ–d** (*hanged*)

combinations are fairly common within syllables, often being used to form the past tense of verbs. The illustrations above show the nasals combined with the **d**, and also present the **m–b** (*amber*) and the **ŋ–g** (*finger*) combinations in order that the nasal-voiced stops may be compared with the nasal-voiceless stop combinations. Recognition of the nasal sound may be based upon both its pattern and influence; but the voiced stop, unless its influence is apparent on the succeeding sound, must be identified primarily by its pattern. The duration of the stop gap is frequently of assistance, since it is shorter between sounds produced from similar articulatory positions than between sounds made from dissimilar positions. Comparison of the patterns above with those of the nasal-voiceless stop combinations illustrates that, like the vowels, the nasal sounds spoken with voiced stops tend to be longer in duration than those said before voiceless stops.



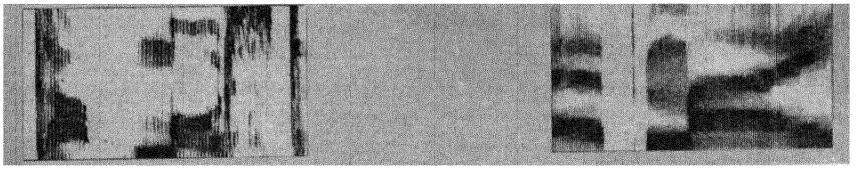
burden *burden,*



certain *certain*

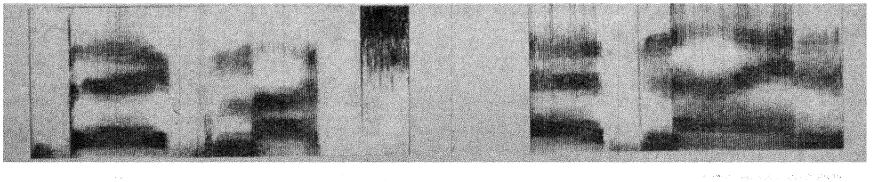
Stops and nasals are combined most frequently at the end of polysyllabic words or between words in continuous speech. When such a combination occurs terminally in an unstressed syllable as in *burden* or *certain*, the vowel between the stop and the nasal sound may or may not be pronounced. If spoken, the vowel usually is some form of the *schwa* or unstressed neutral vowel. If no vowel is used, the syllable consists of the stop and nasal alone, and the nasal is termed a syllabic sound. In the latter case, the stop either is exploded directly into the nasal or is imploded. Syllabic nasal sounds are most apt to occur in combinations of stops and nasals articulated from similar positions such as the **tn** or the **pm**. Since the **tn**, **dn** combinations are encountered more often than others in English, the patterns above illustrate both the syllabic

n and the *schwa*n with the **t** and **d**. The nasal generally is not stressed, and its pattern therefore tends to be short in duration and rather light in intensity. The stop may be identified both by its pattern and influence; the nasal may be identified by its pattern and by its influence on the preceding or succeeding sound as the case may be.



top most ***at my***

Combinations of terminal voiceless stops in one syllable and initial nasal sounds in a succeeding syllable are illustrated in the patterns above. The stop may or may not be exploded, depending upon the rate and precision of articulation and the stress. When a stop is combined with its homophenous nasal, or in other words when both sounds are produced from similar articulatory positions, the stop frequently is imploded or only partially exploded. Sometimes a sharp spike precedes the nasal sound, representing a rapid shift in the position of the soft palate. In the patterns above, the **p** of *top* has no significant influence on the **m** of *most*, since their hubs are similar. However, the **t** of *at* influences the **m** of *my* quite noticeably from its higher hub. The **p** spike is absent, although the stop gap and influence on the vowel present clues for identifying the stop. The **t** spike is short, and there is no fricative pattern because the abrupt explosion was not followed by aspiration or frictional modulation. The gap between the **t** and **m** portrays the time taken for the articulatory shift. In both cases, the stop may be identified primarily by its influence on the preceding and succeeding sounds, and by the stop gap.



bag much ***add nine***

The **g**–**m** combination as portrayed above shows only a partial **g** spike, but the voiced stop gap and the **g** influence help identify the sound. The variable nature of

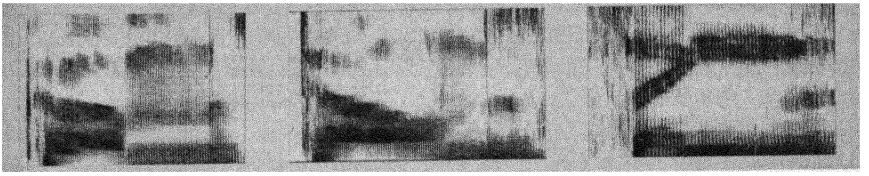
the g is evidenced in the lack of strong influence on the m, though some influence may be observed. The d–n combination illustrates a combination of homophenous sounds in which the d is exploded sharply and an immediate shift is made from the non-nasal to nasal resonation. As in the voiceless stop-nasal sound combinations, changes in rate, precision of articulation, and stress may modify the appearance of the spike, the amount of fill following the spike, and the nature of the nasal pattern.



warmth *once* *length*

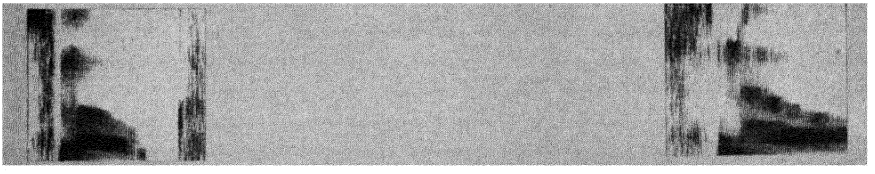
A nasal-voiceless fricative combination may be spoken with or without an intervening voiceless stop sound. In cases where a stop is employed, it is that stop which is similar in articulatory position to the nasal sound, since the explosion is attendant upon the release of tension and pressure following an abrupt movement from the nasal articulatory position. The intervening stop sound also is believed to be incident to the speed of termination of vocal cord modulation and the subsequent initiation of frictional modulation. If the nasal-fricative transition is made gradually, the stop gap and spike are absent because the fricative follows the nasal directly. Frequently the fricative fill begins rather abruptly as in *warmth*, portraying the shift in position of the soft palate upon termination of the nasal sound. Stress and rate and precision of articulation are reflected in the duration of the intervening stop gap, in the sharpness of the stop spike, in the nature of the termination of nasal resonance, and of course in the relative intensity and duration of each sound pattern. Consistencies in the articulation habits of an individual frequently are evident in these nasal-fricative combinations and are of assistance in helping to identify the individual. Differences in the relative proximity of the articulatory positions of the combined sounds are also reflected in the rapidity with which transitions can be made and thus in the appearance of the patterns. Such a combination as the n–t–s or n–s in *once* is relatively simple to articulate due to the similarity in articulatory positions for these sounds.

In the patterns above, stop gaps and spikes are evident in *once* and *length*. Since the stop hub in each case is similar to that of the preceding nasal, the effect is an absence of influence on the nasal and a reciprocal stop-fricative influence as previously discussed in the Unit on stop-fricative combinations. The pattern of *warmth* illustrates the absence of a stop and the resultant direct influence of the θ (*thin*) hub on the m resonances.



comes *towns* *sings*

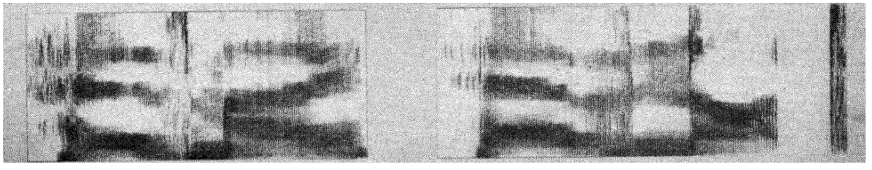
Nasal-voiced fricative combinations follow the same general principles as their nasal-voiceless fricative equivalents with the exception of the continuation of vocal cord-cavity modulation. The patterns above illustrate each of the nasal sounds as combined with the **z** (*zoo*). A **z** terminal is encountered often in plural endings and also in certain of the verb forms. Continuation of vocal cord-cavity modulation is portrayed in the voice bar which extends through both nasal and fricative. Abrupt termination of nasal resonance and shift from the articulatory position of the nasal are portrayed by a single spike in both *comes* and *towns*. In *sings*, the shift from nasal to fricative was not made very rapidly, as evidenced in the gradual transition between the characteristic patterns for each sound. Differences in the hub positions of the **m** and **z** in *comes* as opposed to the similar hubs of the **n** and **z** in *towns*, may be observed in that portion of the nasal and fricative on either side of the spike. In *sings*, the **ŋ** influence is perceptible on the second resonance of the **z**.



smoke *snow*

Combinations of voiceless fricatives and nasals except for the **sm** and **sn** illustrated above, do not occur frequently in monosyllables. Usually the nasal is quite short in duration, and is initiated and terminated rather abruptly. However, even in the short **m** in *smoke*, the influence of the **s** hub may be observed in the transitional movement of the **m** resonances. There is little influence of the **s** on the **n** due to the correspondence in the articulatory positions of the two sounds. The reciprocal influence of the nasal on the fricative may be noted in these illustrations. The nasal may be identified by its pattern and by its influence on the succeeding sound. The fricative may be recognized by its pattern, by its influence on the nasal, and by the nasal influence on the fricative fill. An **s** spoken before an **m** or **p** usually presents a fill covering the entire

frequency range of the present spectrogram, whereas the **s** spoken before an **n** or **t** appears to have an area of concentration towards the top of the present spectrogram. These modifications in the **s** fill pattern are believed to be related to the tension and positions of the articulators, and especially to the relative size and position of the tongue groove in producing the **s**. It seems apparent that in moving from an alveolar-tongue point **s** to a **t** or **n** the tongue groove and the articulatory movement produce an **s** differing in aural and visible pattern from the transition between an alveolar-tongue point **s** and an **m** or **p**. The alveolar-tongue blade **s** and the **m–p** or **n–t** combinations may be differentiated similarly. Detailed discussion of such points is not, however, the province of this volume and must be reserved for future research and writings.



has my ***have not***

Voiced fricatives and nasals are not combined within syllables except in the rare occurrences of syllabic nasals, although they are frequently combined in the course of continuous speech. Vocal cord-cavity modulation is carried through both sounds, but the nature of the cavity modulation is altered during the transition from fricative to nasal. In the illustrations above, the termination of frictional modulation is quite evident, as is the abrupt initiation of nasal resonance. Very rapid speech sometimes makes identification of the sound by its pattern difficult since the resonances may not reach their customary positions. However, the clues afforded by the influences of the sounds are usually sufficient to assure recognition. The movement of the resonance bars from the higher **z** hub to the lower **m** hub in *has my*, and from the lower **v** hub to the higher **n** hub in *have not*, in addition to the influences exerted on the associated vowels, and the individual sound patterns, illustrate these clues to recognition.

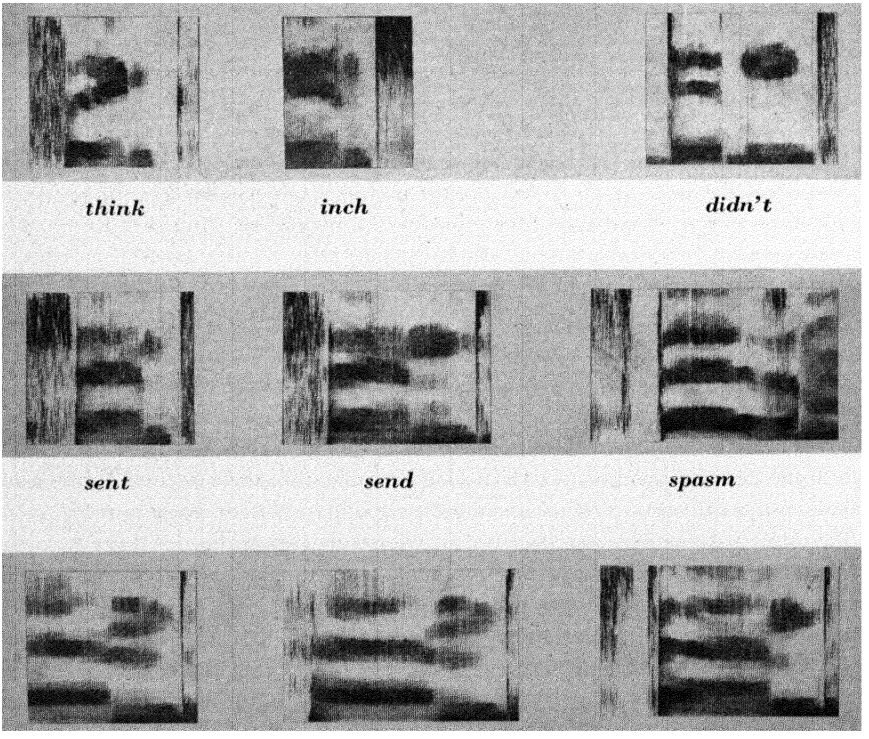
The combinations of nasals with stops, fricatives, and nasals that have been described and illustrated in this unit are among the common blends encountered in reading visible speech and in connection with problems of speech development, correction, or improvement. There are numerous other combinations occurring between syllables or words in continuous speech, but it is hoped that the general principles discussed may be applied to the identification and analysis of any of the fricative-stop-nasal combinations. Although the articulatory movements and positions and the resultant visible and aural patterns are modified in varying degrees by differences

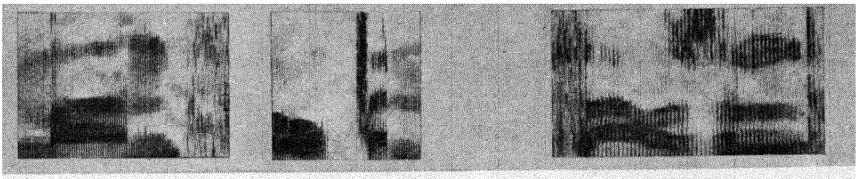
in rate of speaking, precision of articulation, rhythm, and stress, the fundamental principles of sound combination provide consistent clues to the recognition of the phonetic elements combined.

Review words

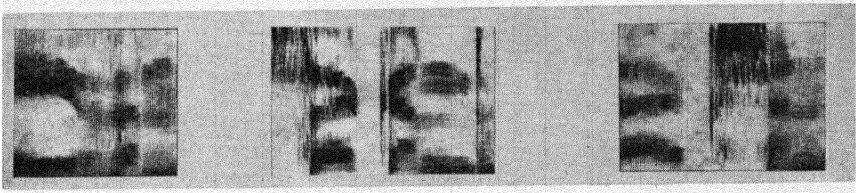
seemed, page 192	*end*, page 192	*sings*, page 196	*amber*, page 192
camp, page 192	*can't*, page 192	*length*, page 195	*finger*, page 192
warmth, page 195	*once*, page 195	*thank*, page 192	*burden*, page 193
comes, page 196	*towns*, page 196	*hanged*, page 192	*certain*, page 193
smoke, page 196	*snow*, page 196	*given*, page 180	*person*, page 180
mountain, page 180			

New words to learn

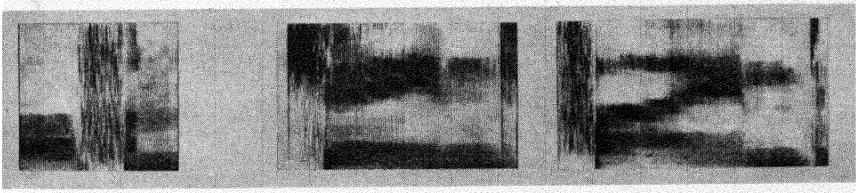




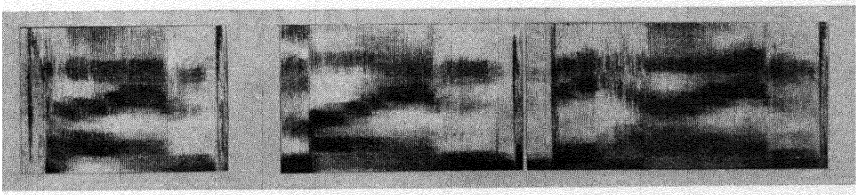
month *open* *thousand*



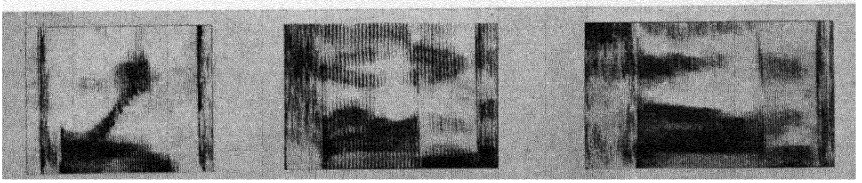
even *second* *action*



often *change* *find*

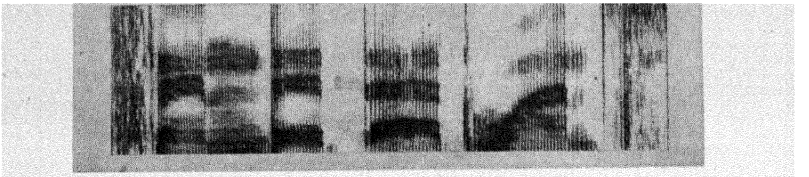


kind *mind* *behind*

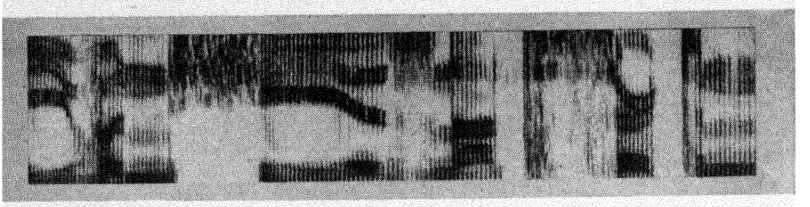


point *found* *sound*

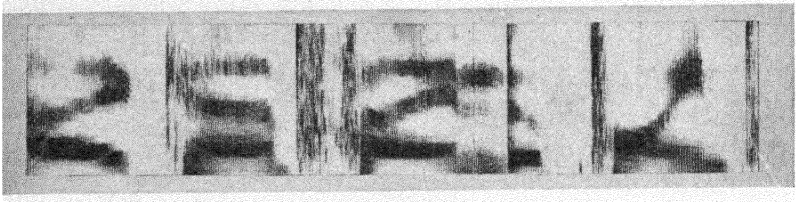
Drill phrases and sentences



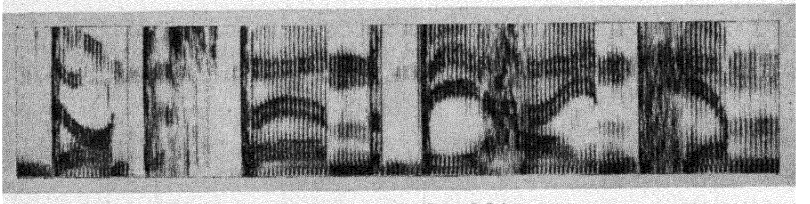
Send it at once



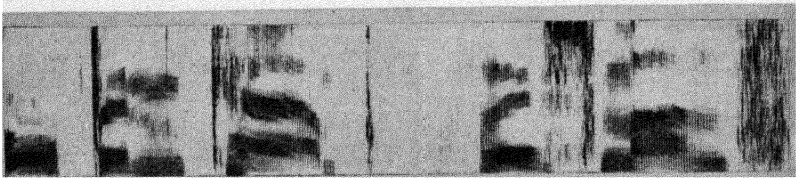
Even she is not certain



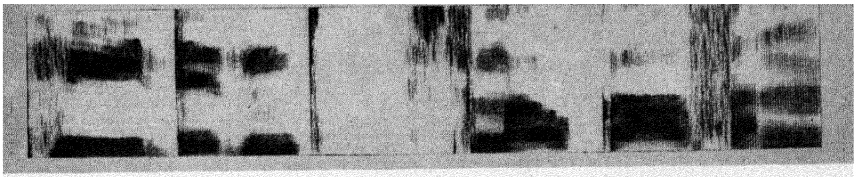
I can't find a point



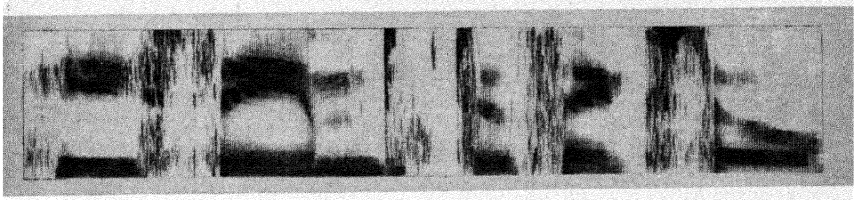
Don't stand behind him



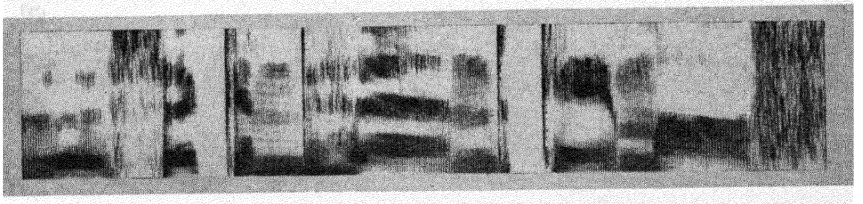
Open camp this month



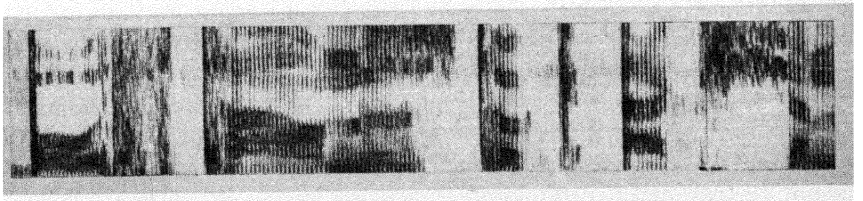
He didn't smoke often



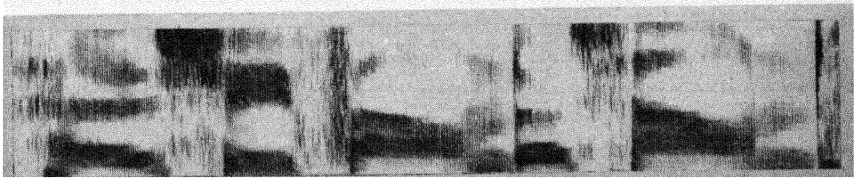
He seemed to think so



The second hand came off



Both towns took action



Has she found a sound?

Chapter 11

Reading the Vowel-Like Sounds—The Glides

Characteristic patterns of the four vowel-like sounds frequently designated as glides, **w** (*we*), **j** (*you*), **r** (*read*), and **l** (*let*), are discussed in this chapter. The combination sound **hw** (*when*) is presented in the combination unit at the end of the chapter. Patterns of the vowel-like glides are recognized principally by the following clues:

1. Vowel-like appearance of the resonance bars.
2. Positions of the beginnings of the resonance bars when the glide precedes another sound.
3. Positions of the endings of the resonance bars when the glide follows another sound.
4. Transitional curve of the resonance bars.

Glides, like vowels, are produced by vocal cord and cavity modulation, but unlike the vowels, the cavity modulation used in producing a particular glide is variable. Variability in the size and shape of the cavity modulators results from the movements of the articulators during the production of the glide. Since the glides are produced only while the articulators are in motion, their patterns are characterized by a sloping or curving of the resonance bars. The slope of the bars, like a curve in a road, may be sharp or gradual, and may vary in length; and the rate of articulation determines the sharpness of the curves. If the rate of utterance is fast, the curves of the resonance bars of the glides will be sharp and steep; if the rate of utterance is slow, the resonance bars of the glides will be gradual and longer. Another important determinant of the length of the curved resonance bars of the glides is the distance through which the articulators are required to move, that is, the distance between the position of origin of the glide or its hub, and the hub of the combined sound. If the positions of the origin of the resonance bars of a glide are near to the positions of the resonance bars of the combined sound, as in *ye* (**ji**) and woo (**wu**), the transitional length of the glide between the sounds is minimal. Maximal transitional length of a glide results when the positions of the resonance bars of the origin of the glide and the combined sound are separated as far as possible, as in *we* (**wi**) and *you* (**ju**). Many variations in the transitional length of the glides occur within the range established by the minimum and maximum combinations.

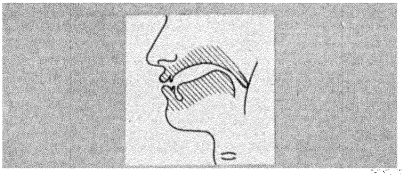
The sound made at the origin of a glide may be produced from a steady state position of the articulators but, as previously mentioned, the glide itself is not produced except during the interval in which the articulators move from the steady state position to the position for the combined sound. Although the duration of the pattern of the sound made at the origin of a glide will vary with the rate of utterance, the positions

of its resonance bars remain quite constant for each of the glides except the l (*let*). The l (*let*) may be initiated from one of several articulatory positions, depending upon the nature of the combined sound; therefore, the location of its resonance bars at the beginning of the glide varies accordingly. Each of the glides has particular and characteristic points of initiation or termination, and a distinctive curvature of its resonance bars. Because of their unique characteristics, the patterns of the glides are among the easier patterns to differentiate.

These general observations will be illustrated and explained in the Lesson Units that follow. Each glide will be discussed in a separate unit and under the same headings used in the Lesson Units in the preceding chapters. Illustrations with several selected vowels were chosen to provide the reader with the essential principles basic to the visualization of other combinations of glides and vowels. Also, the similarities and differences of the several glides are illustrated and discussed in a Comparative Summary Unit.

Unit 1—Reading the w (we)

How the w (we) is made



***Profile of* w (*we*)**

The w (*we*) is produced by:

1. Rounding the lips, without too much tension, to a position similar to that used in saying u (*boot*).
2. Raising the back of the tongue towards the palate to a position similar to that used in saying u (*boot*).
3. Closing the opening between the throat (pharyngeal) and the nasal cavities.
4. Emitting a continuous stream of voiced breath while the articulators are moving towards their positions for the combined adjacent sound.

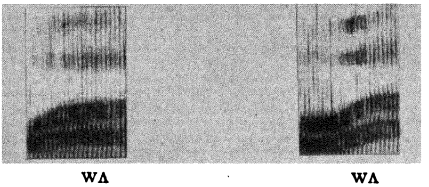
The w (*we*) is produced while the articulators move from positions similar to those used in saying u (*boot*) to the positions used for the sound which follows. The precise articulatory positions from which the w (*we*) glide starts depends somewhat on the sound with which it is combined. For instance, the rounding of the lips and the opening between the lips may vary with each sound with which the w is spoken, although in each combination, the sound made at the origin of the w glide resembles the u except that bar 2 is a little lower. Like the u, this sound can be continued from a

steady state position of the articulators. The w glide itself cannot be continued. It is produced during the movements of the articulators which vary the cavity modulation, it is always terminated in another sound, and it cannot be produced in isolation. Because of these characteristics the w is described phonetically as a glide. The above profile illustrates the general position of the articulators from which the w glide is initiated, but no attempt is made to portray the movements used in producing the glide.

The w may be spelled in the following ways:

w (*weed*) *oi* (*choir*) *u* (*quick*) *o* (*one*)

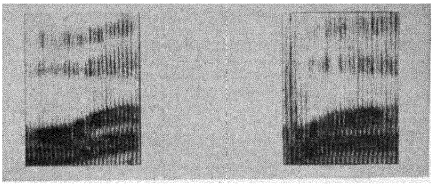
Visible patterns of the w



WA WA
unstressed ***stressed***

Since a glide cannot be made as an isolated sound, the w is shown above in combination with the central vowel ʌ (*up*). The w pattern may be described in the two phases that correspond to the production of the sound. The first phase is the sound produced by vocal cord and cavity modulation, similar to the type of modulation used in producing u (*boot*). This part of the pattern is characterized by the location of bar 1 on the baseline and bar 2 in a slightly lower position than bar 2 of u (*boot*). The second phase is the glide produced by vocal cord and varying cavity modulation. This part of the pattern is characterized by the curving of the resonance bars from their positions of origin to the positions of the bars for the combined sound, which is the ʌ (*up*) in the illustrations above.

When the w glide is unstressed, the resonance bars may be light or partly missing, their length is shorter than when stressed, and the two phases of the sound are less clearly defined. Stressing increases the darkness and length of the glide pattern and emphasizes both phases of the sound. The position of origin of the w glide is quite clear in the stressed pattern above. Sometimes resonance areas in the upper portion of the pattern are visible for the first or second phase of the w. These resonances seem to be associated with variations in the proportions of cavity-vocal cord modulation used in producing the sound.

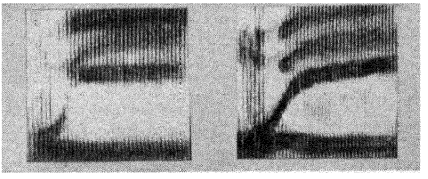


WA
vocal cord-varying cavity modulation

WA
vocal cord-varying cavity and frictional modulation

Frequently, frictional modulation, in addition to vocal cord-cavity modulation, is used in producing the w glide. Frictional modulation is portrayed by vertical striations either superimposed on the resonances or substituted for them, depending upon the relative proportions of the different types of modulation. When frictional modulation is present, it generally is evident in the first phase of the glide and it is believed to be related to the tension and position of the lips in producing this part of the sound. As shown above, the w may be produced without frictional modulation, in which case there are no such vertical striations in the pattern.

How the w influences vowel patterns

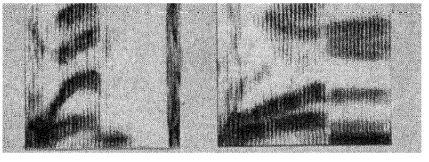


we *way, weigh*

The influence of the w glide on i (*we*) and eɪ (*way*) is illustrated above. In the pattern of *we*, bar 1 of the vowel remains on the baseline uninfluenced by the w since bar 1 of the i is also on the baseline. Bar 2 of the w curves upwards and joins bar 2 of the vowel. Upper resonances of the w are partially visible in the pattern of *we*, and may be observed to curve downwards slightly during the glide phase and then to curve upwards during the portion of the transition immediately preceding the vowel. In the pattern of *way*, bar 1 of the vowel curves upwards from the beginning of the glide to its position for the first vowel in the diphthong eɪ. Bar 2 curves upwards and joins bar 2 of the e and a similar curve is repeated to lesser degree in bars 3 and 4 of the vowel. Resonances 3 and 4 are visible in the origin phase of the w, drop out during the glide, and then the transitionally curved resonance bars of the vowel appear. Patterns such as this illustrate in a general way the separate phases involved in a glide-

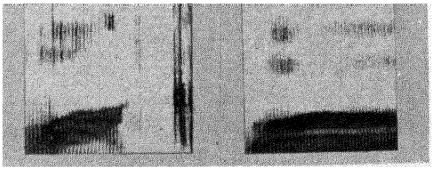
vowel combination. However, it is not possible at this time to draw specific conclusions concerning the relative duration of these phases. Frictional modulation may be observed in the first phase of the w in the patterns of *we* and *way*.

Since a glide is made during the movement of the articulators from the position of origin to the position from which the combined vowel is produced, it may be argued that a glide has no influence on combined sounds. Academically, this may be true; but for descriptive purposes we have designated that portion of the transitional pattern immediately before the combined sound as the zone of influence on that sound rather than the ending of the glide. As a group, glides are considered as but a few of the many transitional sounds that have been symbolized as sound entities in our language.



went ***won, one***

Patterns of *went* and *won* are used to illustrate the influence of the w on the ɛ (*met*) and ʌ (*up*) vowels. Bars 1 and 2 of the w begin at the bottom of the patterns and curve upwards in both patterns. The slope of the curve of bar 2 in *went* is steeper than it is in *won* because the bar 2 position of ɛ (*met*) is much higher than the bar 2 position of ʌ (*up*). The beginnings of bars 3 and 4 curve upwards for ɛ, but for ʌ their patterns are indistinct. As in the i (*we*) and eɪ (*way*) patterns, the w influence on ɛ (*met*) and ʌ (*up*) is to curve the beginning of the vowel bars upwards. Again, vertical striations due to frictional modulation are present in the first phase of the w patterns.



walk ***woo***

In the pattern of *walk*, the w is combined with the vowel ɔ (*all*). Bars 1 and 2 of the w originate near the bottom of the pattern and curve upwards together to the pattern of ɔ. The transitional curves are shorter than those in the preceding patterns because of the proximity of the w and ɔ hubs, and also of the first bars of both sounds. In

u (*woo*), the positions of resonance bars 1 and 2 are as low as they are for any vowel, so that the **w** influence is at its minimum in this combination. Bar 1 is on the baseline for the beginning of the **w** and also for **u**. Therefore, bar 1 of **u** (*boot*) is not curved by the **w**, but a shift in the width of this bar is evident in the pattern. Bar 2 of the beginning of **w** is just a little lower than bar 2 of **u**, so there is a slight, short curve in the beginning of bar 2 of **u** due to the influence of the **w**. Bars 3 and 4 of both the **ɔ** and **u** are uninfluenced by the **w**.

When the **w** glide occurs before a vowel, it curves upwards from the **w** hub to join, at some point, the transitional curves of the vowel resonances. Since the hub of the **w** is towards the bottom of the pattern, the **w** resonance bars 1 and 2 curve from this position of origin to their point of juncture with the bars of the combined sound. If the positions of the resonance bars of the combined sound are known, the influence and the transitional glide of the **w** can be visualized. As the vowels are said in series after **w**, the curve of bar 1 becomes progressively more marked from **i** (*we*) to **ɑ** (*father*) and then becomes less prominent or decreases from **ɑ** to **u** (*boot*). These **w**-vowel bar 1 transitions are similar to the transitions occurring between other voiced sounds and vowel bar 1; and, like the latter, they appear to be related to the change in modulation which is associated with the transitional shift. The curve of bar 2 in **w**-vowel transitions decreases as the vowels are said in series from **i** (*we*) to **u** (*boot*). An upward curve at the beginning of bar 3 is more evident in the patterns of **i** and other front vowels than in the patterns of **w**-back vowel combinations.

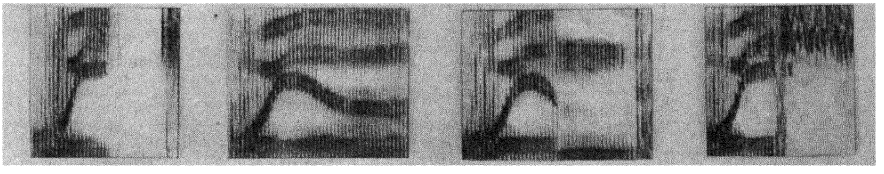
Read the **w** *by:*

1. The appearance and position of the vowel-like resonance bars at the beginning of the glide.
2. The direction of the curves of the resonance bars.
3. The influence of the **w** on the pattern of the combined sound.

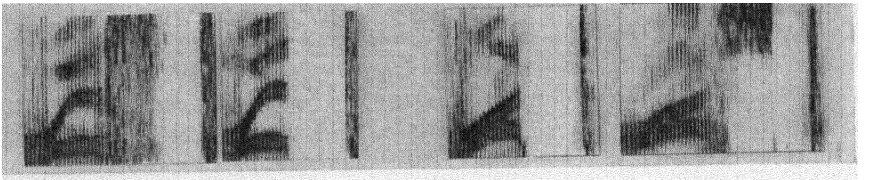
Review words

way, page 205	*went*, page 206	*walk*, page 206
we, page 205	*won*, *one*, page 206	*woo*, page 206

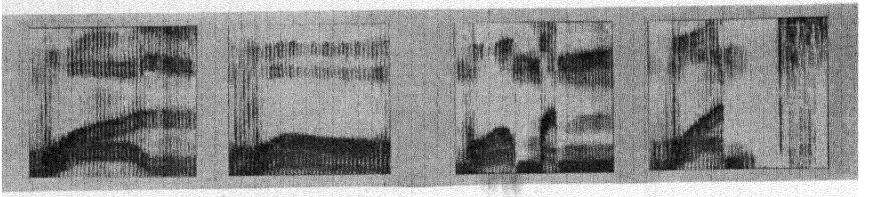
New words to learn



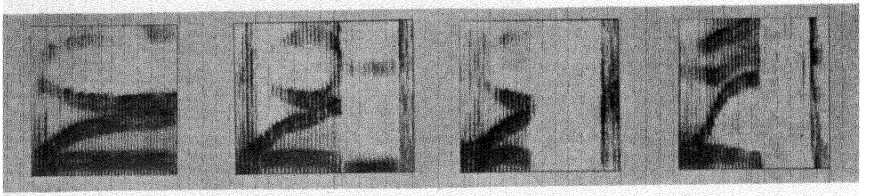
week, weak ***with*** ***wind*** ***wish***



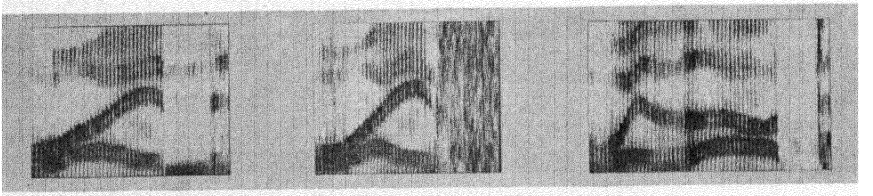
west *wet* *watt* *wasp*



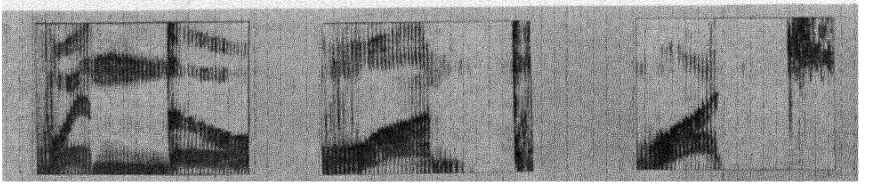
was *woe* *woman* *once*



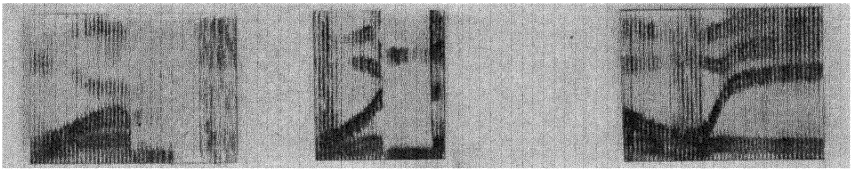
were *word* *work* *wait, weight*



wide *wife* *without*



window *want* *watch*

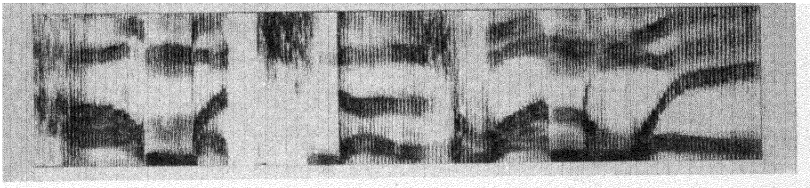


warmth *wood, would* *away*

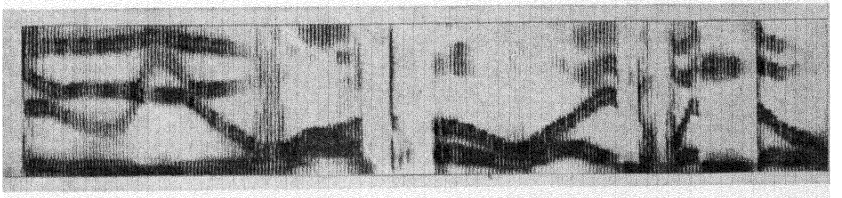
Additional words for practice in visualization

win	*wing*	*within*	*wash*
worth	*waste*	*wave*	*wise*

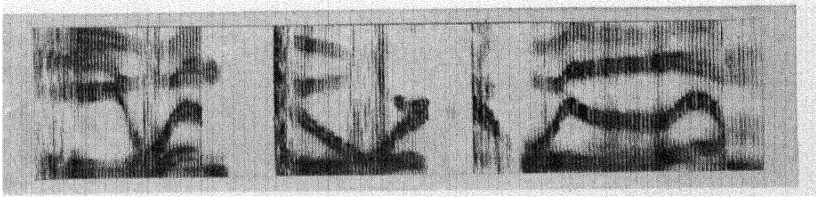
Drill phrases and sentences



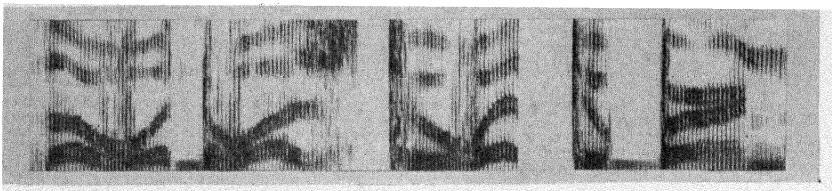
How much does one weigh?



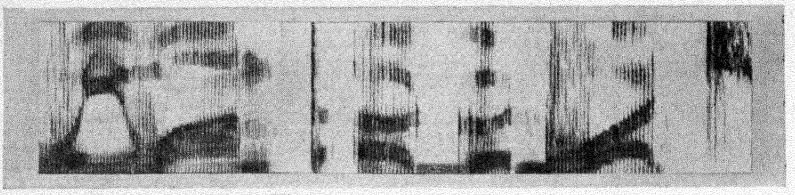
Do you want a wide window?



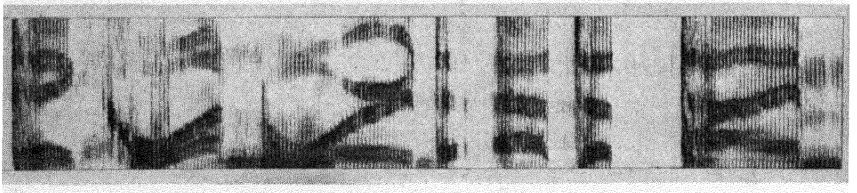
He went to work with them



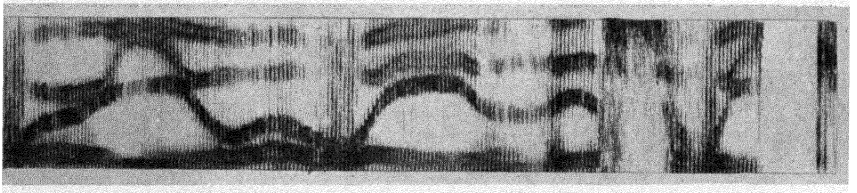
The wood was too wet to burn



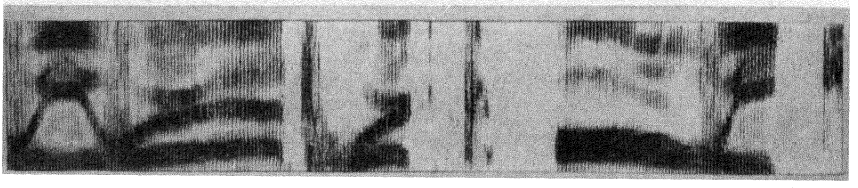
We want a good watch



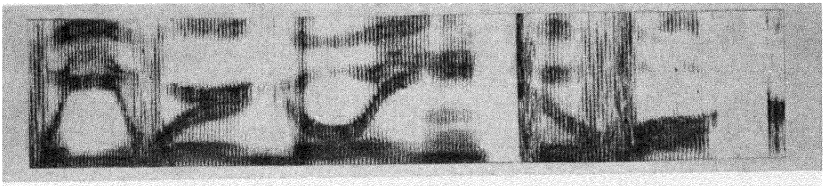
Take one word at a time



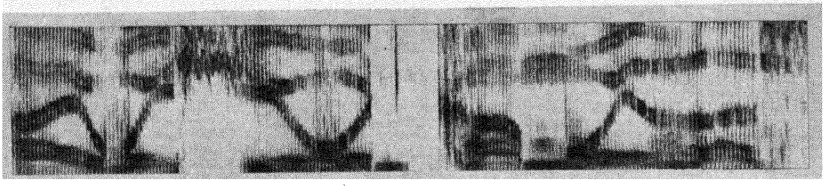
Were you away this week?



We were at work all week



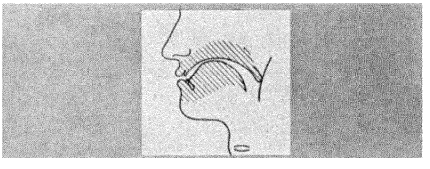
We were going to walk



I wish you would come with us

Unit 2—Reading the j (you)

How the j (you) is made



Profile of j (*you*)

The j (*you*) is produced by:

1. Placing the lips in approximately the same position as for ɪ (*it*).
2. Starting with the other articulators held almost in the i (*eve*) position, except that the elevated portion of the tongue is higher and not so far forward in the mouth.
3. Closing the opening between the throat (pharyngeal) and the nasal cavities.
4. Emitting a continuous stream of voiced breath while the articulators are moving towards their positions for the combined sound.

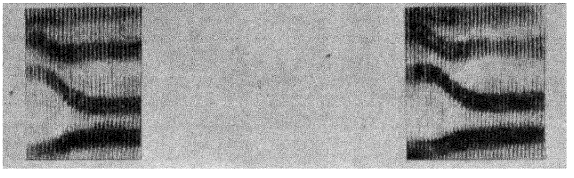
The above profile indicates the positions of the articulators at the beginning of the j. However, the major portion of the j, like the w glide, is produced while the articulators move from their hub or starting position to the position for the sound which follows. The exact position of the articulators from which the j glide starts depends somewhat on the sound with which it is combined. Although the sound produced at the beginning of the j can be continued from a steady state position of the articulators,

no single portion of the glide phase of the sound can be continued, since the glide is made while the articulators are in motion and cannot be produced in isolation. It is a transitional sound characterized by continuously varying vocal cord-cavity modulation between the position of origin and its termination in the next sound. Because the j is produced in this manner, it has been described phonetically as a glide.

The j may be spelled in the following ways:

y (*you*)	*j* (*hallelujah*)	*gn* (*vignette*)
i (*union*)	*e* (*ewe*)	*u* (*unite*)

Visible patterns of the j

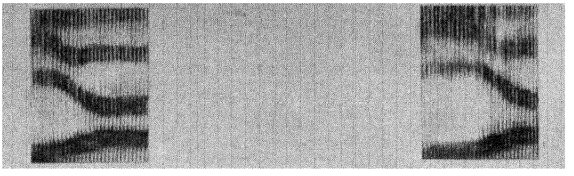


jʌ
unstressed

jʌ
stressed

In the above pattern, the j glide is combined with the vowel ʌ (*up*). Like the w, the j pattern may be described in two phases. Its first phase resembles the pattern of i (*eve*), except that bars 2 and 3 are a little higher for the beginning of the j glide than for the i. Bar 1 of the beginning of the glide is on the baseline as it is for i. This phase of the glide is produced by vocal cord and cavity modulation which may be continued in a steady state condition. The second phase of the glide is the part produced by vocal cord and varying cavity modulation which terminates in a following sound. This part of the glide pattern is characterized by the curving of the resonance bars from their hub positions of origin to the positions of the bars for the combined sound which, in this case, is the ʌ (*up*). Since bar 1 at the beginning of j is on the baseline, it curves upwards to join bar 1 of ʌ. Bars 2 and 3 at the beginning of j are towards the top of the pattern, and they curve downwards to join with bars 2 and 3 of ʌ. The degree of slope of the curves of the resonance bars is determined mainly by the speed of movement of the articulators. If the articulatory movement is fast, the slope is steep; if it is slow, the slope is gradual. Also the duration of the sound produced at the point of origin of the glide varies with the rate of articulation. In patterns having three bars rather than four, the second and third bars are joined in a way similar to the juncture of bars 3 and 4 in these illustrations, and the hub and influence-glide of j are also similar.

Stressing and unstressing of the j modify the patterns in the same general way that they do the **w** patterns. Compared with the stressed pattern, as illustrated above, the unstressed pattern is shorter, lighter, and less clearly defined.

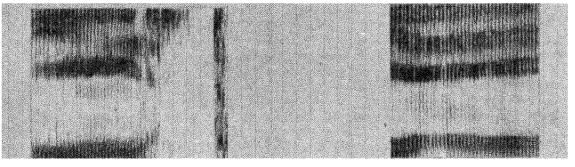


jʌ
vocal cord-varying cavity modulation

jʌ
vocal cord-varying cavity and frictional modulation

Frictional modulation is sometimes present in the normal production of j, but it does not occur as frequently as in the **w**. However, it appears often enough to warrant the inclusion of the above pattern illustrating the modifications in pattern resulting from the addition of frictional modulation portrayed by vertical striations superimposed on the resonance bars at the beginning of the j pattern.

How the j influences vowel patterns

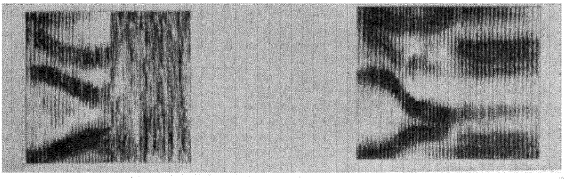


yeast ***yea***

In *yeast*, bar 1 of the j glide joins with bar 1 of the vowel i (*eve*) without changing position, because both bars are on the baseline. However the bar may be observed to widen slightly during the i. Bar 2 of j continues into bar 2 of the vowel with little influence on its position. Bars 3 and 4 of j glide downwards slightly from their j positions into bars 3 and 4 of the vowel. When there are only 3 bars in the pattern, the second and third bars glide downwards to continue into bars 2 and 3 of the vowel.

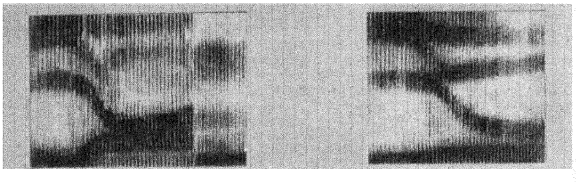
In *yea*, bar 1 of j again joins bar 1 of the vowel, but it has to curve upwards because the bar 1 position of the vowel **e** (*hate*) is off the baseline. The second bar of **j** curves slightly downwards into bar 2 of the vowel. Bar 3 curves more to join bar 3 in *yea* than it does in *yeast*, since the positions of bars 3 and 4 are lower in **e** (*hate*) than they are in **i** (*eve*). Variability in the point of origin or hub of the j is illustrated in these

patterns. The j hub appears to shift slightly from its hub with the i to a lower position with the e. It does not seem to move much below this latter position in other combinations.



yes *young*

Both *yes* and *young* show an upward curve of the first bar of j into bar 1 of the vowel. In each pattern, the second and third bars of j curve downwards to continue into bars 2 and 3 of the vowel. Bar 4 of j also curves downwards into bar 4 of these vowels when it is present. The positions of the bars for ɛ (*yes*) and ʌ (*young*) determine the direction and extent of the transitional curves from their positions at the beginning of the j.



yawn *you*

In *yawn*, the first bar of j curves upwards into bar 1 of the vowel. The second j bar curves downwards to continue into bar 2, which is close to bar 1. The beginning of bar 3 of j curves downwards during the glide phase through a position near bar 2, but curves upwards as it joins the transitional portion immediately preceding bar 3 of the vowel. This shift in bar 3 occurs frequently in combinations of j with the vowels in series from the ɑ (*father*) through the ɔ (*all*). If the fourth bars are present, they join without significant shift of positions.

Bar 1, in *you*, is on the baseline for the beginning of j and also for u. Therefore, bar 1 of u is not curved by the j, but there is usually a noticeable shift in the width of the bar. The j hub, or bar 2 at the beginning of j, and the vowel u are farther apart than they are for any j-vowel combination, so in this pattern the influence of j on a vowel is at its maximum. This is portrayed in the long curve of bar 2 from the beginning of j to u. Bar 3 of j curves with bar 2 until it reaches the level of bar 3 of

the u, at which point it straightens out to join with the vowel bar. Bar 4 of j curves downwards slightly before straightening to continue into bar 4 of u.

It may be noted that in the three-bar patterns the second and third j bars usually are joined at their position of origin. In the four-bar patterns, the third and fourth j bars are joined at their beginning when combined with the front vowels, but the second and third j bars usually are joined at their beginning when j precedes the back vowels. The initial position of bar 2 of j shifts slightly with different vowels. It is slightly higher with the front vowels than with the back vowels. However, this shift is not great enough to cause confusion, and the j glide influence, when it occurs before any vowel, can be visualized provided the positions of the resonance bars of the vowel are known. Bar 1 of j always starts on the baseline, so unless bar 1 of the combined vowel is on the baseline too, bar 1 of j curves upwards to join bar 1 of the vowel. The extent of the curve will depend on how far bar 1 of the vowel is off the baseline. Bar 2 of j begins in approximately the same position as bar 2 of the vowel i (*eve*). With all the other vowels as said in series after i, bar 2 of j curves downwards to join bar 2 of the vowel. The extent of the curve increases with each vowel in the vowel series and reaches its maximum in ju (*you*). Bar 3 of j begins a little higher than bar 3 of i (*eve*), and curves downwards with bar 2 until it reaches the level of bar 3 of the combined vowel with which it joins. Bar 4 of j, when it is present, begins at the top of the pattern and curves downwards or continues straight, depending on the position of bar 4 of the vowel with which j is combined.

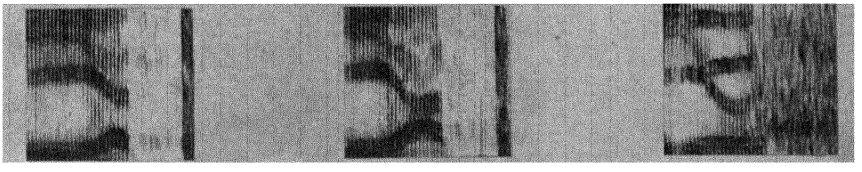
Read the j *by:*

1. The appearance and position of the vowel-like resonance bars at the beginning of the glide.
2. The direction of the curves of the resonance bars.
3. The influence of the j on the patterns of the combined sound.

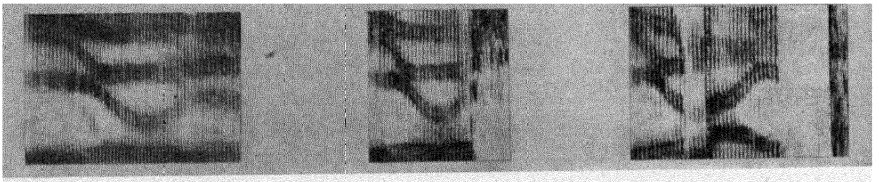
Review words

yeast, page 213 *yes*, page 214 *yawn*, page 214
yea, page 213 *young*, page 214 *you*, page 214

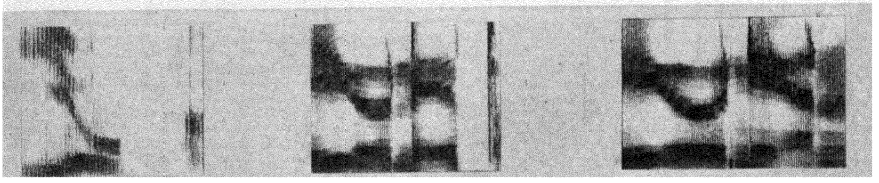
New words to learn



yet ***yacht*** ***youth***



use (juz) *use* (jus) *unite*

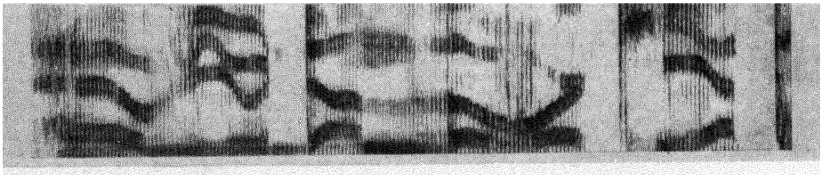


yoke, yolk *unit* *union*

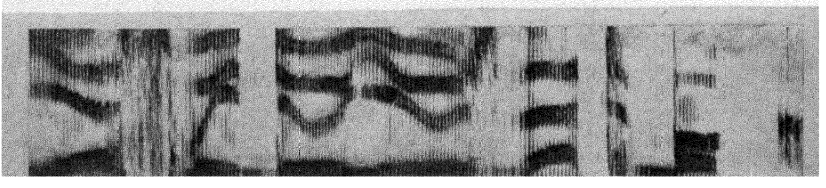
Additional words for practice in visualization

ewe *useable* *utility* *yard*

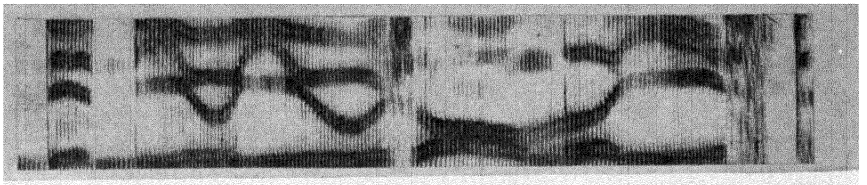
Drill phrases and sentences



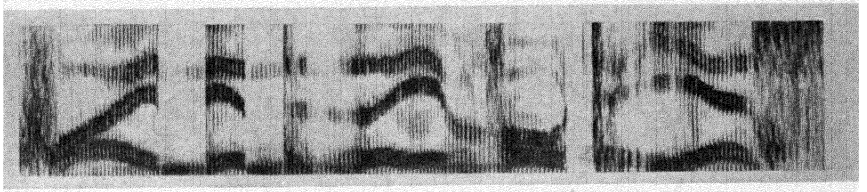
Have you done the work yet?



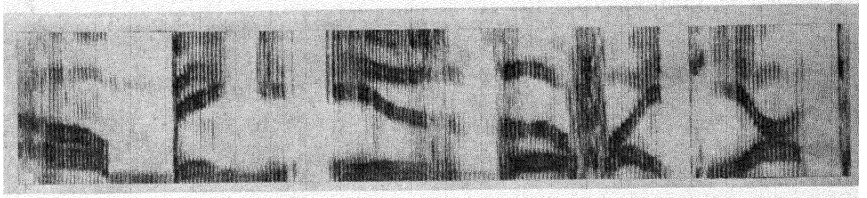
Yes, we do use that book



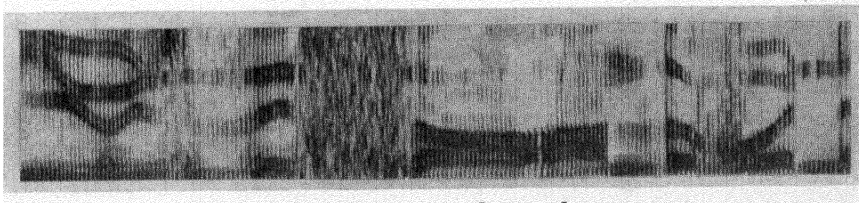
Did you use all the yeast?



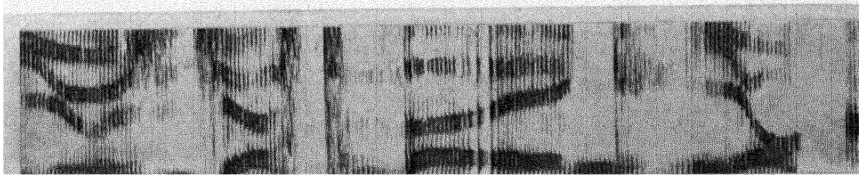
Why did they vote yes?



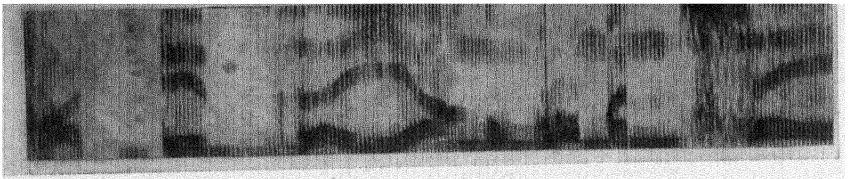
How big is the white yacht?



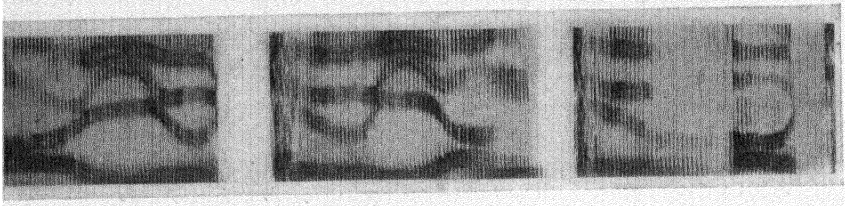
Use this saw on the wood



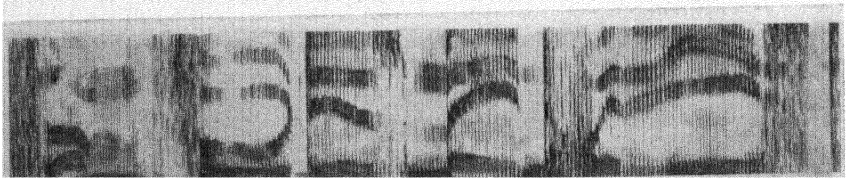
Use just the egg yolk



What did the young woman say?



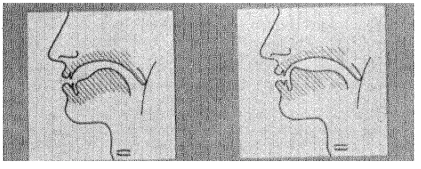
Are you too young to vote?



Some food is made with yeast

Unit 3—Reading the r (read)

How the r (read) is made



Profile of r (*read*) *with tongue tip down*

Profile of r (*read*) *with tongue tip up*

The r (*read*) is produced by:

1. Starting with the mouth open as in making t, with slight rounding of the lips.
2. Retracting the tip or blade of the tongue from an elevated position near the alveolar ridge towards the center of the hard palate, or by moving the tip or blade of the tongue forward from an elevated position near the center of the palate towards the alveolar ridge.
3. Closing the opening between the throat (pharyngeal) and nasal cavities.
4. Emitting a continuous stream of voiced breath while the articulators are moving towards their positions for the combined adjacent sound; or,
5. Emitting a continuous stream of voiced breath while the articulators are moving from the positions of a combined sound towards the positions for the r.

Like the w and j, the r glide is produced either by vocal cord and varying cavity modulation or by vocal cord, varying cavity, and frictional modulation. The articulatory positions from which the glide r starts or terminates vary somewhat with individual speakers. Two commonly used positions are shown in the profiles above; but a speaker may use one or both of these positions, or variations of them, in producing the several r glides. In general, the articulators move from or to a position much like that of the fixed vowelized sound, ɝ as in *word*. Because the major portion of the r is made while the articulators are moving from one position to another, it is described phonetically as a glide. For the same reason, the sounds that have been described as front fricative and back fricative r's can be classified also as glides. As will be illustrated in this chapter, frictional modulation usually is introduced by the transitional articulatory movements essential to producing certain combinations with the r.

The terminal r, as in *better*, generally is designated by phoneticians as the unstressed form of the stressed vowelized sound ɝ as in *word*. Since it is an unstressed sound, its duration is short. The sound made during the glide approach to the short terminal part of the r constitutes a considerable portion of the entire r sound. If no sound were produced while the articulators were moving towards the fixed ɝ position, the terminal r could not be considered as a glide. But a transitional glide sound is produced, and for this reason the final r, as in *better*, may be considered as one of the several r glides.

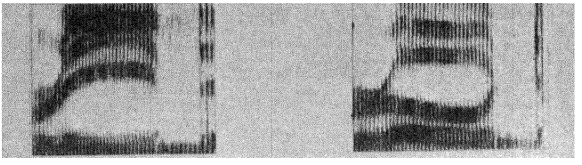
Another variation of the r is characteristic of the speech of many natives of Eastern and Southern United States. There are speakers who use the *schwa* ə (*about*) in place of the final r, and others who use a variation that is neither a *schwa* nor the General American r. Prevalence of so many variations in the production of r makes it impossible to include profiles of the articulatory positions used in producing this sound. In general, the so-called front fricative r (*try*), the back fricative r (*green*), and the unstressed vowelized r (*better*), are glide r's produced as the articulators move to or from the positions described above. The principal difference in their production is the proportion of frictional and cavity modulation in the transitional phase of the

sound. The proportion of frictional and cavity modulation used in the production of these variations of the glide r is determined normally by the nature of the sound combined with the r. Therefore, the illustrations and descriptions of the r's commonly designated as the front and back fricative r and the unstressed vowelized r, are presented in the combination unit at the end of this chapter.

The glide r may be spelled in the following ways:

r (*room*)	*re* (*fare*)	*rh* (*rhetoric*)	*rt* (*mortgage*)
rr (*merry*)	*wr* (*wrap*)	*rps* (*corps*)	

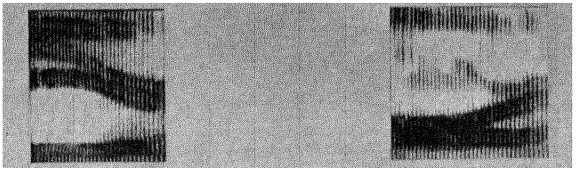
Visible patterns of the r



read ***road, rode***

The r from which and to which the glide r moves can be made from a steady state position, but in continuous speech little time is given to this part of the sound compared to the amount of time given to r made while the articulators are in motion. When the glide r precedes a vowel, as in the word *read*, its beginning is recognized easily by the characteristic position of its resonance bars. Bar 1 is on or very near the baseline. Bars 2 and 3 are in close proximity, a little below the center of the pattern. The group of r sounds are the only sounds of the language that have bars 2 and 3 in this position. Since the glide r is made as the articulators move from this r beginning to the position of the vowel resonances, bars 2 and 3 may curve in parallel, as in *read*, or they may separate as in *road*. Whether or not they stay together, depends upon the relative positions of the resonance bars of the vowel to which or from which the r glides. Like the other glides, the slope of the curves of the resonance bars depends upon how fast the articulators move, the extent of the curves upon the distances through which the articulators travel.

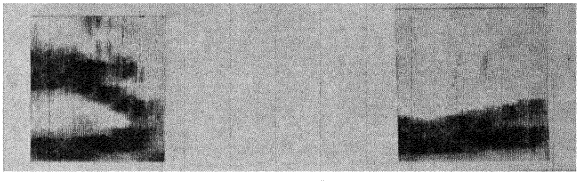
In General American pronunciation, when the glide r follows a vowel as in the words *ear* and *or*, bars 2 and 3 terminate in a position similar to that of the r made before the vowel. However, bar 1 of the terminal r is usually separated from the baseline, the extent of separation depending upon the position of bar 1 of the vowel which precedes the r, and bars 2 and 3 are slightly higher in the terminal r than in the initial r.



ear
General American Pronunciation

or
General American Pronunciation

The positions of bars 2 and 3 of the vowel before the r also determine the direction which these resonances take in combining with the r bars. Since the r glide may be stopped before it has reached its customary conclusion in the terminal r position, the r pattern in continuous speech frequently shows only a portion of the glide. In *ear* bars 2 and 3 of the vowel ɪ (*it*) curve downwards to join with bars 2 and 3 of r, while in *or* the direction of their curves is reversed. When the vowel is the same, the direction of the curves of the bars for an r-vowel glide is reversed in a vowel-r glide.



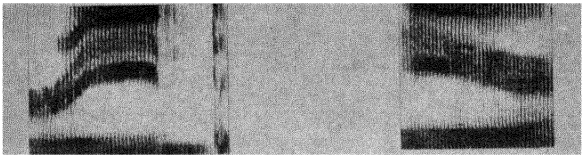
ear
Eastern Pronunciation

or
Eastern Pronunciation

To permit a visual comparison of the General American and Eastern pronunciation of r in the same words, patterns showing both pronunciations are necessary. In the patterns of the Eastern pronunciation of *ear* and *or*, a *schwa* vowel, ə (*about*), is portrayed in place of the r. The visible pattern of ə differs from the final General American r pattern mainly in the positions of bars 2 and 3. Bar 2 for ə is lower than bar 2 of r; while bar 3, instead of being joined with bar 2 as it is for r, is in a position similar to that of bar 3 of the vowel ɪ (*it*). Bar 1 is off the baseline about the same distance for both the ə and the final glide r. The *schwa* that follows the vowel ɔ in *or* is not noticeably different from that used with the ɪ in *ear*.

How the r influences vowel patterns

The initial glide r in *read* (see the illustration at the top of the next page) illustrates the r influence on the vowel i (*eve*). Bars 2 and 3 at the beginning of the r are together below the middle of the pattern. Bar 1 of the r begins on the baseline. Since bar 1

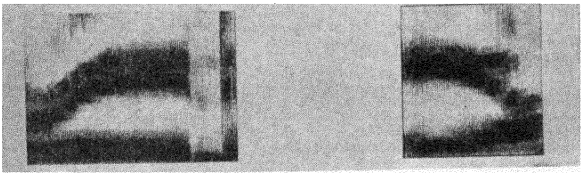


reed, read
General American
Pronunciation

ear
General American
Pronunciation

for both the r and i are on the baseline, their positions in *read* remain unchanged. Bars 2 and 3 of r curve upwards to join bars 2 and 3 of i, and bar 4 of i also curves upwards from a hidden bar 4 resonance of r.

In *ear*, the glide r is in the final position following the vowel ɪ (*it*). Bar 1 of ɪ curves upwards a little to join with bar 1 of r, which is usually off the baseline in vowel-r combinations. Bars 2 and 3 curve downwards from their ɪ positions to the r positions, which are slightly higher for the final than for the initial r. Bar 4 of ɪ slopes downwards a little, but is missing in the r part of the pattern. It is important to note the strong resemblance, in visual as well as aural patterns, among the initial and terminal glide r's and the General American r.



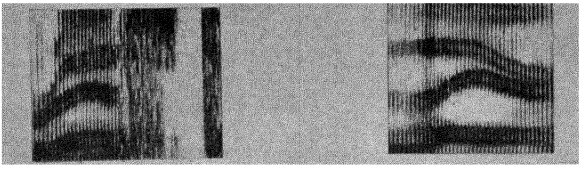
reed, read
Eastern Pronunciation

ear
Eastern Pronunciation

The pattern of the Eastern pronunciation of the initial glide r, as in *read*, is the same, except for variations due to vocal pitch and quality of the speakers, as the pattern of the General American initial glide r. Consequently, the initial r influence on the vowel i, as in *read*, is the same in both pronunciations.

In the Eastern pronunciation of *ear*, the *schwa* vowel ə (*about*) is used in place of the final glide r. This ɪ–ə combination is essentially a diphthong, and its pattern is unlike the ɪ–r pattern.

The positions of the resonance bars at the origin of the r in *rest* are approximately the same as they are in *read*. Because second and third vowel bars are lower for ɛ (*rest*) than for i (*read*), they are nearer to the beginning positions of bars 2 and 3 of r. This makes the curves in bars 2 and 3 in *rest* shorter than they are in *read*. Bar 1 of r curves

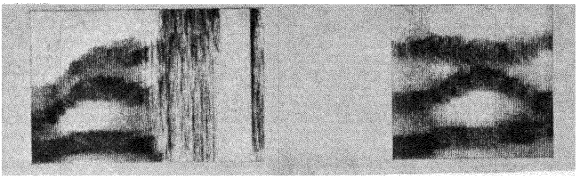


rest
General American Pronunciation

there, their
General American Pronunciation

upwards from its position on the baseline to join bar 1 of ɛ, which is off the baseline. The influence of r on ɛ is to curve the beginnings of all the resonance bars upwards from the r.

In *there*, the r follows the vowel ɛ and terminates the word. Bars 2 and 3 of ɛ curve downwards to r positions that are slightly higher than their beginning positions in *rest*. Bar 1 of the final r is off the baseline about the same distance as bar 1 of the vowel, therefore there is no shift in position of bar 1 in this combination.

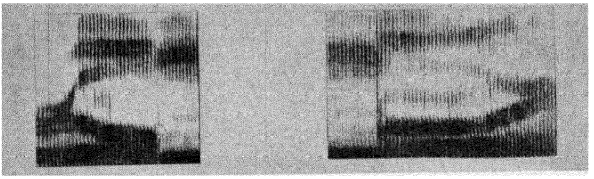


rest
Eastern Pronunciation

there, their
Eastern Pronunciation

Except for differences in the pattern due to vocal pitch and quality, and to speed of articulation, the initial glide r in *rest* in Eastern pronunciation is the same as it is in General American pronunciation. The final glide r in *there* is again replaced by the *schwa* ə (*about*), and may be read as such in visible speech, or transposed to r, depending upon the speech and purpose of the reader.

The glide r made before the vowel u as in *room* (see the illustration on the next page) alters the beginning of the vowel bars in a way that makes this combination easy to recognize. The beginnings of bars 1, 2, and 3 of the r are in their usual positions. Bar 1 of the u is on the baseline so that bar 1 of r joins with it without changing position. Bar 2 curves downwards from the r hub into bar 2 of u. This curve is greater than it is with any of the other r-back vowel combinations, because bar 2 position for the u is lower than it is for any of the other back vowels. Bar 3 of the r curves upwards from its beginning position to the position of the hidden or visible

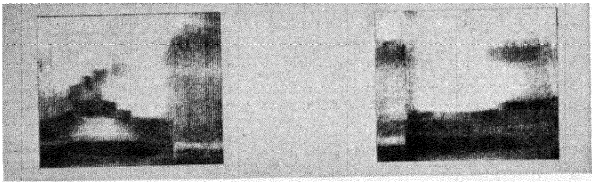


room
General American Pronunciation

moor
General American Pronunciation

bar 3 of u. Bar 4, when present, also curves upwards from its r position to join with bar 4 of the u when it is present.

When the glide r follows the vowel u, as in *moor*, the curves of bars 2 and 3 from the u to the r are the reverse of the r–u curves. The ending of bar 2 of the u curves upwards to the r hub, and the ending of bar 3 curves downwards to bar 3 of r. Bar 1 of the vowel joins bar 1 of the r with little change in position.



room
Eastern Pronunciation

moor
Eastern Pronunciation

As previously mentioned, the Eastern pronunciation of the initial glide r is the same as the General American pronunciation of this sound. This is illustrated again in the pattern of the word *room*. In *moor*, the final glide r is replaced with a *schwa* ə (*about*), just as it was in *ear* and *there*.

In r-vowel and vowel-r combinations, the part of the r pattern from which and to which the glide curves is relatively fixed. However, the positions of the resonance bars of this part of the r sound are generally higher in the final position than in the initial position. Bar 1 is generally off the baseline in the pattern of the terminal glide r, while it is usually on the baseline in the pattern of the initial or beginning glide r. Bar 2, or the hub of r, customarily is just below bar 3, and frequently these bars appear to overlap. Their general position, at the beginning of an r-vowel, or at the termination of a vowel-r glide, is below the center of the pattern. There is some variation in this position, depending upon the vowel that is combined with r; but for the purpose of reading, the variations are not great enough to cause confusion.

The glide portion of the r is portrayed by the curves of the resonance bars. The direction and extent of the curves are dependent upon the positions of the corresponding resonance bars of the combined vowel. If bars 2 and 3 of the combined vowel are near to each other as in i (*read*), bars 2 and 3 of the initial r will curve upwards in parallel. After i, the second and third resonance bars of the vowels as said in series have progressively lower positions. Bar 3 of the vowels never reaches as low a position as bar 3 of r, so that its curve from the initial position is always upwards. Vowel bar 2 is higher than r bar 2 for all of the front vowels, and is lower for most of the back vowels. Therefore, the curve from the beginning r may be upwards, negligible, or downwards, depending upon the combined vowel. Bar 1 of r is on or near the baseline, so its least influence is on bar 1 of vowels i (*eve*) and u (*boot*), and its greatest influence is on bar 1 of the vowel ɑ (*father*).

Eastern pronunciation of the initial glide r is the same as the General American pronunciation, but the terminal glide r is pronounced by many native Easterners as a *schwa*, ə (*about*). The vowel-ə combination is as easily recognized in visible speech as the vowel-r and the ə may be read as ə or transposed to r depending upon the purpose of the reader.

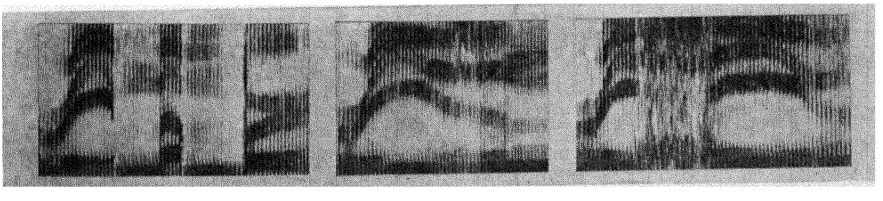
Read the r *by*:

1. The appearance and position of the vowel-like resonance bars at the beginning of the glide in r-vowel combinations.
2. The appearance and position of the vowel-like resonance bars at the ending of the glide in vowel-r combinations.
3. The direction of the curves of the resonance bars.
4. The influence of the r on the pattern of the combined sound.

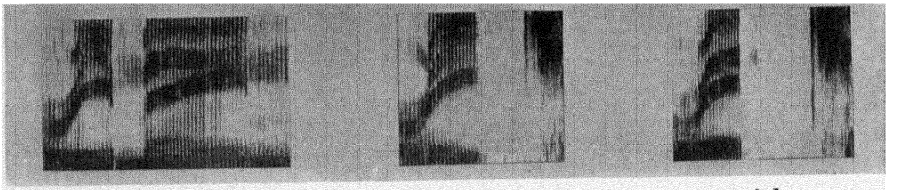
Review words

are, page 70	***rouge***, page 176	***finger***, page 192	***warmth***, page 209
rest, page 223	***read***, page 220	***room***, page 224	***there***, page 223
ear, page 222	***moor***, page 224	***four***, page 62	

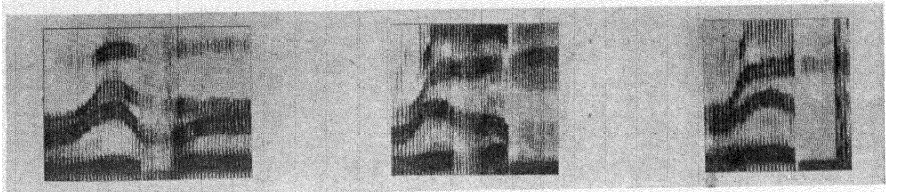
New words to learn



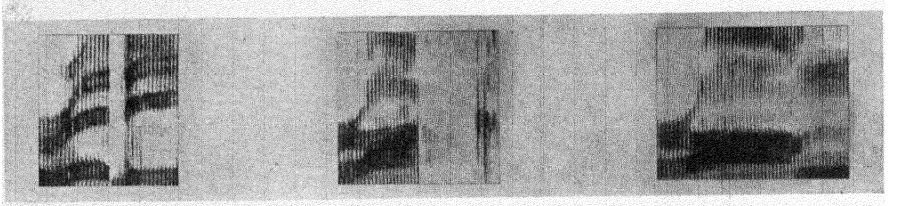
remember ***reason*** ***receive***



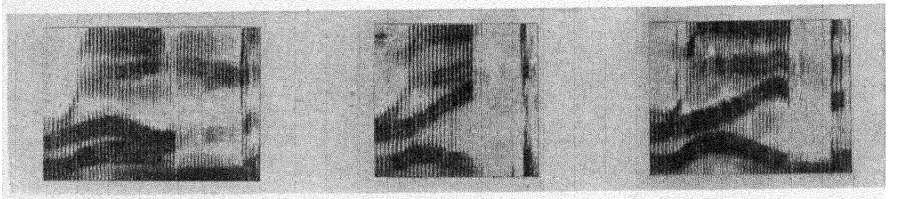
remain *reach* *rich*



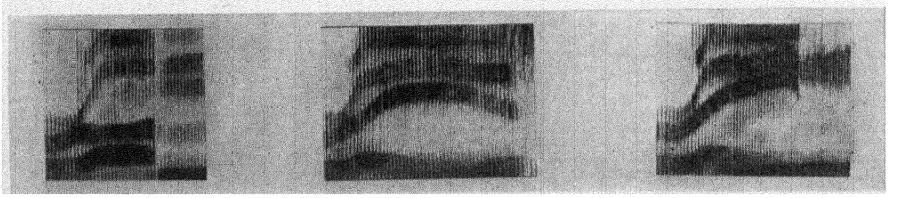
river *rhythm* *red, read*



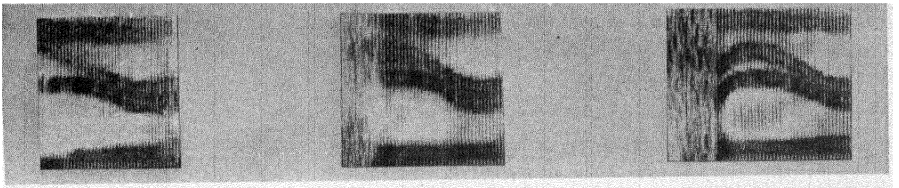
ready *rock* *wrong*



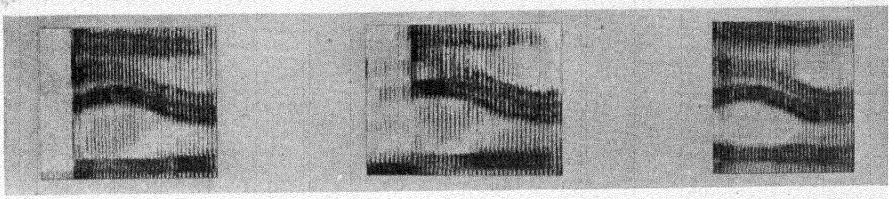
round *right, write* *ride*



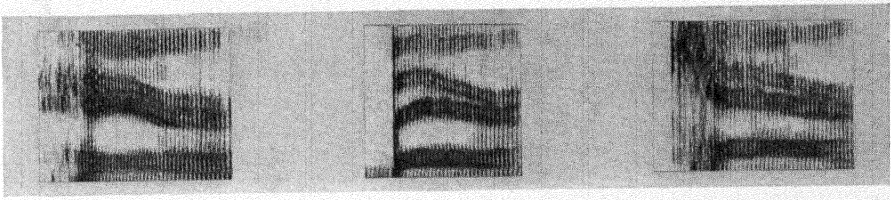
run *raise* *rain, reign*



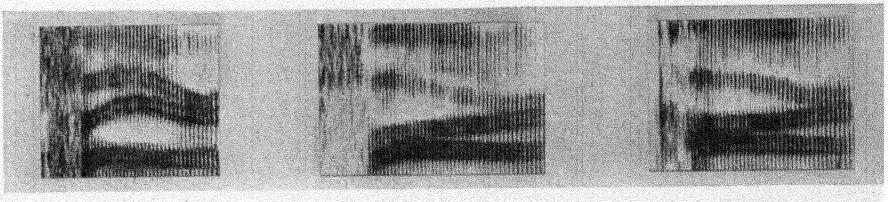
year *here, hear* *fear*



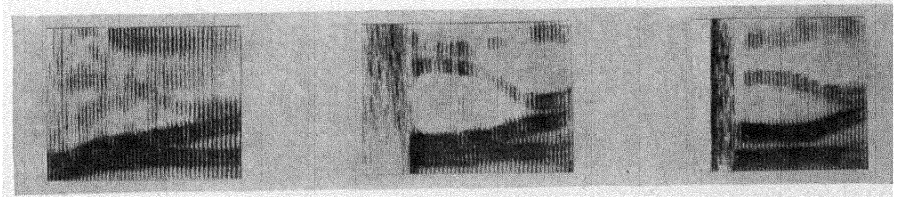
dear, deer *near* *air*



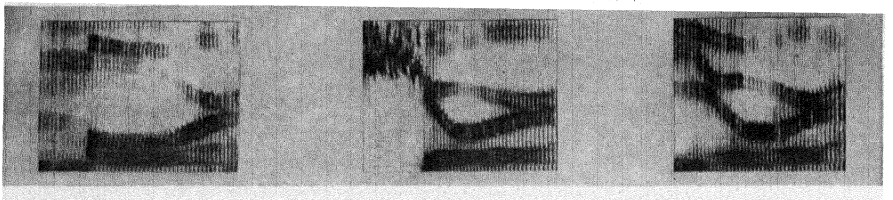
hair, hare *bear, bare* *care*



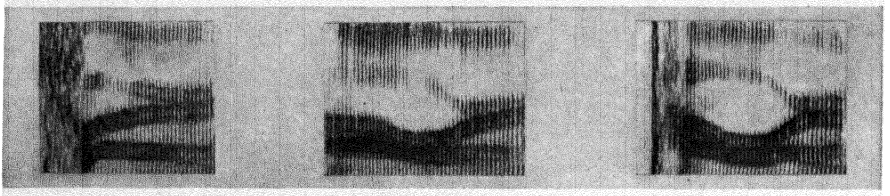
fair, fare *far* *car*



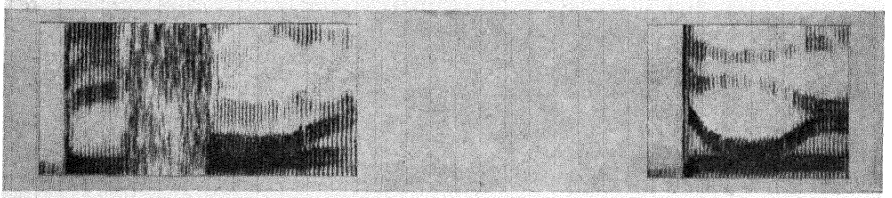
war *for* *poor*



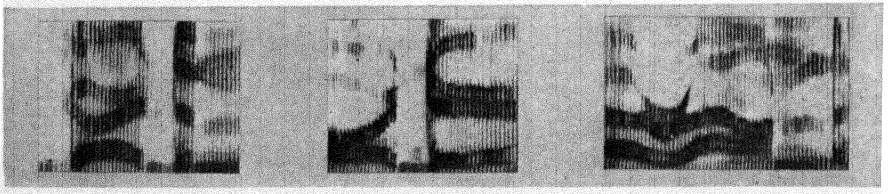
more *sure* *your*



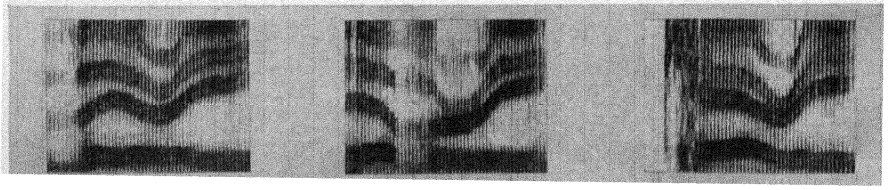
fire *our, hour* *power*



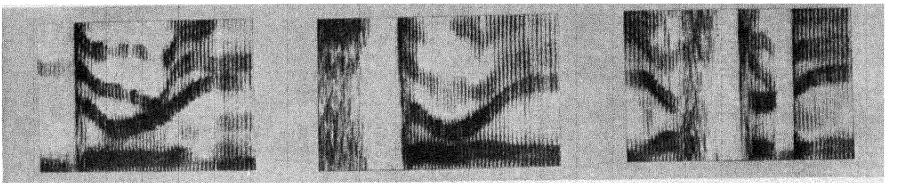
before *door*



garden *order* *around*



very *every* *carry*

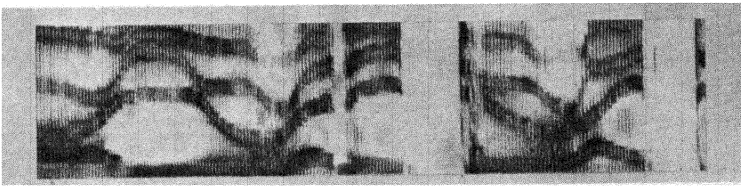


during *story* ***yesterday***

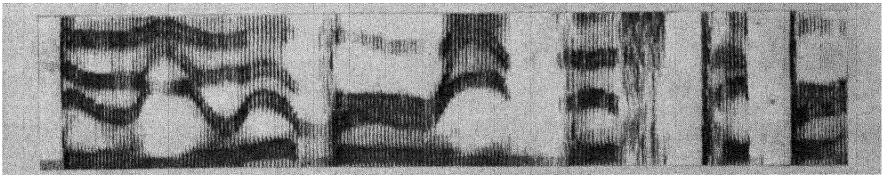
Additional words for practice in visualization

Initial r Words	Medial r Words		Final r Words			
require	*therefore*	*spirit*	*appear*	*tear*	*pair*	*chair*
roof	*market*	*narrow*	*dare*	*star*	*nor*	*therefor*
rose	*tomorrow*	*forest*	*pure*	*figure*	*tire*	*entire*
rather	*exercise*	*hurry*	*desire*	*store*	*shore*	*teacher*
row	*necessary*	*direct*	*either*	*dinner*	*member*	*weather*
ran	*farmer*	*party*	*whether*	*center*	*manner*	*master*
race	*separate*	*corner*	*rather*	*doctor*	*daughter*	*corner*
rapid	*understand*	*interest*	*quarter*	*honor*	*offer*	*officer*
rise	*everything*	*army*	*sugar*	*suffer*	*wonder*	*discover*
ring	*afternoon*	*forget*	*neighbor*	*nature*	*enter*	*however*
rush	*arrive*	*quarter*	*wear*	*require*	*father*	*butter*
ruse			*square*	*neither*	*former*	*favor*

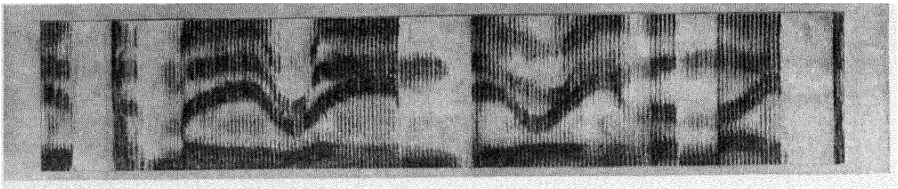
Drill phrases and sentences



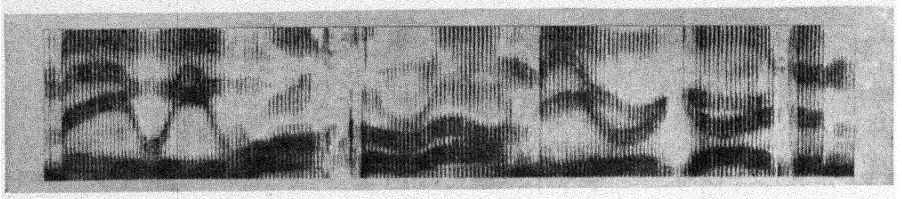
Are you ready to write?



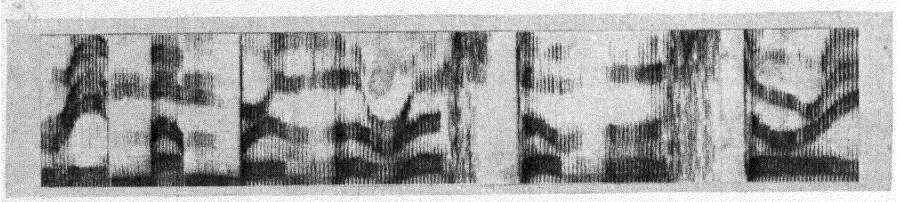
Do you ever read this paper?



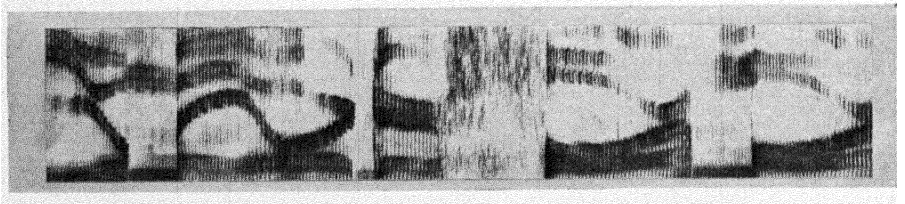
It may rain during the night



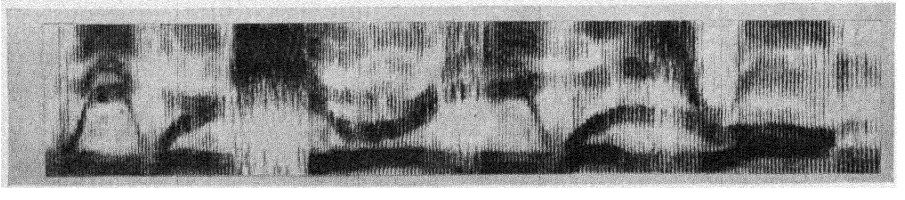
May we walk around your garden?



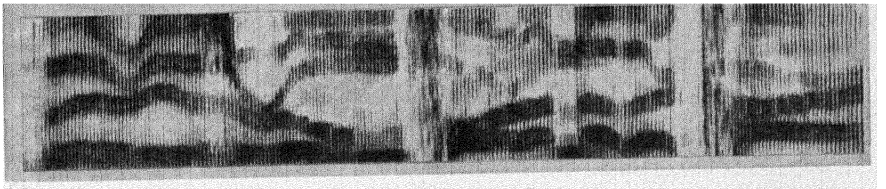
Remember the rest of the story



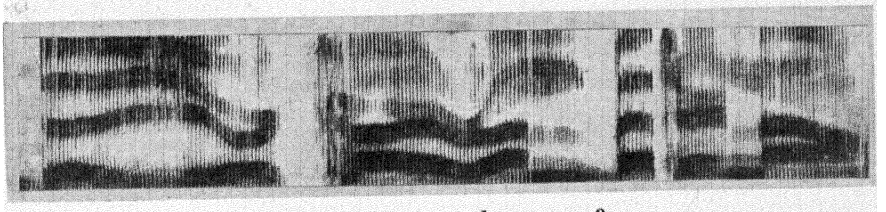
You may order four more



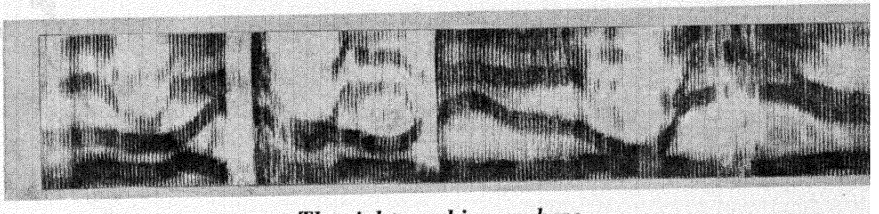
We were sure he was wrong



There is room for another car



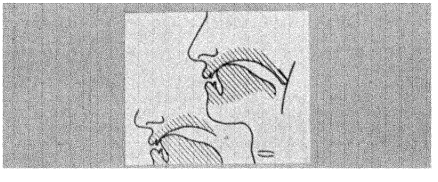
Does your car run better now?



The right road is over here

Unit 4 — Reading the l (let)

How the l (let) is made



Profile of l (let)

The l (*let*) *is produced by:*

1. Assuming a position of the lips similar to the lip positions for ʌ (*up*).
2. Placing the tip of the tongue in contact with the alveolar ridge or the adjacent hard palate depending upon the l being produced.
3. Holding the sides of the front of the tongue against the upper front teeth.
4. Maintaining openings between the back sides of the tongue and the upper molars on each side of the mouth.
5. Closing the opening between the throat (pharyngeal) and the nasal cavities.
6. Emitting a continuous stream of voiced breath while the articulators are moving towards the positions of the combined adjacent sound.

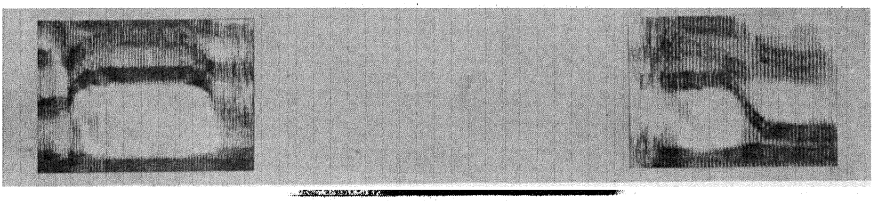
The above description applies to the initial glide l. Like the other glides, the l is produced by vocal cord and varying cavity modulation. But unlike every other sound in the English language, the l is made normally by a lateral emission of the voiced breath stream. This provides a unique cavity modulation which is varied by the position of the tip of the tongue against the alveolar ridge or anterior hard palate, the extent of depression of the back part of the tongue which widens the lateral openings between the sides of the tongue and the upper molars, and the positions of the lips.

When the tip of the tongue is in contact with the alveolar ridge, it is in the position to make the *front* l; when the tongue is retracted and the tip is in contact with the anterior hard palate, the position is that for producing the *back* l. From this physiological point of view, the terms *front* and *back* refer to the positions of the tip of the tongue in producing the l sounds. However, the terms are also used to designate the l sounds used with the front vowels and those used with the back vowels. Since a front tongue point l position normally is used with the front vowels and a back tongue point l position generally is used with the back vowels, both uses of the terms are descriptively appropriate in a broad sense. Each of these positions is illustrated in the above profiles. But actually the l phoneme consists of a series of l's made from several tongue point positions, depending upon the types of modulation used in producing the combined sound. Frictional modulation is present frequently when the l is made with a fricative, and in this respect the l is similar to the r. Also, like the r, the l may be described as being produced from two articulatory phases. One phase is that portion of the sound that may be continued from a steady state position of the articulators. The other phase is the glide portion of the sound which is produced while the articulators are in motion and thus varying the size and shape of the resonance cavities. In the glide l, the moving phase of the articulators predominates, while in the production of other l's the steady state phase may predominate.

The glide l may be spelled in the following ways

l (*lie*)	*ll* (*holly*)	*el* (*shrivel*)
le (*little*)		*sle* (*aisle*)

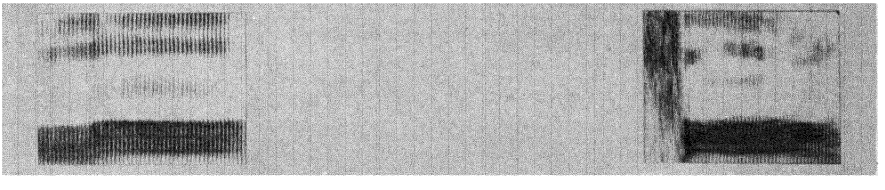
Visible patterns of the l



leave *heel*

Like the other glides, the l is made with vocal cord and varying cavity modulation; therefore, its pattern consists of resonance bars. The positions and curves of the resonance bars of the l are as variable as the articulatory positions used in producing the sound. As explained above, the articulatory positions shift with the different sounds that are combined with l, so there is a concomitant variation in the pattern of the l. The principal change in the initial l pattern is portrayed by the position of bar 2 in the portion of the sound that may be produced from a steady state position of the articulators. This is the first part of the sound, and it determines where the glide begins.

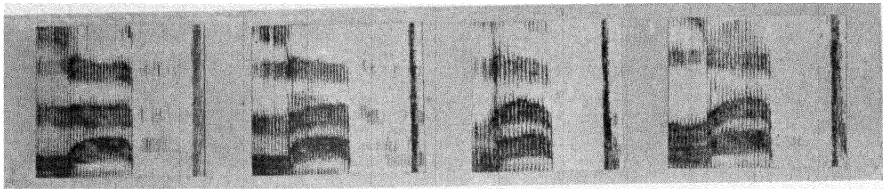
When l precedes the front vowel i as in *leave*, bar 1 of the beginning part of the l is on or near the baseline, bar 2 is near the position of bar 2 of the vowel æ (*at*), and bars 3 and 4, when present, are towards the top of the pattern. The glide portion of the l, made while the articulators are moving towards the i position, is portrayed by the curves in the resonance bars. In *heel*, the l follows the vowel i and terminates the word. The pattern of the final glide l differs from the initial glide l pattern in that bar 1 is farther from the baseline and bar 2 is in a much lower position. The glide portion of the l is shown in the curves of the bars from the i positions to the l positions. These vowel-l curves are unlike the l-vowel curves because of the difference in the positions of the resonance bars of the initial and final l's. Patterns of the final l's are not as variable as the patterns of the initial l's; and, regardless of the vowel they follow, they resemble the l's made with the back vowels.



law *fall*

The first part of the pattern of the l in *law* is characterized by bar 1 near the baseline, bar 2 immediately above bar 1 and adjacent to it, and bar 3 towards the top of

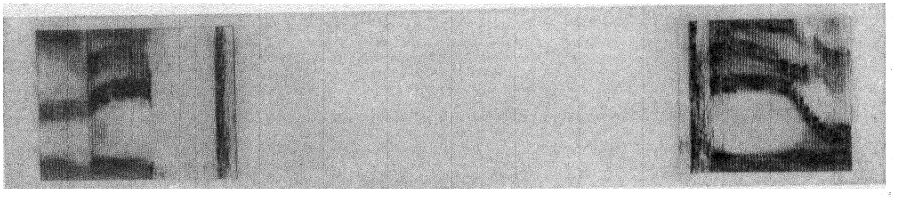
the pattern. Because of the proximity of the l and ɔ (*all*) sounds, the glide phase of the l is short. In *fall*, the l follows the vowel ɔ, and the positions of bars 1 and 2 are a little higher than they are in the initial l–ɔ pattern. The glide phase of the ɔ–l combination is almost negligible since the positions of the ɔ–l resonance bars are almost identical. (For this reason the ɔ–l combination is difficult to identify unless the upper bars are present.) Patterns of the final l's in *heel* and *fall* are very much alike, but the initial l patterns in *leave* and *law* are very different. In *leave*, bar 2 of the l pattern is near its maximum distance from bar 1, while in *law* bar 2 is as close to bar 1 as it can be. These patterns illustrate the extreme positions of the variable l hub.



lap ***lap*** ***lap*** ***lap***

Four different l's were spoken with the same vowel to produce the four patterns of *lap*. In the first pattern, bar 2 of l and æ (*at*) are about in the same position. This l is similar to the l in *leave*, which frequently is called the *front* or *light* l. The second pattern shows an l pattern with bar 2 in a slightly lower position than it has in the first pattern. This change in position of bar 2 is due to a shift in cavity modulation, which is the result of varying the articulatory positions from which the sound is produced. In the third and fourth patterns other articulatory positions were used, and in the fourth, bar 2 of l is so low that it appears to be adjacent to bar 1. This l is very similar to the l in *law*, and it is designated as the *back* or *dark* l. There are, however, many other variations of these l's, and they may be differentiated in visible speech patterns mainly by the position of bar 2 or the hub of their patterns.

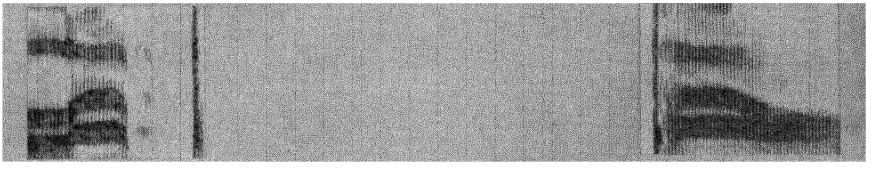
How the l influences vowel patterns



leap ***peel***

The initial l influence on the vowel i (*eve*) is shown in *leap*. Although both bar 1 of the l and of the vowel are on the baseline, there is a shift in width and position of this resonance in the transitional period immediately before the vowel. A similar transition has been observed in the case of the l–u (*boot*) combination, and it is believed to be related to the shift from lateral emission to the high front or back tongue position characteristic of these vowels. Bar 2 of l is near the center of the pattern, and bar 2 of the vowel is in a higher position. So the l influence curves the beginning of bar 2 of the vowel upwards. The extent of the curve varies with the hub position of l. Bar 3 of the vowel curves upwards from the position of bar 3 of l in a curve that resembles that of bar 2.

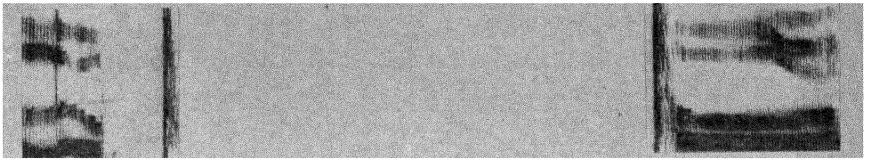
In *peel*, the l follows the vowel i (*eve*) and the l bars are not in the same positions as they are in *leap*. Therefore, the final l influence on i is not merely a reversal of the initial l influence on this vowel. Bar 1 of the final l is off the baseline more than it is for the initial l. This causes an upward curve in bar 1 from its position on the base line for i, to its position above the baseline for l. Bar 2 of the final l is much lower than bar 2 of the initial l, therefore bar 2 of i curves downwards a greater distance in the i–l combination than it curved upwards in the l–i combination. Although bar 3 of the vowel is curved downwards by the final l influence, its curve may be reversed to join with bar 3 of the final l. This sharp reversal of the curve in bar 3 in the i–l combination is associated with the building up of the third bar resonances of the l.



lap ***pal***

In *lap*, the beginnings of bars 1 and 2 are curved upwards from their l positions to their æ positions. Bar 1 of l is near the baseline and bar 2 is much lower than it is in *leap*. However, the l in the l–æ combination may be made, as illustrated previously, so that bar 2 is joined without curving. Therefore, the influence of bar 2 of l on the vowel æ may vary in the speech of an individual speaker or with different speakers. Bar 3 of l glides downwards to join with bar 3 of the vowel, but the shift in position is not as great as it is in bars 1 and 2.

When the l follows the æ, as in *pal*, the position of its bars are about the same as they were in *peel*. Bars 1 and 2 of the vowel curve downwards to join with bars 1 and 2 of l and the curves in this case are more gradual in the æ–l than in the l–æ combination. Bar 3 frequently is absent in the terminal l glide.



loop *pool*

When the l precedes the vowel u as in *loop*, bar 1 continues along the baseline with little change in position. Frequently, there is a shift noticeable in the width and position of the bar at the transitional juncture of the sounds. There is some variation among speakers in the nature of the u vowel used after the l. These modifications in the vowel may be observed in the transitional pattern immediately preceding the vowel resonances. Bar 2 of the l in the l–u combination is usually higher than it is for most of the other l-back vowel combinations. This higher position of bar 2 curves bar 2 of u so that the space between bars 1 and 2 is greater than it is when the u is made in isolation or when the u precedes the l. Bar 3 curves downwards from its l position to join with bar 3 of u. The beginning of the curve in bar 3 is similar to that of bar 2, but the bar 2 curve continues under the influence of the movements to make the closure for the p.

Final l influence on the vowel u is shown in *pool*. Bars 1 and 2 of the l are in approximately the same positions as they are in *peel* and *pal*. Bars 1 and 2 of u are in slightly lower positions than bars 1 and 2 of the l; therefore, there is not much shift in the position of these bars in the u–l combination. The endings of bars 3 and 4 curve downwards from their u positions to their l positions, and are helpful in identifying this combination of sounds.

The positions of the l bars from which the initial glide l begins are variable, depending in part upon the vowel with which the l is combined. Since the beginning of bar 1 is always on or near the baseline, variations in the beginning of the initial l patterns are portrayed mainly in the position of bar 2. It may have its highest position in combination with the vowel i as in *leap*, and its lowest position in combination with the vowel ɔ as in *law*. Within this range, the beginning of bar 2 of l may be varied, so that it is difficult to predict its influence on bar 2 of a combined sound. In general, the beginning of bar 2, or the hub of the initial l, is lower than bar 2 of the front vowels and is higher or in the same position as bar 2 of the back vowels. Bar 3 of the beginning of the l glide is near the position of bar 3 of the vowel ɛ (*let*). This position may vary, but not as much as the position of bar 2. The direction of the glide phase of the initial l is determined by the positions of the beginning of the l bars as well as by the positions of the bars for the combined sound.

Final glide l differs from initial glide l in the positions of its bars, and in the relatively fixed nature of its terminal pattern. The terminal part of the final l glide is that phase of the sound that can be made continuously or from a steady state position. It is the sound to which the glide moves, and it corresponds in this respect to the first phase of the initial glide l from which the glide moves. Bar 1 of the terminal part of the final glide l is usually off the baseline and is separated from bar 2 by a small space. Bars 3 and 4, when present, are towards the top of the pattern. The glide phase of the final l glide is usually more gradual than the corresponding phase of the initial glide, and the direction of the curves are predictable from the positions of the bars of the preceding sound.

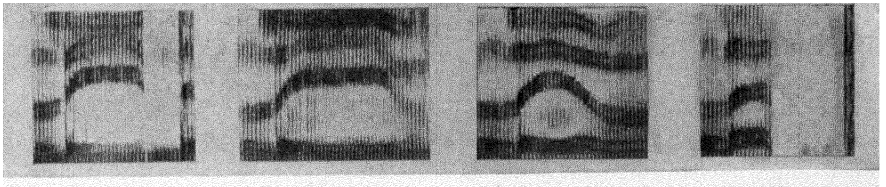
Read the l *by:*

1. The appearance and position of the vowel-like resonance bars at the beginning of the glide in l-vowel combinations.
2. The appearance and position of the vowel-like resonance bars at the ending of the glide in vowel-l combinations.
3. The direction of the curves of the resonance bars.
4. The influence of the l on the pattern of the combined sound.

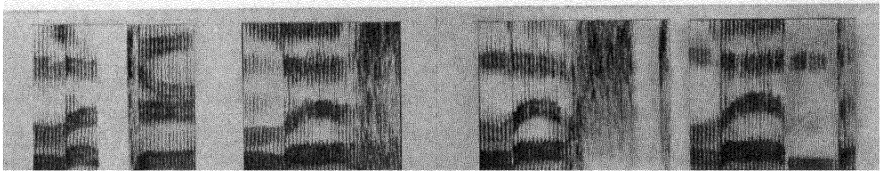
Review words

leave, page 233	*law*, page 233	*heel*, page 233	*length*, page 195
leap, page 234	*loop*, page 236	*peel*, page 234	*pool*, page 236
lap, page 234, 235	*fall*, page 233	*pal*, page 235	*casual*, page 145
			all, page 70

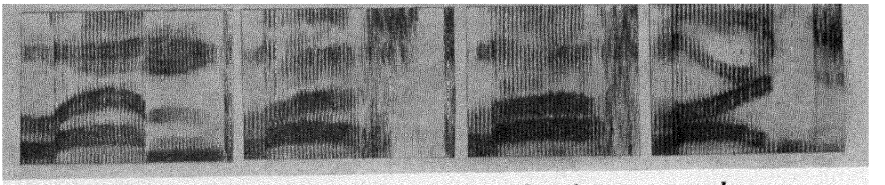
New words to learn



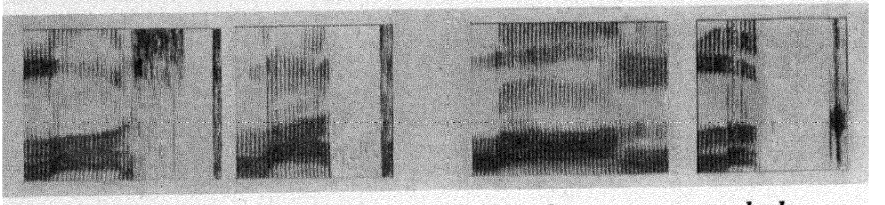
lead *leave* *live* *let*



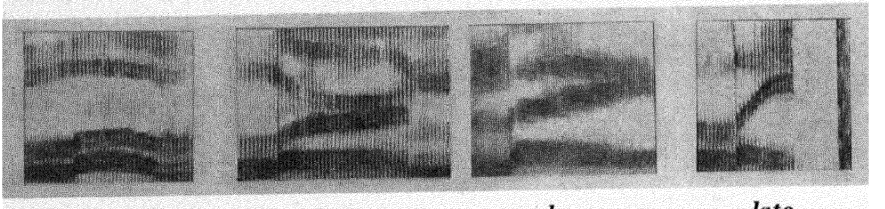
letter *less* *left* *led, lead*



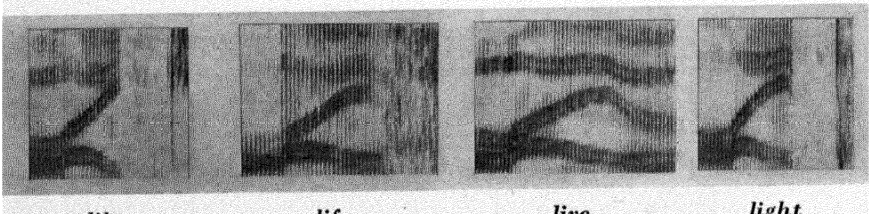
land *last* *laugh* *large*



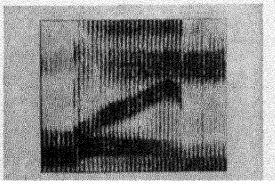
lost *lot* *long* *look*



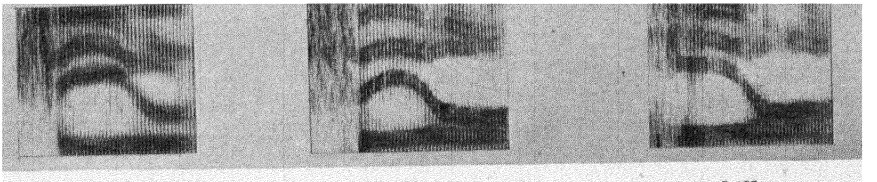
love *learn* *lay* *late*



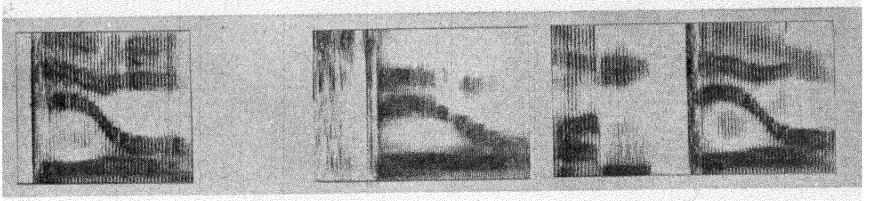
like *life* *live* *light*



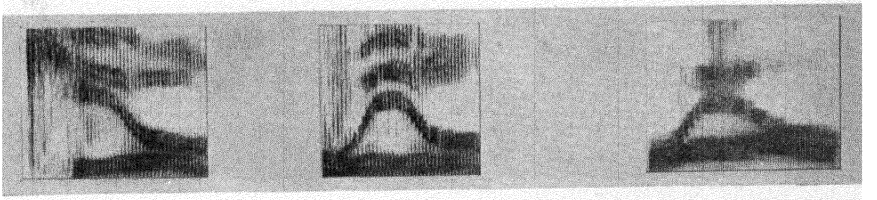
line



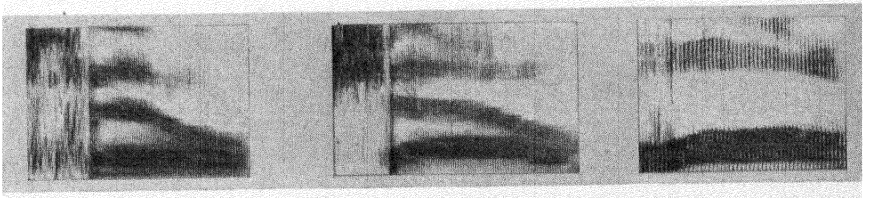
feel *fill* *hill*



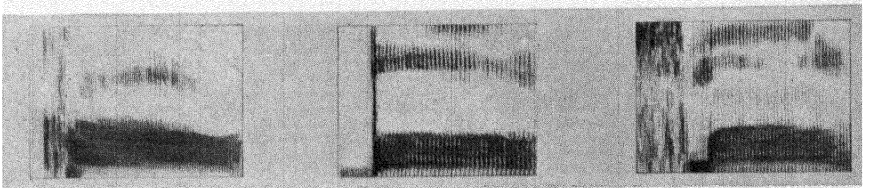
till *still* *until*



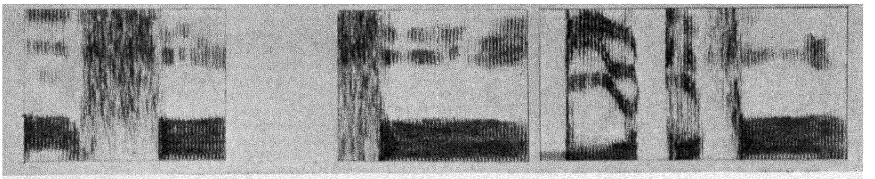
kill *will* *well*



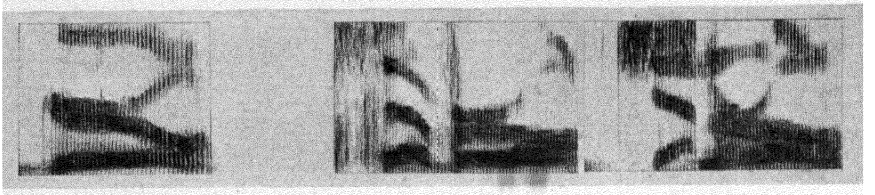
sell *shall* *wall*



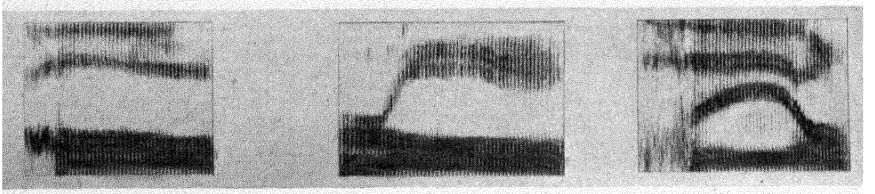
call *ball* *small*



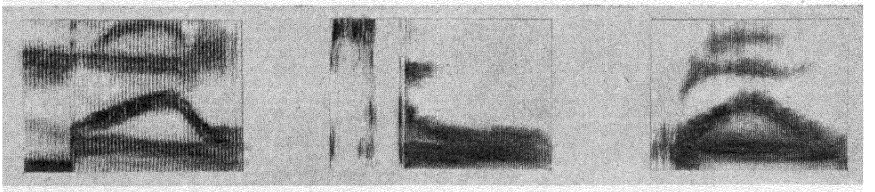
awful *full* *beautiful*



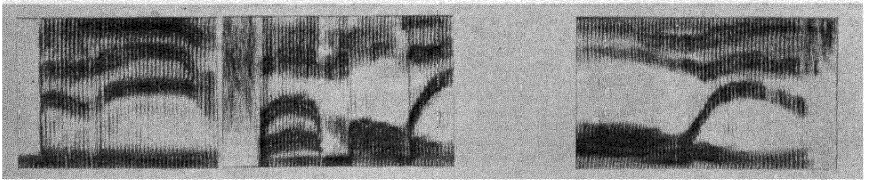
girl *several* *general*



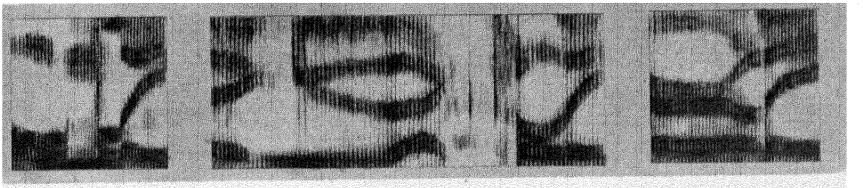
hole, whole *roll* *sail, sale*



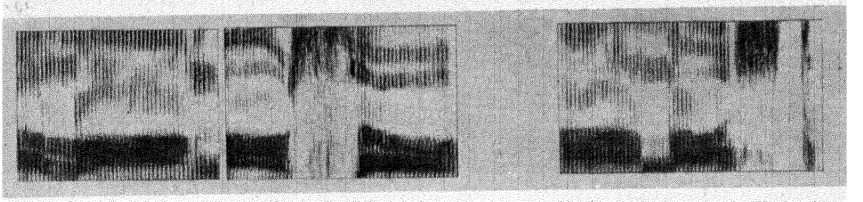
mile *school* *while*



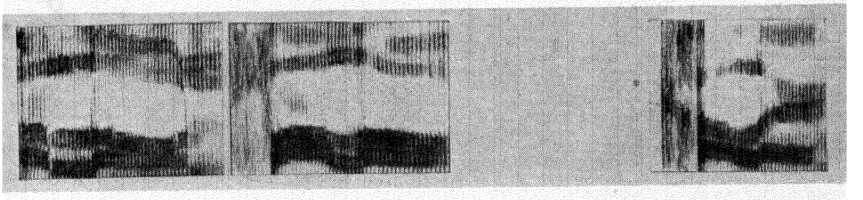
believe *family* *always*



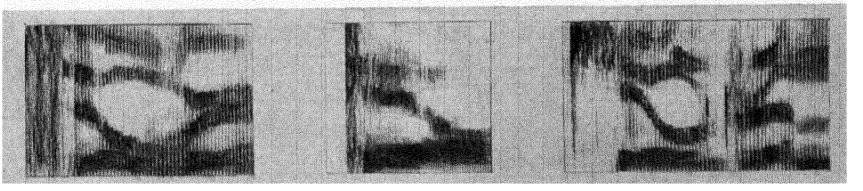
only *exactly* *early*



along *also* *almost*



alone *follow* *color*

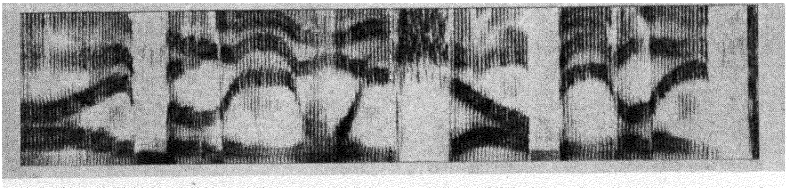


silver *tell* ***children***

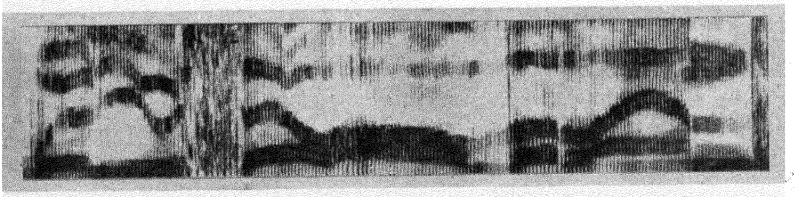
Additional words for practice in visualization

Initial l Words		Medial l Words		Final l Words	
least	*lift*	*village*	*valley*	*deal*	*wheel*
listen	*lesson*	*fellow*	*already*	*bill*	*fell*
lord	*laid*	*following*	*delight*	*tall*	*pull*
lake	*lie*	*shoulder*	*jelly*	*rule*	*smile*
loud	*lower*	*yellow*	*public*	*equal*	*travel*
lip	*list*	*although*	*value*	*animal*	*wonderful*
leg	*loss*	*allow*	*island*	*ill*	*real, reel*
lose	*lady*	*cellar*	*colder*	*bell*	*mill*
labor	*lion*			*cool*	*hall*
load	*lodge*			*soil*	*tail, tale*
				natural	*careful*
				towel	*usual*

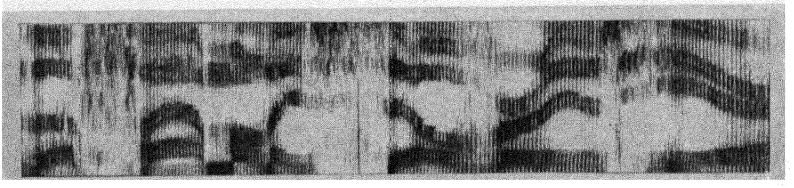
Drill phrases and sentences



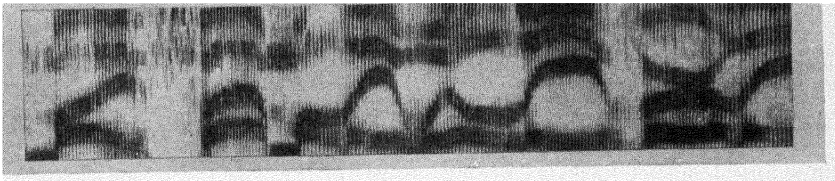
I believe we shall be late



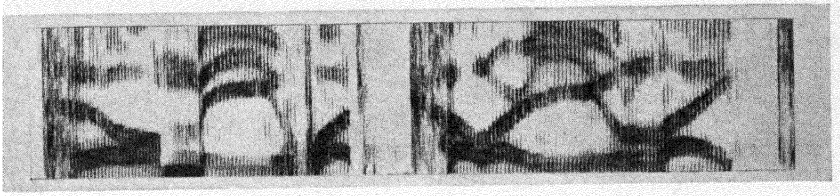
Will you sell all the land?



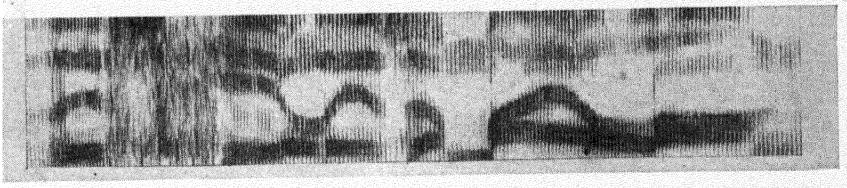
The family silver is here



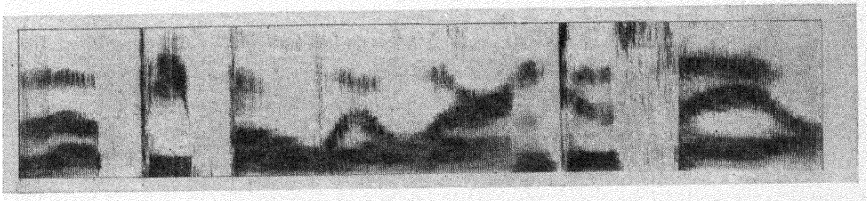
My family will leave early



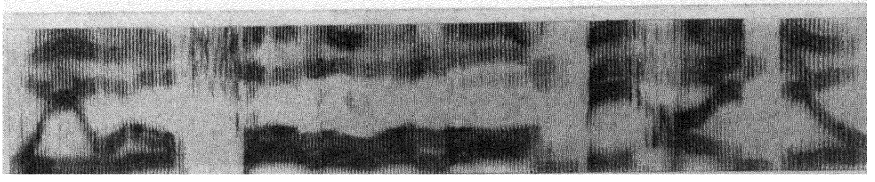
Tell me what color you like



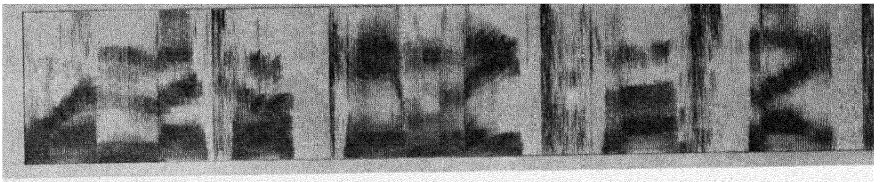
This hill is a mile long



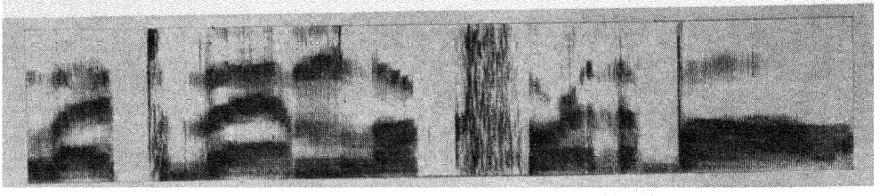
The people will learn to sail



We will follow along behind you



One letter came late last night



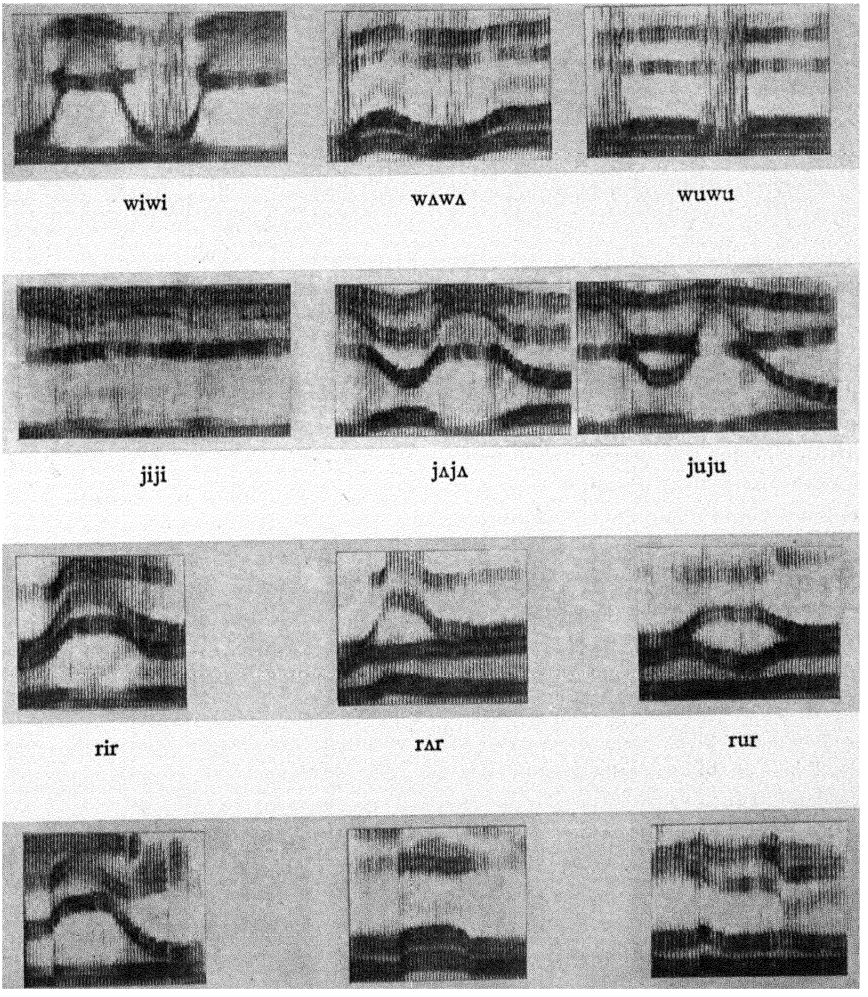
Let them look for the ball

Unit 5 — Comparative Review of the Glide Sounds

A comparative review of the visible speech patterns of the glides, their influences, and the unique characteristics by which they may be differentiated, is presented in this unit. Each of the four glides is produced by a characteristic combination of vocal cord and cavity modulation. Cavity modulation is constant in the steady state portion of the glides, but is varied during the glide portion of these sounds. Patterns of the glides are differentiated by the positions of their resonance bars during the steady state phase, by their hubs, and by the transitional variations in the curves of their resonance bars to patterns of combined sounds. These differences in patterns among the several glides are related to the different articulatory positions and movements used in producing the sounds.

As in the comparative reviews of the preceding chapters, this unit will present both a pattern and written comparison of the glides. The figure below contains patterns of the **w** (*we*), **j** (*you*), **r** (*read*), and **l** (*let*) spoken before and after the vowels **i** (*eve*), **ʌ** (*up*), and **u** (*boot*). In the **w** and **j** patterns, the vowel is repeated because these glides do not ordinarily terminate a syllable. For example, **wiwi** was spoken instead of **wiw**. Consequently, in each case, the second glide occurs in medial position, and influences both the vowel which precedes and the vowel which follows. It is of especial interest to note the effect of the initial and medial glide on the first vowel in these **w** and **j** combinations. The **r** and **l** glides are shown before and after a single vowel, since they occur that way in spoken English and since their patterns are not the same in the

initial and final positions. The hub of l following a vowel is usually lower than the hub of l that precedes the same vowel. The hub of r following a vowel is higher than the hub of r that is spoken before the same vowel. The hub of w is similar to that for the vowel u (*boot*), and the j hub is approximately the same as the vowel i (*eve*).



Each glide has a different hub position

1. The **w** hub is towards the baseline, similar to the **u** (*boot*) hub.
2. The **j** hub is towards the top of the pattern, similar to the **i** (*eve*) hub.
3. The **r** hub is variable between the positions of the **a** (*ask*) and **æ** (*at*) hubs. The hub of a final **r** is higher than the hub of an initial **r**.
4. The **l** hub is variable: towards the middle of the pattern with front vowels, towards the baseline with the back vowels. The terminal **l** hub is usually lower than the initial **l** hub.

Each glide influences the vowels differently

1. The **w** influences the front vowels the most, with a gradual decrease in influence as the vowels are said in series from **i** (*eve*) to **u** (*boot*).
2. The **j** influences the back vowels the most. Its influence on **i** (*eve*) is negligible, but progressively increases as the vowels are said in series.
3. The **r** influences the vowels **i** (*eve*) and **u** (*boot*) the most, and it has the least influence on vowels that occur towards the middle of the vowel series.
4. The initial **l** influence is less marked than the final **l** influence. In the vowel series, the final **l** influences **i** (*eve*) the most, and the back vowels from **ɔ** (*all*) to **u** (*boot*) the least. The initial **l** exerts the same general kind of influence as the final **l** except that the degree of influence is less due to the greater variability of the initial **l** hub.

The glides are among the most easily recognized sounds in visible speech. Their patterns are differentiated from the patterns of vowels and vowel-like sounds by the positions of their resonance bars during the brief steady state portion of the sounds, plus the direction and the appearance of the curves in the resonance bars during the glide portion of the sounds. The extent of any glide varies directly with the distance between the hubs of the glide and the combined sounds. When the hubs of a glide and a combined sound are in similar positions, the glide portion is very short. When the hubs are as far apart as possible, the extent of the glide is at its maximum. Speed and precision of articulation, and concomitant variations in cavity modulation, vary the patterns of each of the glides. To those interested in learning to read visible speech, it is suggested again that words, phrases, and sentences containing glides be visualized and reproduced in graphic form for comparison with the visible speech patterns of the illustrations in the lesson units. Also, as mentioned in the preceding chapters, the association of the articulatory positions and movements with the patterns made from the positions and movements, will facilitate the learning and understanding of the general and the specific concepts of visible speech.

Unit 6 — Reading Combinations of Glides and Other Sounds

In the preceding units in this chapter, the individual glide patterns, hubs, and influences on vowels have been discussed. This unit presents illustrations and explanations of the principal clues for recognition of patterns made by combinations of glides

and sounds other than the vowels. Such combinations include glide and glide, glide and stop, glide and fricative, in elisions that occur within words and also between syllables and words in continuous speech. In general, the pattern of any combination of sounds is characterized by the type of modulations used in producing the separate sounds. However, there are additional characteristic clues inherent in each type of combination, such as glide-stop and glide-fricative, that aid in their recognition.

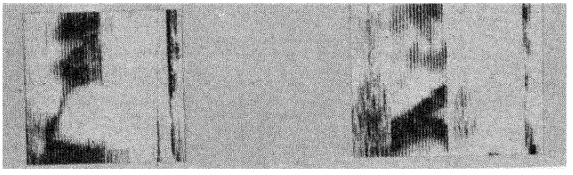
Reading the hw (when) combination

We have classified the **hw** (when) as a combination or Group 6 sound because it is produced with a voiceless **h** preceding a partially unvoiced **w** (*we*) glide. The **h** part of the combination is normally made through a rounded lip position similar to that for **u** (*boot*). The rounded lip position causes the greater portion of this **h** fill to be concentrated at the bottom of the pattern. Usually, the **w** phase of the **hw** combination is classified as voiceless and described as produced from moving articulatory positions comparable to those used for **w** (*we*). Thus, the **hw** has been considered as a voiceless combination. However, the patterns of the **hw** reveal that vocal cord and cavity modulation are introduced near the beginning of the **w** while the articulators are moving. Therefore, the **w** phase of the **hw** pattern is similar to the voiced **w** as in *we*, except that there is some variation in the point at which the voicing begins. A considerable portion of the glide usually is apparent in the curves of bars 1 and 2 from positions near the bottom of the pattern towards the resonance positions of the following sound.

The hw is spelled in the following way:

wh (*when*)

Visible patterns of the hw:



wheat *what*

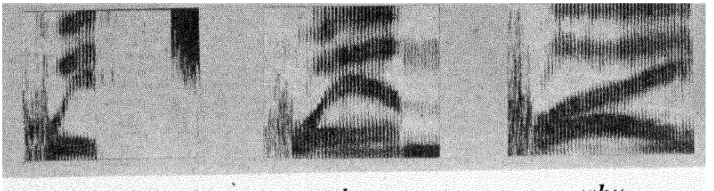
Patterns of *wheat* and *what* illustrate the **hw** pronunciation of the **wh** spelling. In this combination of sounds, the fill for the **h** usually is concentrated towards the bottom of the pattern. The **w** phase of the sound is partly unvoiced and partly voiced. The amount of unvoicing varies with different speakers but, in general, a considerable

portion of the glide is voiced. In *wheat*, the voicing, as shown by the beginning of bars 1 and 2, starts near the positions of bars 1 and 2 for **u** (*boot*). Curves in the bars are produced while the articulators are moving towards the **i** (*eve*) position. A similar proportion of unvoicing and voicing is portrayed in the pattern of *what*, but the curve in bar 2 is not as steep as it is in *wheat* due to the difference in position of bar 2 of the vowels in these words. Patterns of **hw** vary in the amount of **h** fill and in the amount of unvoicing of the **w**. The hub of this combination is approximately that of the **u** (*boot*), and since it does not occur in the final position of words its influence is primarily on the sound that comes after it. The **hw** influence in continuous speech may be observed in the sentences in this section.

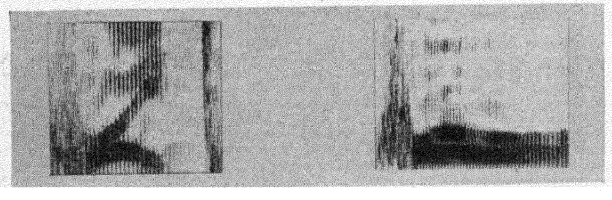
Review words

wheat, page 247 *what*, page 247

New words to learn



which *when* *why*

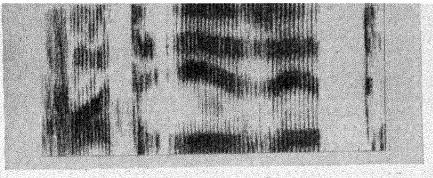


white *whoa*

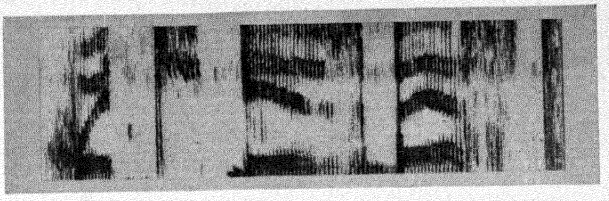
Additional words for practice in visualization

whether	*wheel*	*wharf*	*whip*
whisper	*whisk*	*whistle*	*whirr*
whine	*whim*	*while*	*whiff*

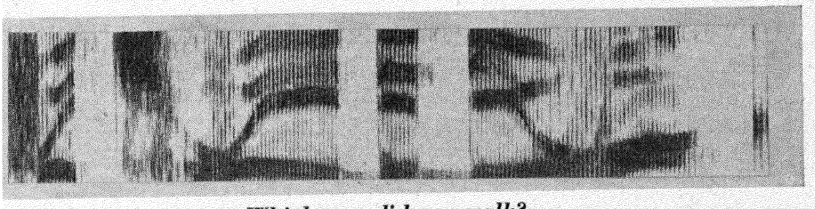
Drill phrases and sentences



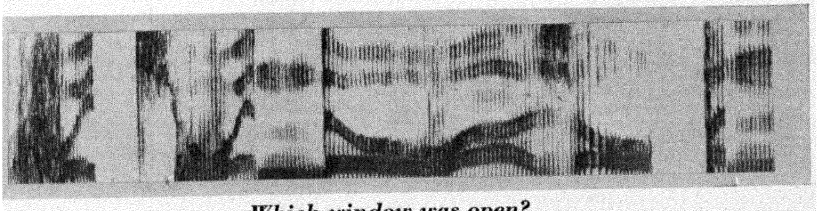
What is it?



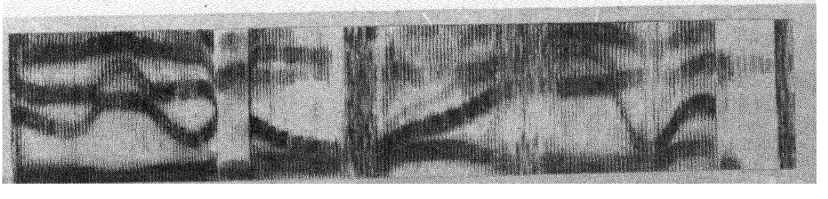
Which is best?



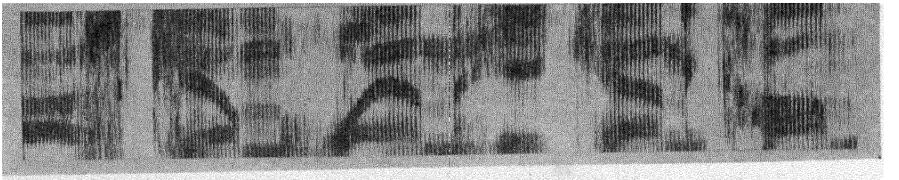
Which way did you walk?



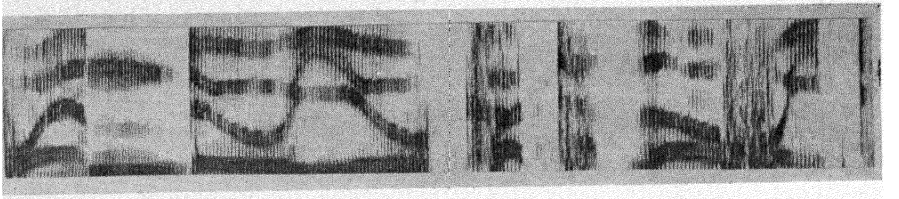
Which window was open?



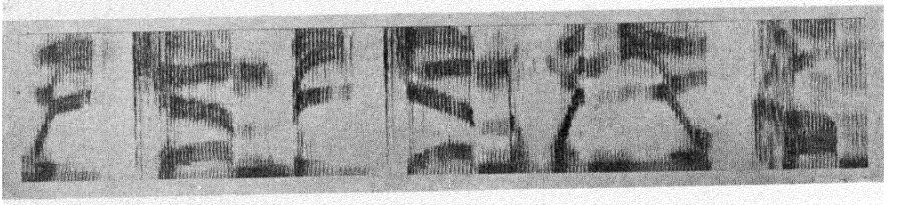
Do you know why he went?



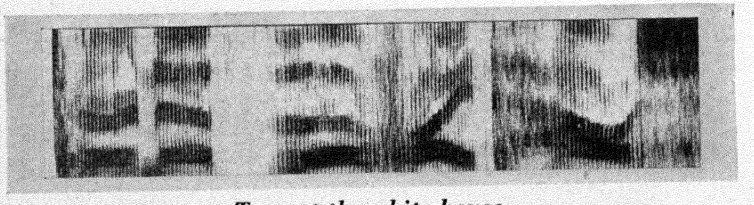
Ask him when he can come



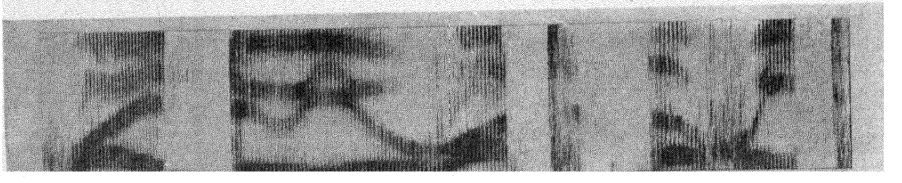
When do you cut the wheat?



We can begin when you come

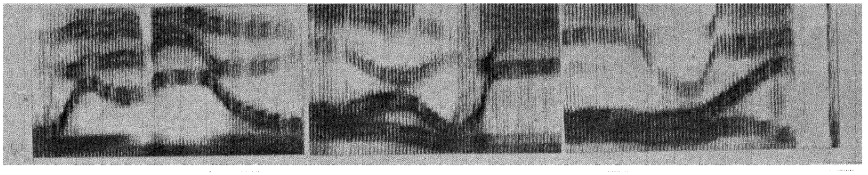


Turn at the white house



Why do you want the wheat?

Unit 7—Reading Combinations of Glides, Stops, and Fricatives

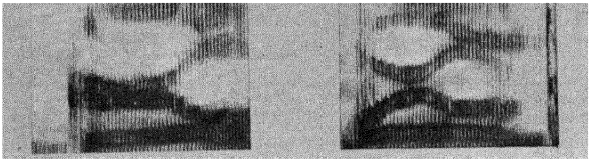


will you *are we* *all right*

Patterns of glide-glide elisions illustrate combinations of two transitional patterns. If they are compared with vowel-vowel, vowel-voiced fricative, or voiced fricative-voiced fricative combinations, differences may be noted in the relative duration of that phase of the combined sounds which may be made from a steady state position of the articulators. For instance, such vowel-vowel combinations as the diphthongs are portrayed by a transitional pattern between two vowel patterns. In terms of articulatory positions and movements, a vowel-vowel elision is produced by a sequence of steady state-varying-steady state articulation. During continuous speech, the duration of the steady state portion of most combined sounds is shortened, but it is reduced to a minimum in glide-glide combinations. Regardless of the duration of the steady state elements of combined sounds, if the sounds are produced at all, the articulators must assume momentarily their respective steady state positions.

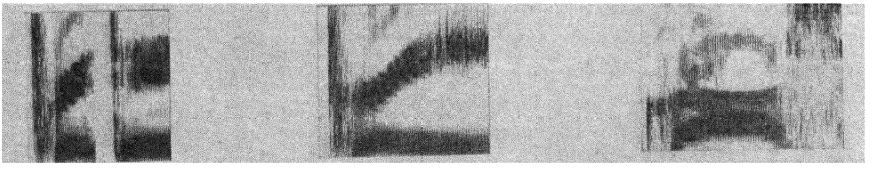
Illustrations of glide-glide combinations are presented in the patterns of *will you*, *are we*, and *all right*. The l–j elision in *will you*, the r–w in *are we*, and the l–r in *all right*, are characterized by continuous curving of the resonance bars. It would be difficult to identify these curving bars if it were not for the unique characteristic previously mentioned. In order to say the glides, the articulators necessarily move from or to their steady state positions. That is, glides, like all other sounds, have a characteristic point of origin or termination both aurally and physiologically. Therefore, glide-glide patterns may be identified in general by the vowel-like appearance of their resonance bars, and specifically by the positions at which the resonance curves change their course. This point of change approximates the position of the hub of the glide, or in other words that phase of the glide which may be produced from a steady state position of the articulators. Even though the resonance curves appear to be continuous, if the combined sounds are articulated, their resonances will curve to, from, or through their hub positions.

The words *girl* and *world*, as spoken with General American pronunciation, contain the ɝ (*bird*) sound. This r is generally considered to be vowelized and to be made from a fixed or steady state position of the articulators. In the pattern of *girl*, the ɝ



girl ***world***

occurs between a voiced stop **g** (*go*), and a glide **l**. Although both the initial and terminal sounds influence the ɜ characteristically, the duration of the ɜ is long enough to illustrate its fixed, vowel-like properties. *World* shows the ɜ occurring between two glides **w** (*we*) and **l** (*let*). In this case, the duration of the ɜ component is short and its pattern resembles that of the glide **r** (*read*). The uninterrupted resonance curves reveal that the articulators were in continuous movement from the **w**, through the ɜ, to the **l**; but the hub position of each sound is indicated at the point where the resonances originate, change course, or terminate. Whenever the larger portion or phase of a sound, such as the ɜ in *world*, is produced while the articulators are in motion, its pattern resembles that of the glides. It is significant that many sounds customarily considered as steady state sounds, assume the characteristics of glides in certain combinations in continuous speech.

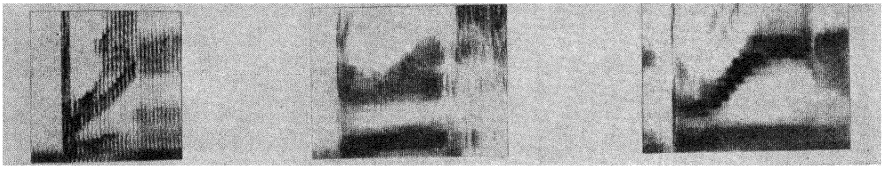


pretty ***tree*** ***cross***

Voiceless stop-**r** glide combinations are made by varying combinations of start stop-frictional-vocal cord-cavity modulation. The transitional pattern between the stop and glide appears partly in the voiceless fill that follows the stop spike and partly in the beginning of the resonance bars of the glide. This is illustrated in the patterns of *pretty*, *tree*, and *cross*. Such combinations are readily identified by the nature of their patterns, and may be differentiated primarily by the positions of the hubs of the stop sounds. In *pretty*, the darker area in the fill following the spike curves upwards from the **p** hub towards the hub of the **r**. Similarly, the darker areas of the **t** fill in *tree* and the **k** fill in *cross*, curve to join the **r** hub.

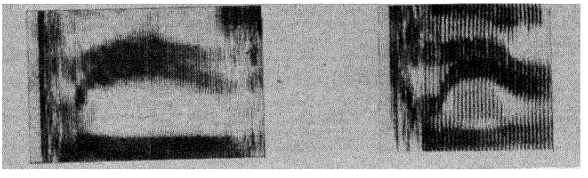
The **r**'s in *pretty* and *tree* usually are referred to as front fricative **r**'s, while the **r** in *cross* is called a back fricative **r**. These terms, front and back, designate the articu-

latory positions from which the sounds are made. The visible patterns indicate that frictional modulation is present in the transition between the stop spike and the first phase of the **r**. This is the component that has been heard by phoneticians and that resulted in the designation of these **r**'s as fricative sounds. However, the positions of the hubs and other resonance bars at the initiation of voicing of these **r**'s, as well as the transitional nature of their patterns, indicate that they are glides produced with frictional, vocal cord, and varying cavity modulation; and that the fricative portion of the glide is related to the transition from stop-frictional modulation. Variations in speed and precision of articulation modify the duration of the steady state phase of the **r** and also the relative durations of the fricative and voiced components.



bring ***dress*** ***green***

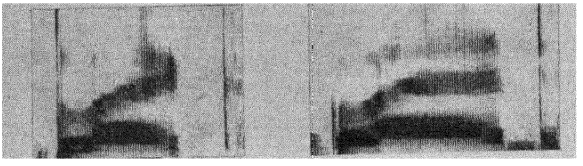
In voiced stop-glide **r** combinations as illustrated in *bring*, *dress*, and *green*, the fricative transitions between the stop and glide are voiced. Because the stop is voiced, the stop spike adjoins the beginning of the resonance bars of the glide. Like the voiceless stop-glide combinations, the voiced stop-glide blends may be identified by the appearance of their combined patterns and differentiated by the hub influence of the stop on the glide. Patterns of the **r**'s following voiced stops indicate the varying duration of the fricative and glide components, and also the stop hub influence on the point of origin of the glide.



please ***clear***

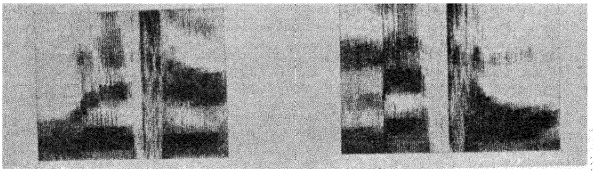
Voiceless stop-glide **l** elisions, as in *please* and *clear*, have characteristics similar to voiceless stop-glide **r** combinations. Frictional modulation following the stop spike continues into the beginning of the **l** just as it does for the **r**. In this respect, these

l's may be considered as fricative glide sounds. Voiceless stop-glide l combinations are identified and differentiated by the appearance of the transitional pattern between the stop and glide hubs. Since both the **k** and l hubs are variable, patterns of the **k–l** may vary among different speakers and in the speech of an individual.



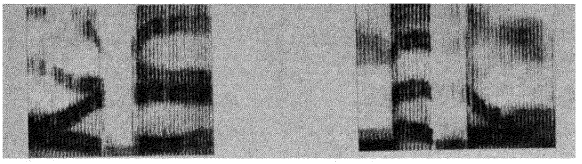
black ***glad***

Voiced stop-l glide elisions, as illustrated in *black* and *glad*, are similar to the voiced stop-r combinations in that the stop spike is adjacent to the beginning of the glide, and the transitional pattern is voiced and joins the hubs and other resonance bars of the two sounds. The proportions of the fricative and voiced components of the transition vary with the rate, precision, and individual characteristics of articulation.



letter ***metal***

In words like *letter* and *metal*, the **r** and l are terminal sounds preceded by a voiceless stop. These combinations are frequently voiceless stop-glide **r** or l elisions. However, when a vowel is used between the stop and the glide, the terminal combination is changed from stop-glide to vowel-glide. The vowel, when present, is usually a form of the *schwa* which varies in duration with the speed of articulation and stress. The **r** in *letter* generally is described as the unstressed form of the fixed vowelized ɜ as in *bird*. Although the resonance bars at the end of the sound are in approximately the same position as the ɜ bars, a considerable portion of the **r** in *letter* is made while the articulators are in motion going from the **t** to the ɜ positions. Therefore, r's of this type may be classified as unstressed fricative glide r's. The **t–r** and **t–l** combinations may be identified by the same clues as the other voiceless stop-glide patterns that have been explained previously.



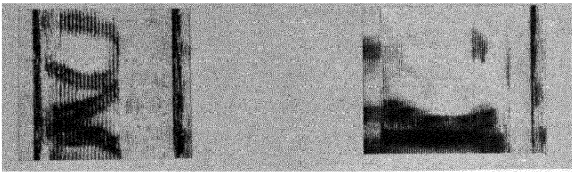
order *medal*

Patterns of *order* and *medal* illustrate representative voiced stop-glide r and l combinations. In these patterns voicing is continuous between the d–r and d–l sounds, so that the transitional patterns are apparent. At first glance, the r pattern, following the stop spike for d in *order*, may appear to be the ɝ (*bird*) pattern. However, upon closer inspection it may be noted that bars 1, 3, and 4 continue to curve as a glide r until they reach a steady state position towards the end of the pattern. Similarly, the transitional pattern between the stop spike for d and the l in *medal* shows that vocal cord and varying cavity modulation are employed until the final part of the l is reached. In this pattern, a schwa vowel was spoken during the d–l transition. When the schwa is used in this way, it generally occurs as part of a prolonged transition, and sometimes the schwa-in-transition appears to be indistinguishable from the voiced stop-glide transition.

between *but you*

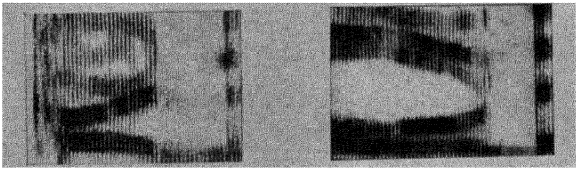
There are many other stop-glide combinations that occur within words, such as *between*, or in phrases like *but you*. The t–w combination in *between* illustrates the continuation of frictional modulation into the w glide. This results in a w which is fricative to the same extent that the r and l are fricative in similar combinations. Such combinations may be recognized by the voiceless stop gap preceding the glide, by the hubs of the combined sounds, and by the characteristic transitional pattern resulting from the shift in modulation. In phrases similar to *but you*, there frequently occurs a voiceless stop-glide elision in which the frictional modulation following the voiceless stop spike may be continued into the beginning of the glide. The beginning of the glide is evident in the darker area at the top of the fill, and the pattern of the j glide above shows that the movements of the articulators are continuous until the

position of the u (*boot*) is attained at the end of the pattern. When the stop and glide are not elided, there is a pause in the movements of the articulators causing a break in the continuity of the glide pattern. The break occurs between the stop fill and the beginning of the bars of the glide. Variations in stress and rhythm, or in speed and precision of articulation, result in effects similar to those discussed in other types of combinations.



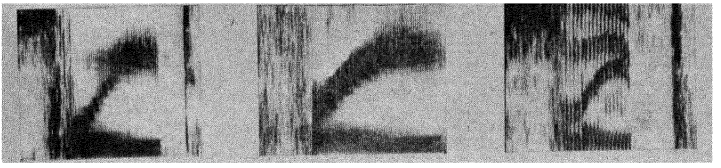
part ***molt***

Glide-to-voiceless stop combinations are characterized by the abrupt termination of vocal cord-varying cavity modulation and the beginning of stop-start modulation. This is illustrated in the patterns of *part* and *molt*. In each pattern, voicing continues while the articulators are moving from the vowel to the glide towards the stop position, as evidenced by the continuous curving of the resonances. That the glide is influenced by the stop is apparent in the positions of both the r and l bars at their terminal points. When the articulators reach the stop position, the influence on the glide ceases, the stop gap begins, and the word is completed by the explosion of the stop. Such combinations are recognized by the characteristic appearance of each element in the combined pattern and by the stop influence on the glide.



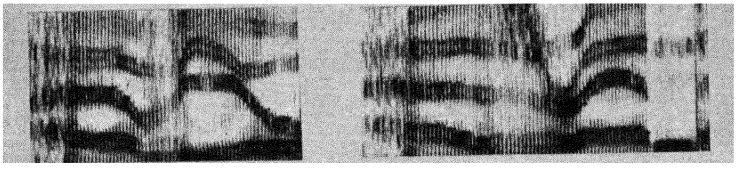
card ***old***

In glide-voiced stop combinations as illustrated by *card* and *old*, the voiced stop influence on the glide is similar to the cognate voiceless stop influence described under *part* and *molt*. The continuous voicing of the glide and stop emphasizes the stop influence and is portrayed in part by the voiced stop gap on the patterns. The nature of the transition, the hubs, and the characteristic pattern of each element in the combination serve as clues in identifying the sounds.



sweet *free* *sleep*

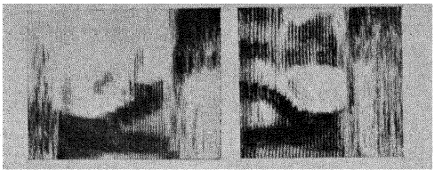
Combinations of voiceless fricatives and glides such as *sweet*, *free*, and *sleep* are portrayed by patterns produced primarily by frictional modulation united with patterns resulting from vocal cord and varying cavity modulation. Since the fills for the fricative sounds are longer in duration than the spike fills in the voiceless stop sounds, their influence on the origin of the glides frequently is observed more readily than that of the stops. However, both stops and fricatives follow the same general principles of transition when they are combined with glides. Voiceless fricative-glide combinations may be identified by the patterns and hubs of the combined sounds, and by their influences.



have you *has read*

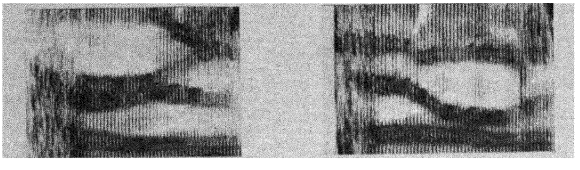
Voiced fricative-glide combinations occur quite often in continuous speech between such words as *have you* and *has read*. Continued voicing is evident in the connected resonance bars during the transitional change from frictional-vocal cord-cavity modulation to vocal cord-cavity modulation. The curves in the bars portray the reciprocal influence of the combined sounds. In addition, the hubs and positions of the bars at the points where the curves change, and especially the voiced fills of the fricatives, are clues by which these combinations may be identified. Variations in stress and in the nature of articulation modify both the appearance of the individual patterns and of the transitions as well as the relative durations of the separate elements in the combination. Stress is of particular importance in a consideration of combinations occurring between syllables in connected speech.

When a voiceless fricative follows a glide, as illustrated in *horse* and *else*, the point of termination of the voiced phase of the glide is influenced by the fricative sound. Voicing usually continues until frictional modulation begins, and the articulatory



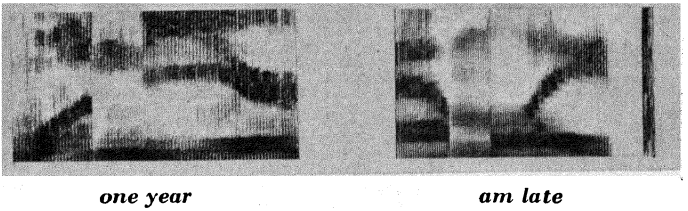
horse *else*

movements may be completed during the fricative sound. It is apparent that the relative proportion of glide elements and fricative elements may vary from time to time for a given speaker, and that the points of termination of the glide and initiation of the fricative may overlap. Occasionally stop-start modulation is introduced during the transition between the glide and fricative, resulting in a brief stop gap followed by a stop-fricative blend. As in other combinations, the appearance of the patterns of the elided sounds, the positions of their hubs, and the nature of the transition, provide the principal clues for identification.



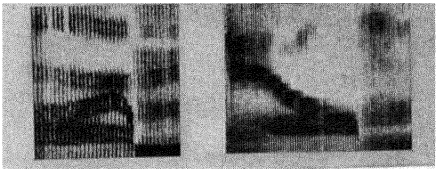
hers *tells*

Glide-voiced fricative combinations portray the mutual influences of the combined sounds in the continuous curves of the resonance bars. The r–z combination in *hers* and the l–z in *tells*, may be identified by the position of the resonance bars in the short steady state phase of the sounds and by the in-between transitional pattern. Beginning of frictional modulation is apparent in the voiced fill of the fricatives. It may be observed that the frictional component frequently is present in a considerable portion of the transitional pattern and that the resonances superimposed across the fill continue to curve until the end of the pattern.



one year *am late*

Nasal-glide elisions occur most frequently between words or syllables in connected speech. Like all combined patterns, the appearance of the individual sounds and their transitions is related to stress, rate, and precision of articulation. A distinctive difference in the appearance of the vowel-like nasal and glide resonance bars is present only when nasals are nasalized and glides are free from nasality. When control of the soft palate is rapid and well coordinated, a sharp transition between the nasal and glide patterns is apparent, and, consequently, the mutual influence of the sounds as in the patterns *one year* and *am late* may be de-emphasized. The same combinations made with slower articulation, or with less precise action of the soft palate, will portray the mutual influence of the combined sounds and the continuation of nasal resonation into the pattern of the glide. In normal speech, the transitional pattern between combined vowel-like sounds becomes more prominent as the distance between their hubs increases.



arm ***elm***

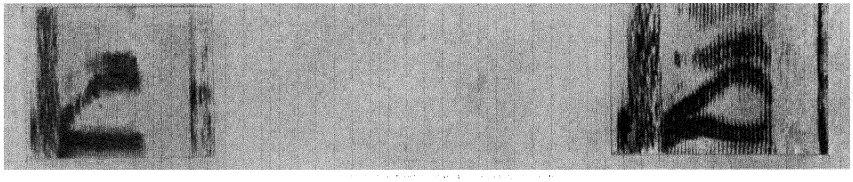
Glide-nasal combinations occur both within words and between syllables and words. Like the nasal-glide elisions, the appearance of the patterns and the transition is contingent upon the manner of speaking. However, since the glide-nasal combinations commonly are used as terminal units for syllables, they are not modified significantly by stress, and their patterns usually portray transitional influences quite clearly. When the nasality is confined to the nasal sounds, as in the patterns of *arm* and *elm*, a distinguishable difference in the appearance of the glide and nasal resonance bars is apparent. This difference in appearance is due to the types of modulation characteristic of the individual sounds. The glides normally are produced by vocal cord-cavity modulation, while the nasals are made with vocal cord-closed mouth-open nasal cavity modulation. Frequently a short gap is present between the end of the glide and the beginning of the nasal sound. This is believed to be associated with the stopping of the emission of the sound from the mouth, the opening of the nasal port, and the lag in building up the nasal resonances. Some speakers interpose a form of the schwa as part of the transition between these sounds, with an effect similar to that of the schwa-in-transition between the voiced stop and glide.

Representative combinations of the glides with glides, nasals, fricatives, and stops have been illustrated and briefly described in this Unit. Although numerous combinations other than those presented, occur between syllables or words in connected speech, it is hoped that the basic general principles discussed here may be applied and extended to the identification and analysis of any specific combination. Such variations in speech as the rate and precision of articulation, rhythm, and stress, serve to modify the visible as well as the aural patterns, affecting chiefly the relative durations and the transitions. Alterations in the nature of the translators also may change certain aspects of the visible display, but the fundamental phonetic observations have been based on relationships between the appearance of individual and combined sound elements, the different types of modulation used in their production, and the physiological functions basic to continuous speech. Therefore, they are not believed to be dependent upon arbitrary assumptions regarding specific, definite pattern forms, but may be considered as general phonetic principles regardless of individual variations.

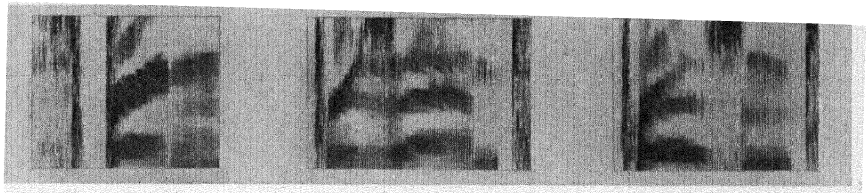
Review words

will you, page 251
have you, page 257
pretty, page 252
cross, page 252
part, page 256
arm, page 259
black, page 254
medal, page 255
tells, page 258
am late, page 258
sweet, page 257
but you, page 255
dress, page 253
are we, page 251
one year, page 258
tree, page 252
green, page 253
card, page 256
has read, page 257
clear, page 253
molt, page 256
sleep, page 257
order, page 255
hers, page 258
please, page 253
girl, page 252
world, page 252
between, page 255
bring, page 253
letter, page 254
horse, page 258
free, page 257
glad, page 254
old, page 256
elm, page 259
metal, page 254
else, page 258
all right, page 251

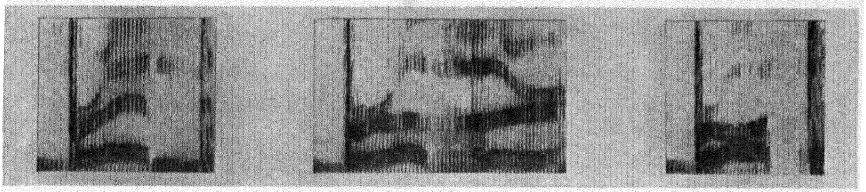
New words to learn



quick *quiet*



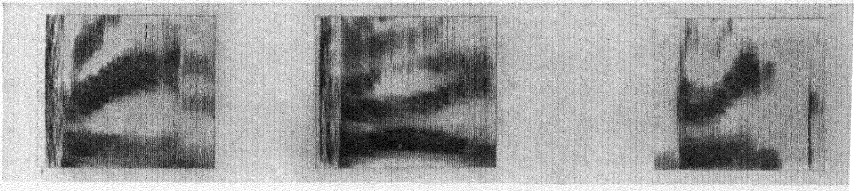
spring *present* *present*



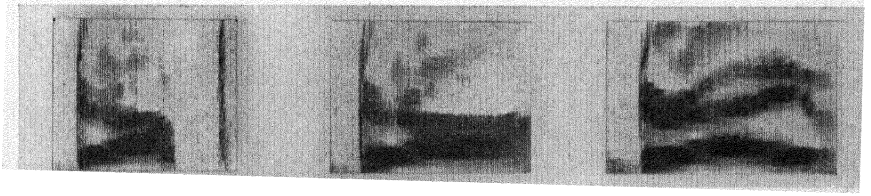
bread *brother* *brought*



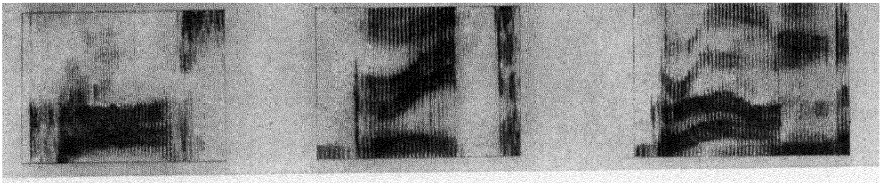
street *strong* *true*



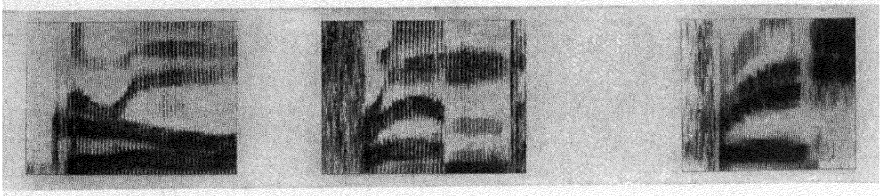
train *try* *drink*



drop *draw* *drive*



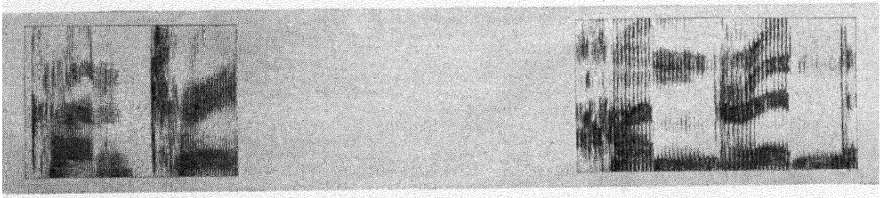
cross *great* *ground*



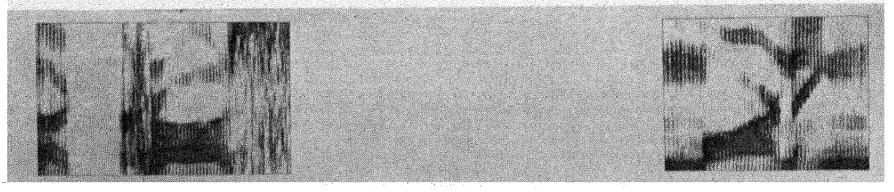
grow *friend* *fresh*



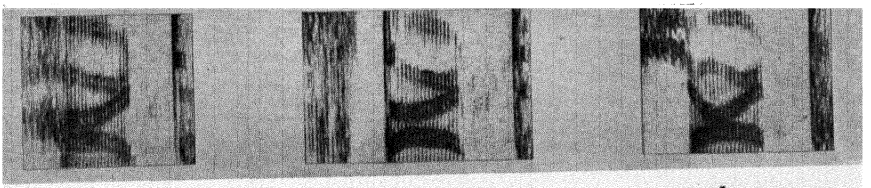
from *front* *through*



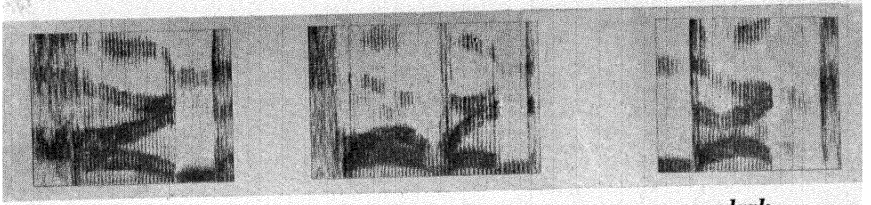
country *hundred*



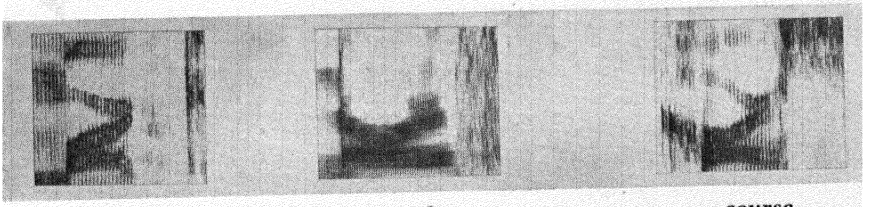
across *morning*



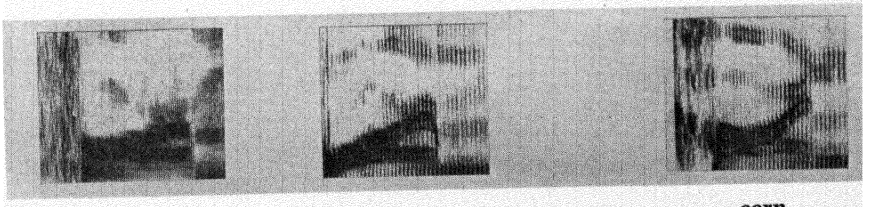
heart *start* *short*



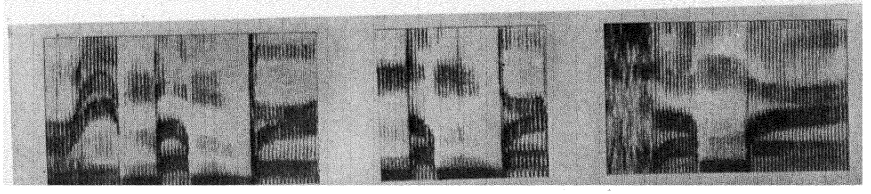
hard *forward* *dark*



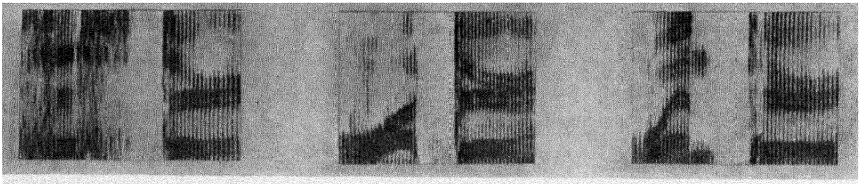
mark *north* *course*



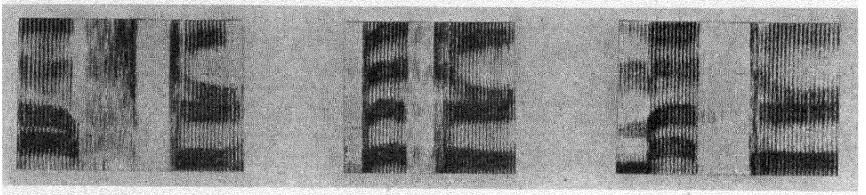
form *warm* *corn*



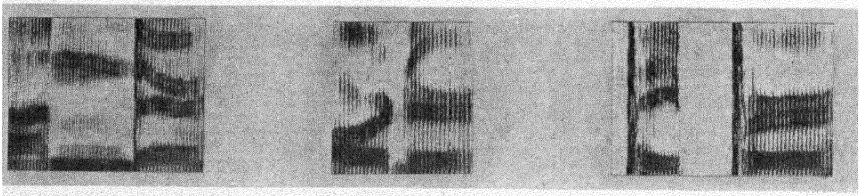
remember *number* *summer*



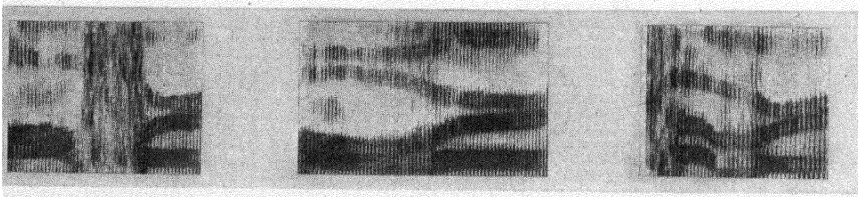
sister *water* *winter*



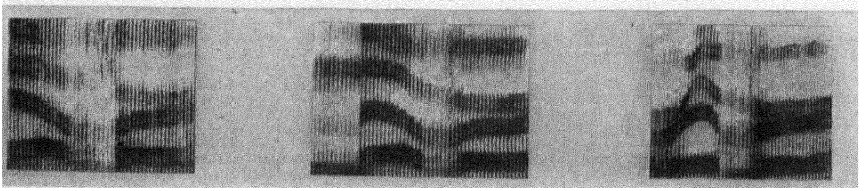
after *better* *matter*



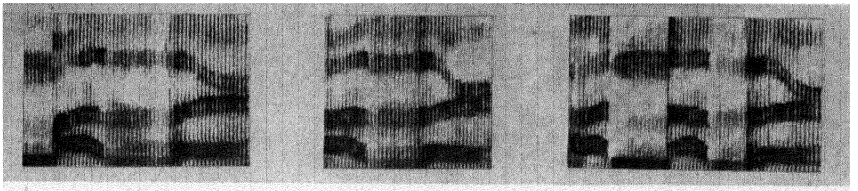
under *order* *paper*



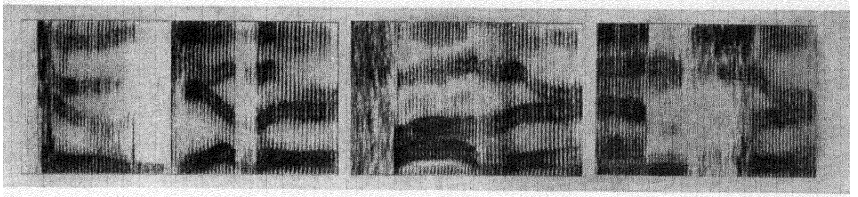
offer *over* *cover*



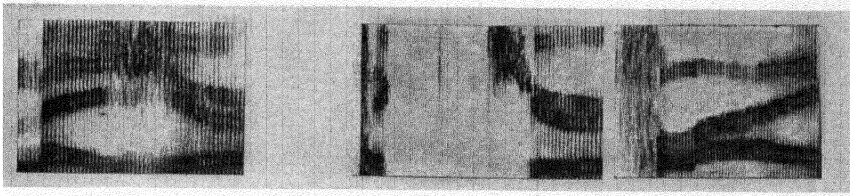
ever *never* *river*



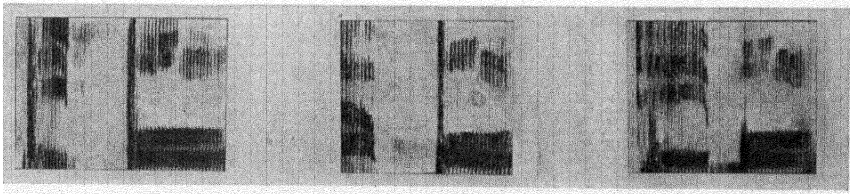
mother *other* *another*



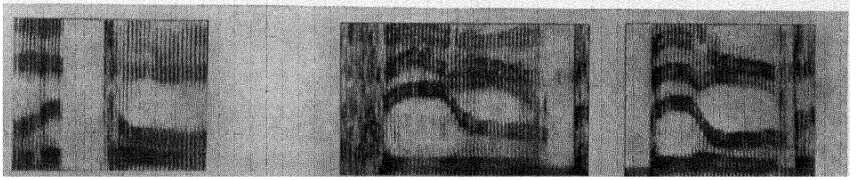
together *father* *answer*



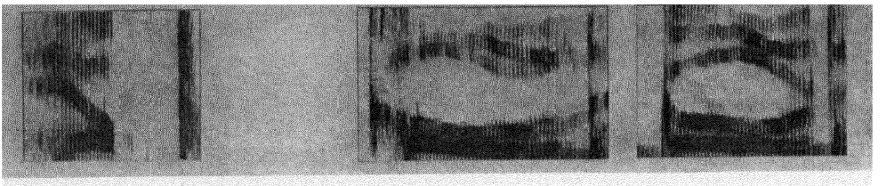
measure *picture* *fly*



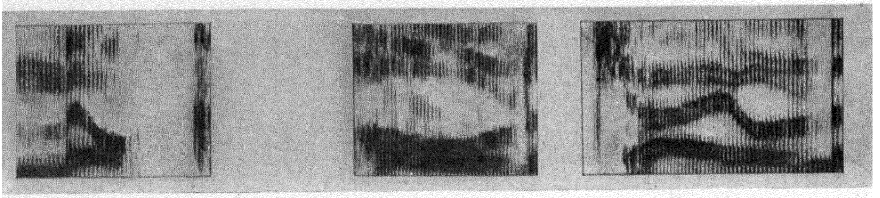
people *apple* *table*



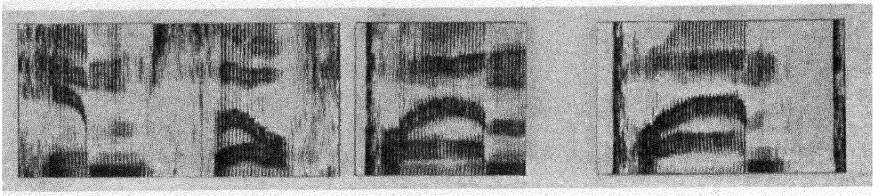
little *field* *build*



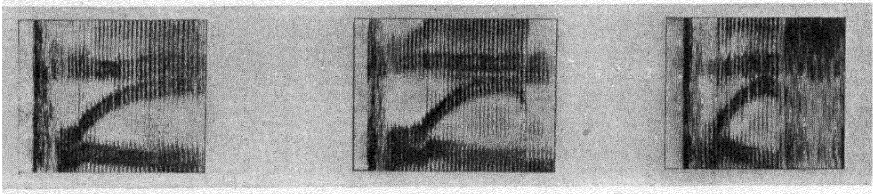
help *cold* *gold*



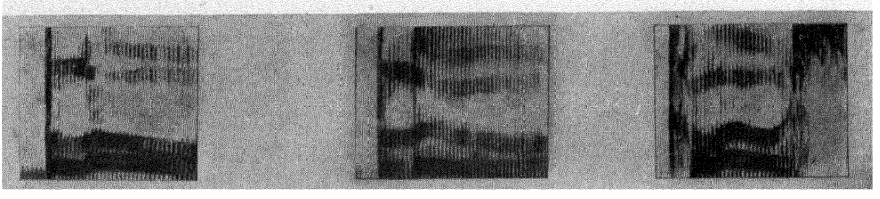
milk *hold* *child*



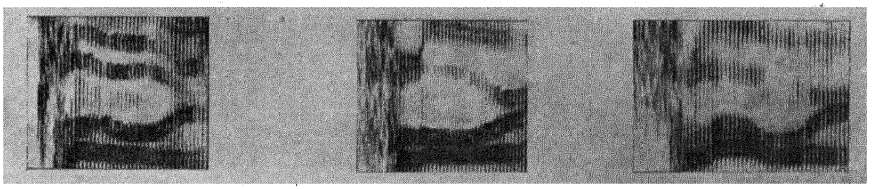
himself *plan* *plant*



play *plain, plane* *place*

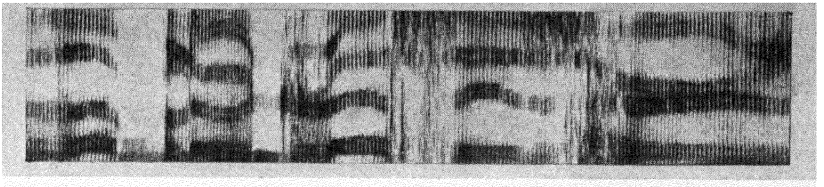


blow *blue* *close*

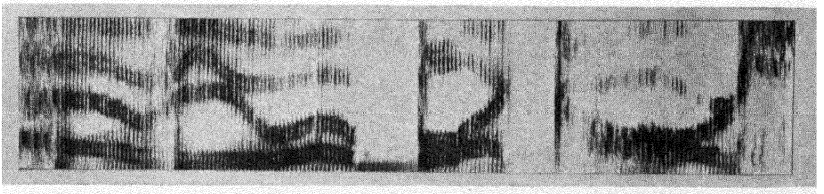


close *floor* *flower*

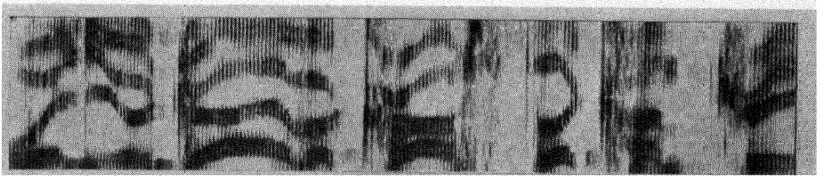
Drill phrases and sentences



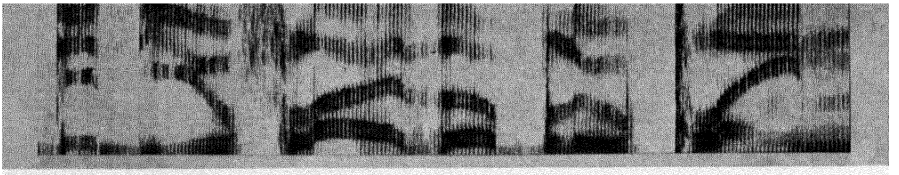
The blue dress is hers



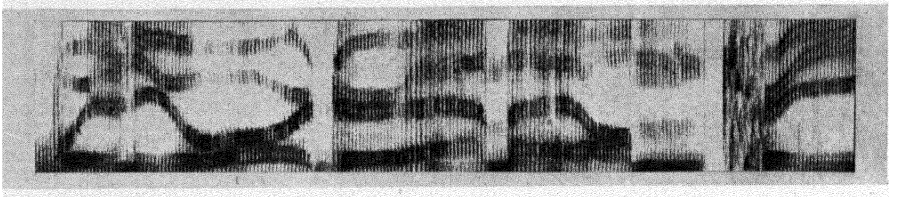
Have you a black horse?



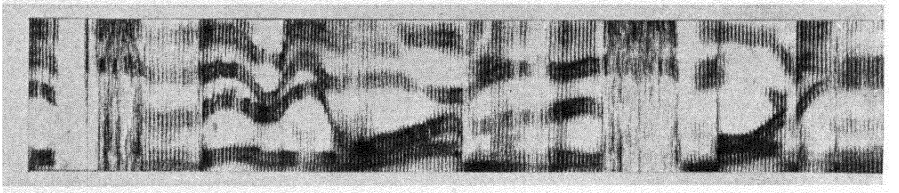
Will you drive across the country?



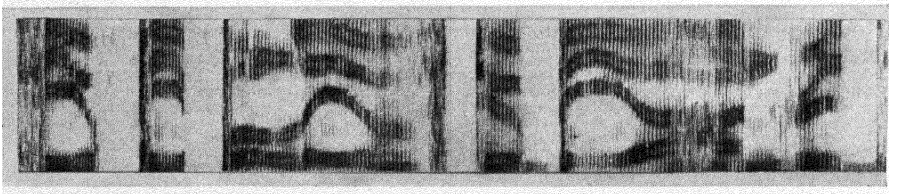
Did you fly the blue plane?



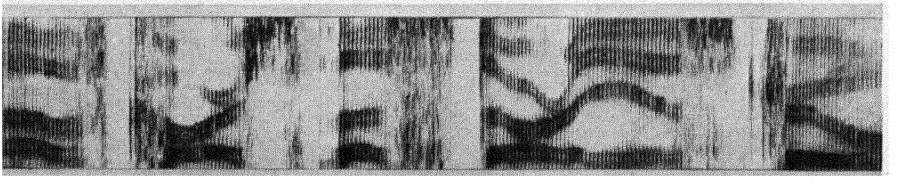
Will you order an elm tree?



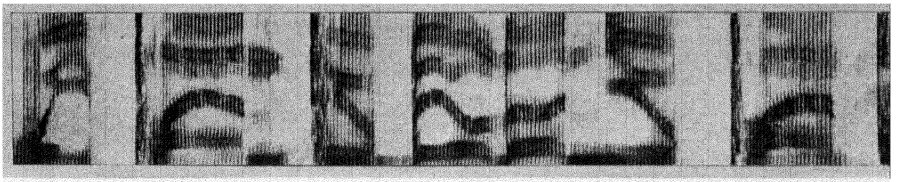
It's very warm this morning



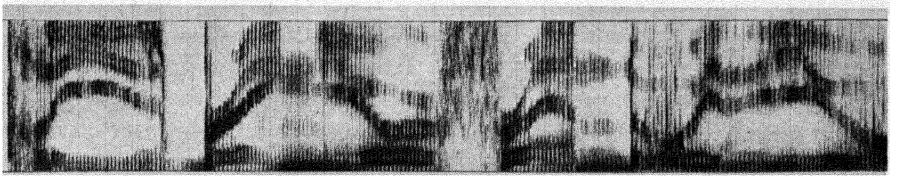
Few people live to be a hundred



Of course, the story is true



We plan to build a new plant



Please bring your friend with you

Chapter 12
Reading the Vowels

In Chapter 7, a general visual reference pattern for each of the vowels and diphthongs was presented and discussed briefly. Since the patterns were intended to serve as references for the discussions of the hubs and influences in the chapters that followed, they portrayed the vowel sounds spoken by themselves, free from influence by associated sounds. Isolated vowels, however, do not occur frequently in connected speech; and therefore the vowel must be considered as it influences and is influenced by other sounds. Chapters 8, 9, 10, and 11 are concerned primarily with discussions of the stop, fricative, and vowel-like sounds as they are related to the vowels and to each other. Chapter 12 is organized to present an illustrated review of the vowels as they are related to the non-vowel sounds discussed in the lesson chapters.

Each of the front, back, and central vowels comprises a separate unit in the early part of this chapter, consisting of: a profile illustrating the usual articulatory positions for the vowel, spectrograms of the vowel in a male and a female voice, and a series of diagrams summarizing the relative resonance positions of the vowel and the hub areas of the non-vowel sounds with which it may be combined. Hub areas of the non-vowels are diagrammed as they occur in the initial position. As discussed in the lesson units, the initial and terminal hubs of these non-vowel sounds are not significantly different with the exception of the variable sounds, **k–g**, **l**, and **r**, and the **ŋ**, which does not occur in the initial position in English. The reader is referred to the respective lesson units for information concerning the terminal hubs of these sounds. Concluding the vowel unit series is a brief unit presenting a profile and spectrograms in a male and female voice for each of the six combination vowels commonly classified as diphthongs. Since each element in the diphthongs is treated separately in the vowel unit, it did not seem necessary to repeat the diagrams of resonance position in relation to non-vowel hub areas.

Also, we have discussed the unstressed or schwa forms of the vowels in some detail in the lesson chapters, and therefore they are not included in the illustrations in this chapter. It will be recalled that any vowel, when unstressed, may take a form intermediate between its stressed position and the position for the stressed central, neutral vowel **ʌ** (*up*). The particular form taken is dependent upon many variables, among which are the nature of the combined sounds, the degree of unstressing, and the rate and precision of articulation. In any event, the unstressed form usually is shorter. lighter, and less well defined than the stressed form.

Following the diphthong unit is an illustration, Unit 5, giving a pictorial review of the resonance positions for each of the vowel groups, front, back, and central, in rela-

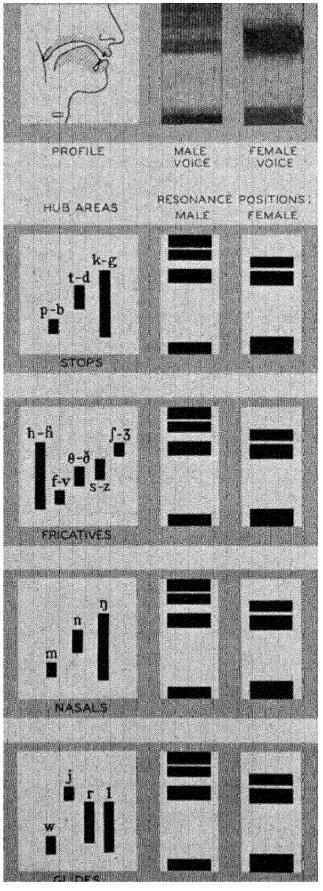
tion to the hub areas of the non-vowel sounds. Several different factors are summarized in this illustration. First, the vowel resonances are shown in series for a male voice and for a female voice so that the progressive arrangements of the resonance positions may be compared within each series and between the male and female voices. Second, the hub areas of the non-vowel sounds may be compared, and their relation to the resonance positions of the vowels in both voices may be observed quite readily.*

We have stressed, in the discussion in the lesson chapters, the concept that the primary hub does not represent a single, fixed point but rather a range of points within which the major zone of influence for a sound may be found. To extend this idea further, it should be mentioned that the resonance positions for any particular vowel also vary within certain limits in the speech of one individual, and in the speech of different individuals. Hub areas and resonance positions which are discussed in the lesson chapters, and which are illustrated in the vowel unit series and in the pictorial review of the vowels, are approximations resulting primarily from the analyses and observations of the 3,500-cycle broad-band spectrograms used in connection with the visible speech training program. In order to clarify somewhat the variations in the resonance positions of the vowels in the speech of different individuals, we have drawn composite diagrams in Unit 6, showing the hub areas and the vowel resonances for nine male and for eight female speakers. Each diagram in this unit represents the averages for the male and the female voices by solid blocks. The ranges of variation are shown by narrow vertical lines attached to the solid blocks. It is evident that, with this rather small sample, there is some overlapping of the limits of the resonance areas, emphasizing the basic concept of the relative nature of the resonance positions within an individual's speech. Perhaps it may be compared to a melodic pattern for each individual speaker in which the separate notes maintain an approximately constant relation to each other. Melodic patterns of different speakers will show similar relationships among the separate notes, although the absolute distances between notes may vary from speaker to speaker. While the available spectrographic evidence has been considered in selecting the frequency range of these resonance areas, there is little question but that future studies and modifications in visible speech translators will define more closely the averages, the limits, and the standard deviations of the resonance and hub ranges.

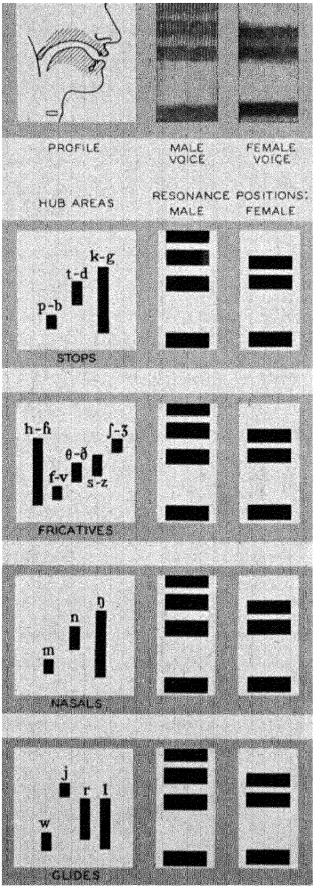
The illustrations which follow are intended to constitute a concise summation, comparison, and review of the information concerning the hubs and influences presented in the preceding chapters.

* A similar summary unit providing an abbreviated pictorial review of the relationships between the profile diagrams and hub areas of the non-vowel sounds is presented in Appendix II.

Unit 1 — Reading the Front Vowels

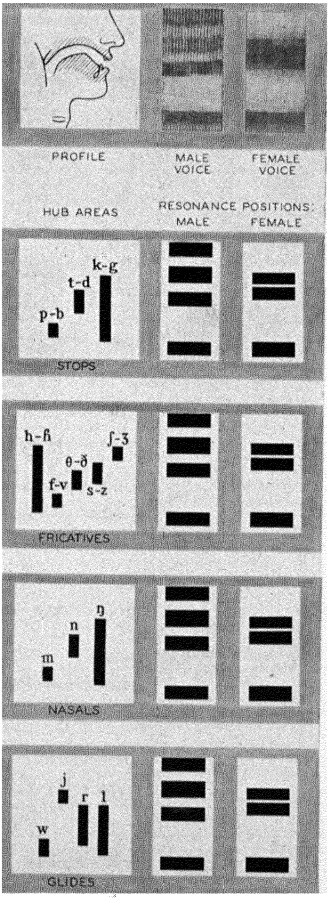


i (*eve*)

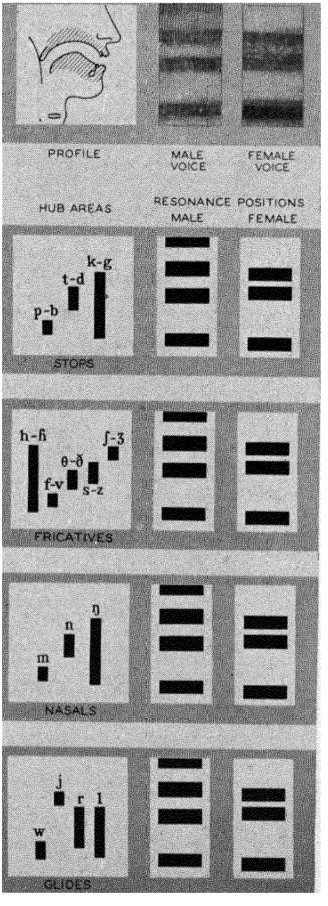


ɪ (*it*)

Unit 1 — Reading the Front Vowels

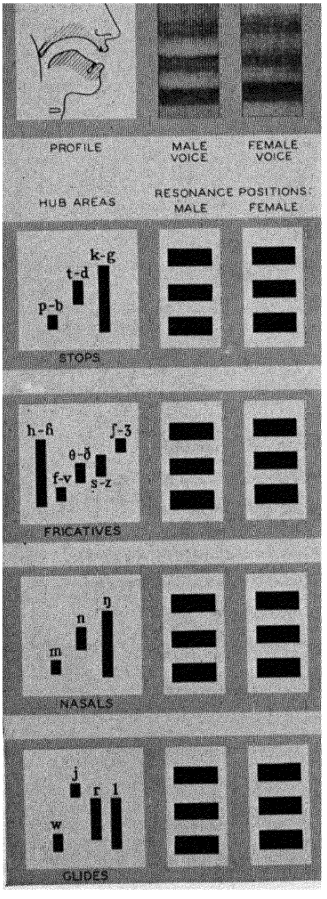


e (*hate*)

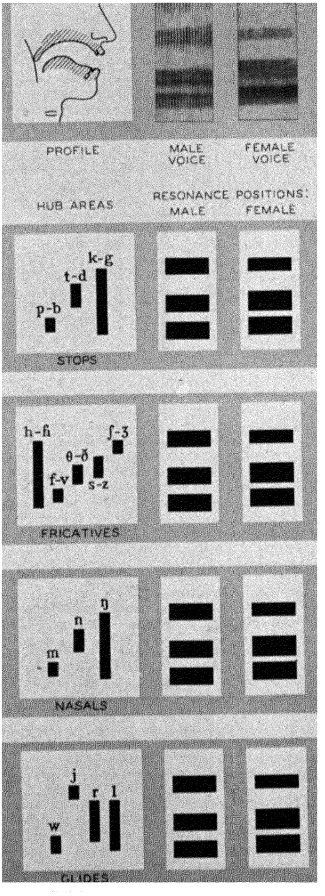


ɛ (*met*)

Unit 1 — Reading the Front Vowels

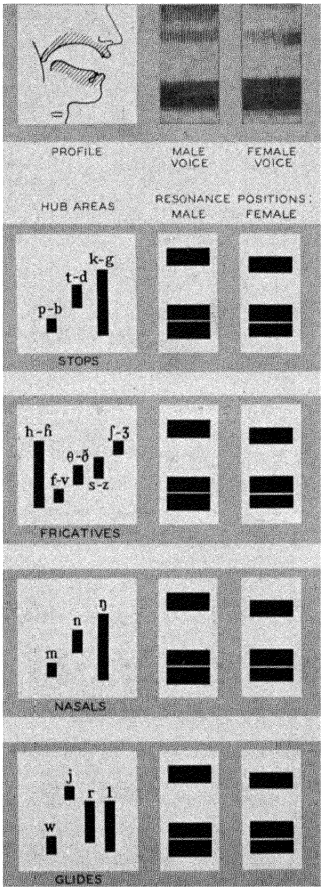


æ (*at*)

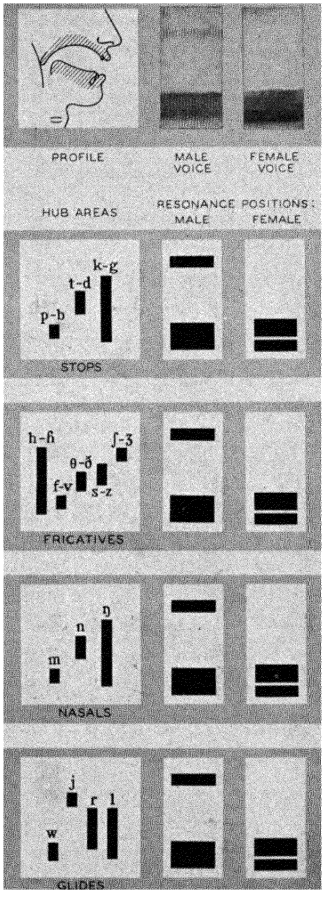


ɑ (*ask*)

Unit 2 — Reading the Back Vowels

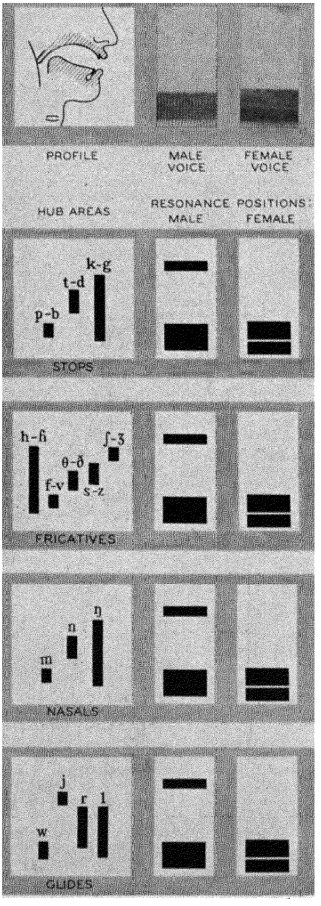


ɑ (*father*)

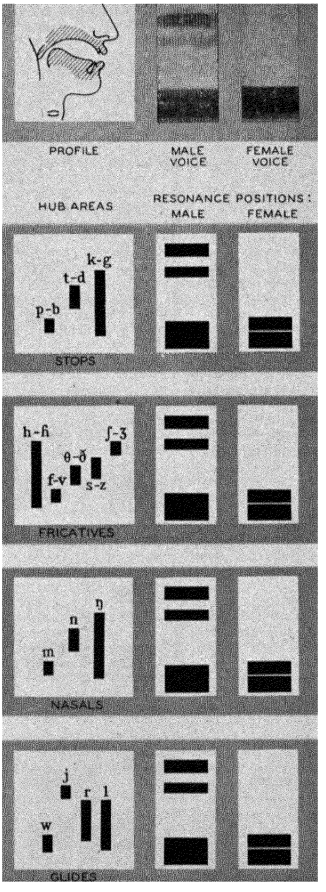


ɒ (*not*)

Unit 2 — Reading the Back Vowels

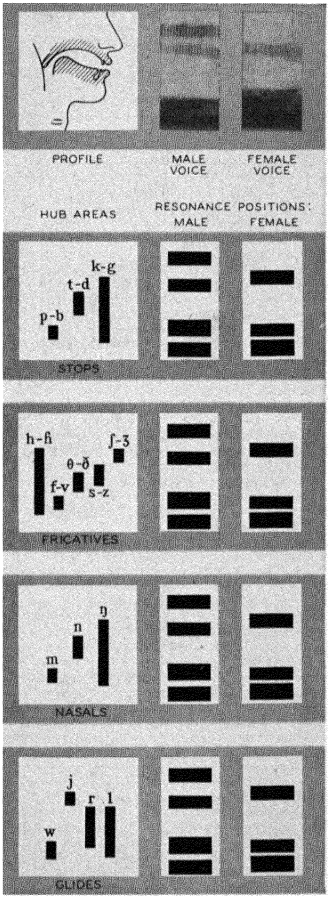


ɔ (*all*)

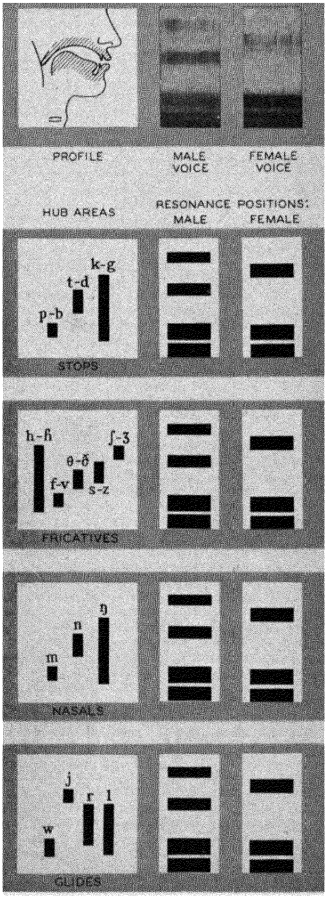


o (*obey*)

Unit 2 — Reading the Back Vowels

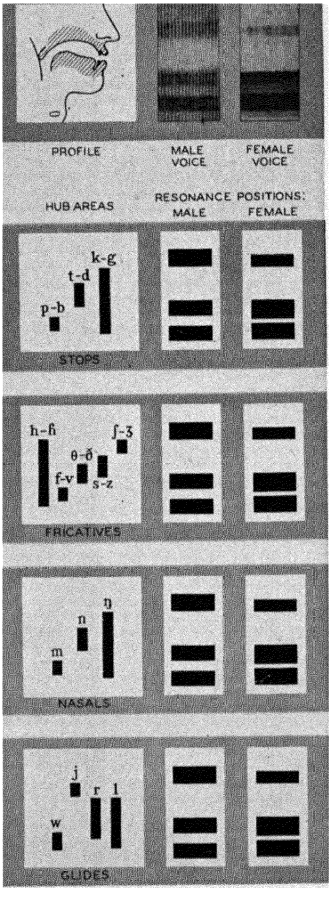


ʊ (*foot*)



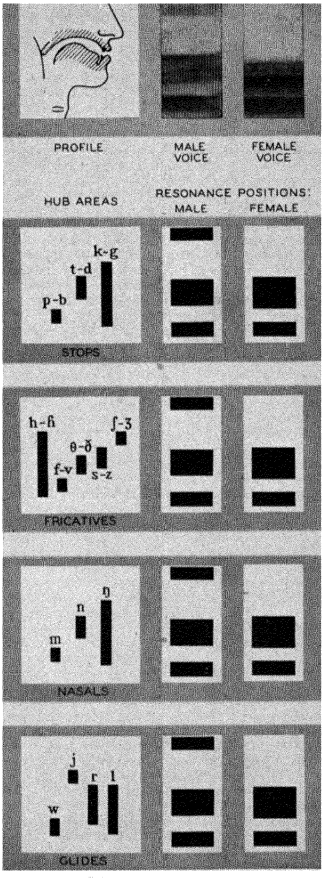
u (*boot*)

Unit 3 — Reading the Central Vowels

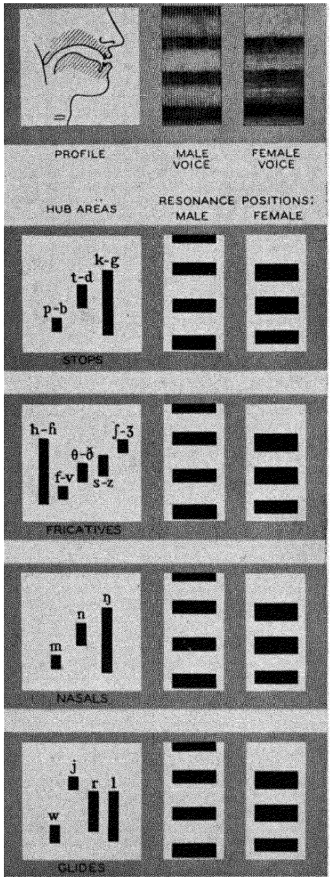


ʌ (*up*)

Unit 3 — Reading the Central Vowels

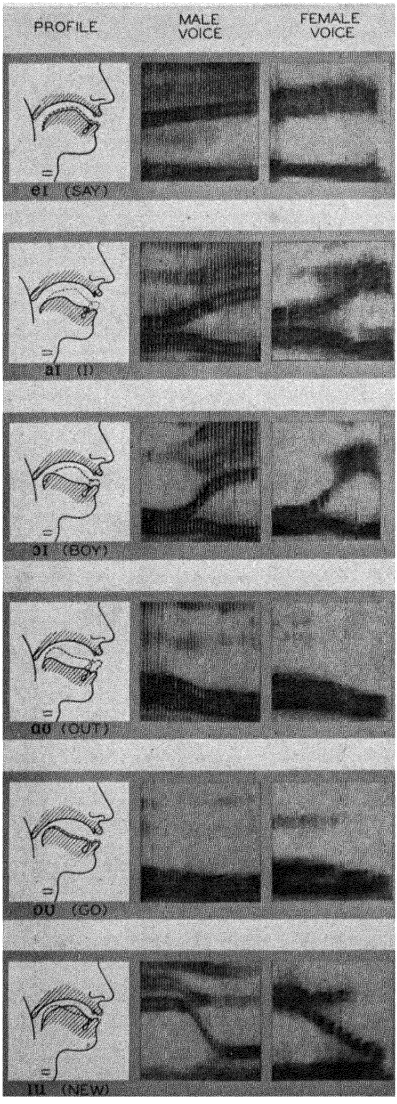


ɝ (*bird*) *General American*



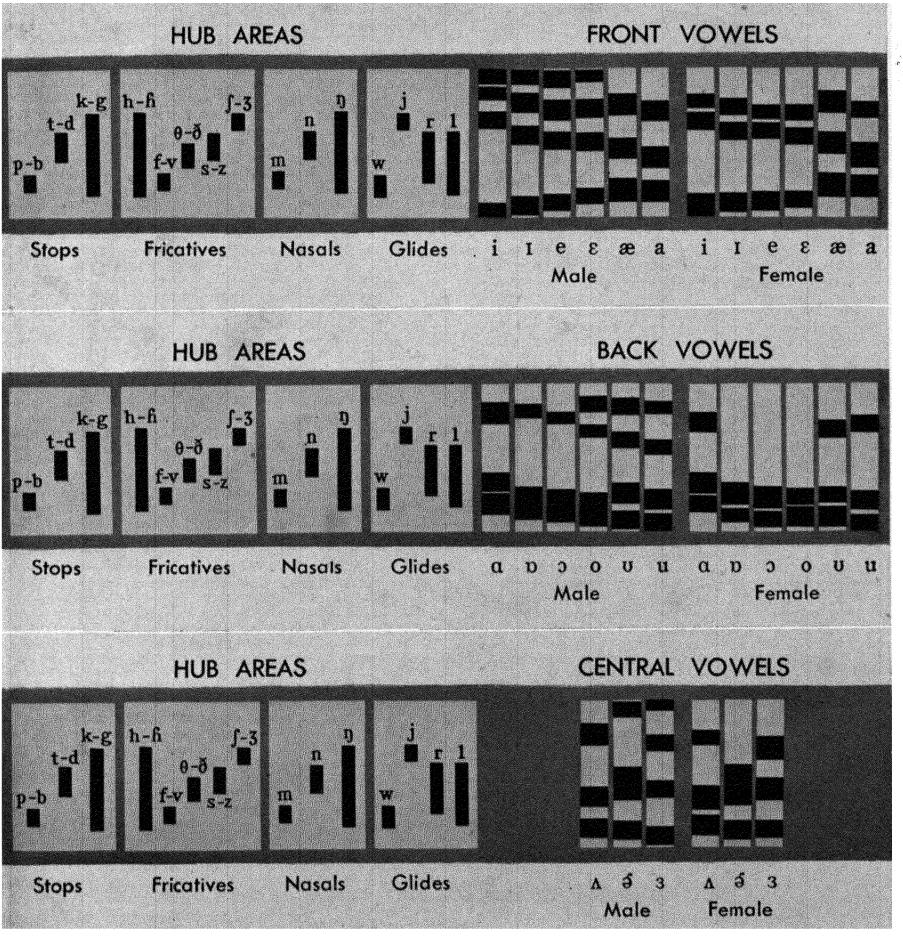
ɜ (*bird*) *Eastern*

Unit 4 — Reading the Diphthongs

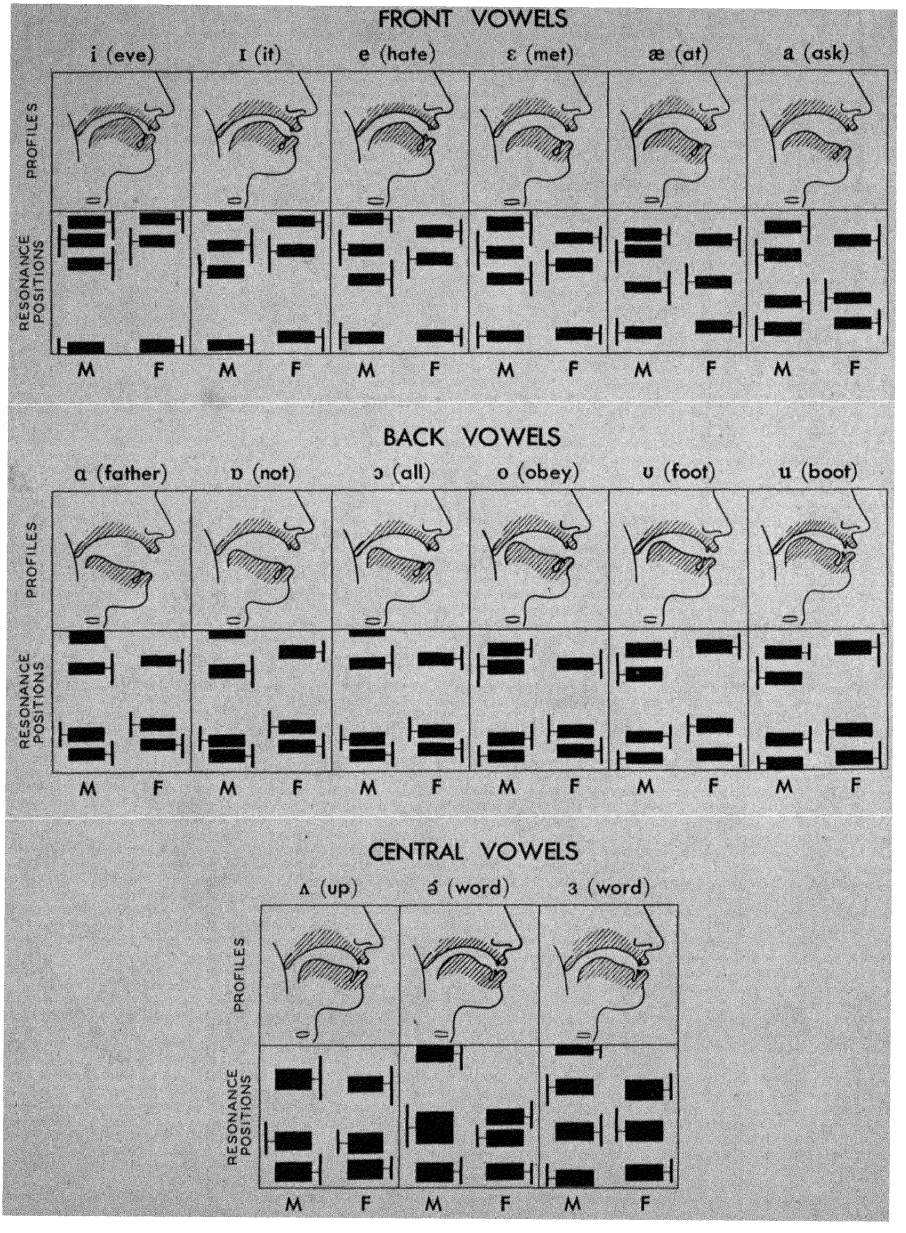


Diphthongs

Unit 5 — Summary Diagrams of the Vowels with the Hub Areas of the Stops, Fricatives, Nasals, and Glides



Unit 6 — Profiles of the Vowels with Diagrams of Resonance Positions in a Composite Male and a Composite Female Voice



There are innumerable combinations of sounds and variations within sound patterns occurring in normal American speech which we have not attempted to discuss or illustrate in this text. The presentation of the basic phonetic principles of visible speech, patterns of the basic sound elements, and a large number of frequently used sound combinations, constitute a broad frame of reference upon which the reader may focus his particular interests and field of experience. Together, Parts One and Two contain the available technical and phonetic information concerning visible speech in a form designed to enable the reader to understand and to evaluate, from his own point of view, the present status of development of visible speech instrumentation and concepts.

PART THREE

APPLICATIONS OF VISIBLE SPEECH

Chapter 13
Deaf Interests

As explained earlier, the experimental training carried on in the Bell Telephone Laboratories was only for the purpose of providing assurance that the patterns of visible speech contained intelligence or meaning in a form that the eye could perceive and interpret. This would be a basic requirement for any method of speech portrayal to fulfill the needs of deaf education, but it would not be the complete answer. Experimental training in pattern reading, and investigation of the phonetic principles involved, indicated that visible speech could be used in speech and general education for the deaf, but did not indicate *how* it could be used most effectively. Nor was this the kind of study that belonged in a telephone communications laboratory. Evaluation and application studies are obviously the responsibility of those directly concerned with the problems of deaf education. At the same time there are several considerations that would affect any plan for carrying on such studies.

Principles, not machines

Important results of the visible speech development at this stage are in the production of principles rather than machines. The possibility of turning speech into patterns that one may read has long been of interest to deaf educators, although in the past the prospects of accomplishing such a translation have not been encouraging. Instruments such as the present oscilloscope and earlier devices for showing the complex wave shapes of speech seemed promising initially but proved on further study to be very limited in their usefulness. There was even some feeling among those acquainted with the fundamentals of the problem that it might be impossible, for, after all, the eye and the ear are very different organs.

The visible speech development shows that translation from speech to readable pictures is possible, and indicates principles that are essential. Up to the present, the emphasis has been upon establishing these principles and comparatively little attention has been given to the production of simpler machines for performing such translations. The translator employed in the initial experimental training work was a large, complicated, and exceedingly expensive device. Although one small unit of a different type was constructed in order to obtain some preliminary idea of what would be involved in a "boiling down" process, the patterns of this compact unit were appreciably inferior to those obtained with the large unit. In an ultimate translator for school or other use, both compact construction and improved patterns would be essential.

The present situation is analogous in a way to the initial accomplishment of airplane flight. When the first plane left the ground for a brief journey through the air, it did

not mean that these machines could immediately be made available for general use, but it did mean that man *could* fly and it established certain principles that were eventually built into machines that, after patient development, made possible aviation as it exists today. Similarly, we now know speech can be made visible in a meaningful form because patterns have been produced that people have learned to read. We also know why these patterns are readable while those produced in the past were not. In other words, principles for the translation of speech into meaningful patterns are established, but a considerable amount of work remains before the deaf will benefit by this result even though it does constitute a major advance.

Before the airplane became a practical machine, a great deal of development and experiment was necessary; and before people generally could use planes, airfields had to be constructed and teachers had to learn to fly in order to teach others. The situation is somewhat the same in the case of visible speech. A great deal of electronic development will be necessary before visible speech translators can be generally available to the deaf. Paralleling any apparatus development there must be experimental instruction, and such instruction should for the present be in the nature of research on *how to use principles rather than how to use machines.*

If the present experimental models of visible speech translators were turned over to a school for the deaf to be "tried out" in the midst of established educational procedures, the results might be as unsatisfactory as if the Wright brothers had submitted an early model of their plane to a transportation company for an appraisal of its possibilities. This is no reflection upon the ability of teachers of the deaf to know what is useful and what is not. It simply means that present speech patterns are still experimental, and that any study of visible speech applications ought to consider not only what may be done with available forms of display, but how these displays might be improved as indicated by training results.

A carefully organized and conscientiously undertaken research plan seems to be a most desirable first step in the visible speech program from this point forward. That the present forms of display or some related type ultimately may be used by the deaf seems clear, but the manner of use remains to be determined, and the time when benefits will begin to be felt depends to a large extent upon the way the exploratory phase of evaluation is handled.

Limited equipment

A point to bear in mind is that the equipment at our disposal for research studies is limited. There is only sufficient apparatus available at the present writing to equip one research project in one school or university. More might be built, but such equipment is very expensive to reproduce, and effort applied in that direction would have to be taken from technical studies concerned with the production of better and less expensive equipment.

Equipment limitations make it doubly important to plan the visible speech research program carefully in order to be certain that what is available will be used to full capacity. As the result of discussions with several individuals who are acquainted with the problems of deaf education and with some of the possibilities and present limitations of visible speech, certain ideas concerning such a plan have accumulated. These are outlined below. No doubt they are deficient in some respects and perhaps wrong in others, but they have served as a first attempt to state certain questions that need to be answered before a well organized evaluation and application study of visible speech can be formulated.

The Problem

Broadly, what we would all like to know is how visible speech may be used in the education of the deaf. This may better be stated in two specific questions:

(*a*) *To what extent can visible speech facilitate speech education and speech improvement?*

(*b*) *To what extent can visible speech be used to advantage as a means of communication in the general education of the deaf?*

Any program organized to answer these questions must clearly avoid interruption of conventional training. It would, for example, be a serious mistake to center the teaching of deaf children around the visible speech translator, and to the disadvantage of lip reading progress, since lip reading is and probably will continue to be a major factor in their daily communication. In addition, it would be most undesirable to make the pupils dependent upon equipment that might not be available to them for some time after their participation in the program had been completed.

The students who take part in the experimental work with visible speech, as well as their parents, should understand clearly that the project is experimental and that there is no assurance visible speech translators will be available for individual use for some time to come. They would, however, be participating in a program that should be of considerable importance to the deaf as a group.

In general, there are four groups to be considered in a research program of the kind under discussion. These are the teachers, the pupils in nursery school and the intermediate grades, and the adults.

The teachers

It is naturally assumed that studies of the application of visible speech to deaf education would involve an established school, although those primarily responsible for the research might conceivably be affiliated with a university. Even though the research project should not interfere with the routine training procedures in this school, there would have to be cooperation on the part of the classroom teachers.

Classroom and special teachers should, in fact, become well acquainted with visible speech and take an active part in the studies of its use.

Initially, therefore, there ought to be an interval devoted simply to preparing to carry on the research, a period in which teachers would have an opportunity to learn at least the fundamentals of visible speech. From this point on, it would be essential that these teachers work with the researchers and be prepared to take over the major responsibility for carrying out specific projects, with the expectation that ultimately the methods they are helping to formulate will become a part of their regular teaching procedures.

Nursery group

The nursery and pre-school group will be of particular interest in relation to these studies. Adults and older children will have had a background of training and an accumulation of habits certain to affect the experimental results, but these pre-school children will be comparatively free from the influence of linguistic experience and of related acquired habits. They would present a unique problem in training and one that is of great importance, because the possibilities of utilizing visible speech at an early stage would affect the whole plan of subsequent training. If visible speech could be utilized early it would be of much more value than if its use, for some reason or other, had to be restricted to later age groups.

But how could a lively youngster with no conception of concentration be expected to look at a visible speech display screen? Is it not unreasonable to expect to teach any tiny tot that this pattern means one thing and that another? It would certainly be useless to apply to them the training procedures effective in adult training. To attempt to teach a very small child the principles of visible speech through lesson chapters of the kind presented in Part Two of this book obviously would be an absurd procedure.

But there are certainly sensible possibilities of procedure for teaching small children visible speech, and one major purpose of the research program would be to try these and, while trying them, to consider ways in which the results might be improved by changes in both method and equipment. The assumption is that requirements for both methods and equipment will be different for different age groups, and a better understanding of these requirements will be one purpose of the educational experiment.

An important question to be answered is, to what extent may vision be substituted for hearing in the speech and language training of the deaf? For instance, can three- and four-year-old deaf children learn to associate visible patterns with sounds in a way similar to the auditory associations made by normal hearing children? Can whistles, horns, footsteps, coughs, sneezes, and all the numerous incidental sounds and noises of our daily lives acquire more complete and natural meaning for the little deaf child

through their visible patterns? Will he be able to combine visual perception with imitation and feelings of pleasure and satisfaction just as the hearing child does in imitating the sounds he hears? If reflex babbling can be stimulated, and enjoyed by the deaf baby, will his voice quality be more natural and will he be more susceptible to speech activities? How best can the deaf child use visual association to acquire a meaningful vocabulary and a natural easy flow of language?

Finding the answers to such questions is a challenge to the ingenuity and skill of teachers and parents of deaf children. Certainly the prevailing practice of having the deaf associate visual, tactile, kinaesthetic, and (when possible) auditory impressions of speech should continue to be used.

Visible speech is a means of extending the present use of vision to include sound patterns that vary with the vocal processes used in their production, thus giving the tactile and kinaesthetic senses a tangible reference pattern. Just how these and other possibilities may best be applied remains to be seen. No doubt various procedures will need to be tried, and it is expected that results will be influenced by added teaching experience and better types of translation. Further experience may indicate that instruction would be facilitated by a larger or smaller screen, by slower or faster screen speeds, by different colors in the display, by different pattern shapes to accentuate certain sounds, and by different ways of showing pitch.

Specific questions that need to be answered in connection with how best to use visible speech in nursery school, and also in the upper grades and with adults, are numerous. Many will occur to those familiar with the problems of deaf instruction. It would be premature to attempt discussion of them here, but by way of illustration one important problem may be mentioned. Should the training be largely individual or group? If group, should the visible speech translator be a single large screen unit with the group scattered in front of it in some way? In this case how should the microphone pickup be handled? Would it be better to use a single large screen and place the pupils in chairs in front of this screen with individual microphones connected together so that all sounds would appear on the screen? Or would it be advantageous to use individual small screen translators in front of each chair and arranged in a small circle with all microphones connected to all translators?

There are many other questions of this sort that have a bearing not only upon educational methods, but also upon the design requirements for ultimate equipment.

Experimental training must not, as explained earlier, interfere seriously with established training procedures. This means that initial exposure to visible speech might have to be much more limited than would be desirable for most effective results, but such restrictions can probably be taken into account in evaluating progress. Probably experimental class work involving visible speech might have to be separated somewhat from the regular class work, for example, in a special room. However, it would

seem very desirable that the research program be integrated with the established educational program, and that classroom teachers assume considerable responsibility in connection with the experimental studies.

Continuation of nursery group

Results of the experimental training research that will apply most directly to the ultimate evaluation of visible speech will be those provided by experience with groups that have learned visible speech at an early age. It will be especially important, therefore, to follow through with the research program at least several years beyond the initial nursery school training. The need for following such groups through several subsequent years is also important in relation to a study of methods and procedures, so that any comprehensive research program concerned with applications of visible speech in deaf education should be one that is organized to continue for perhaps a minimum of five years.

Intermediate group

While in the nursery group, advantages both in speech education and in communication for general education, would be given consideration, the special needs of the intermediate group would doubtless call for initial emphasis upon speech improvement. The first several months might be spent simply in becoming acquainted with the patterns while using them for speech correction, with secondary emphasis upon learning to read them. Such a program would be carried on in addition to the current school program, or with as little interruption of this work as possible, and with the cooperation and participation of the classroom and special teachers.

When a stage has been reached at which the group could begin to use visible speech as a means for communication, some of the emphasis might better be shifted to the question of the possible advantages of visible speech as a medium for communication in general education, and it would then be necessary to decide how to proceed with this part of the study. One possibility might be to teach a single subject, such as history, by conventional methods supplemented with visible speech, comparing progress with that normally obtained by conventional methods alone. Another might be to use visible speech part of the time on all subjects. In either procedure, evaluation of progress would, to a degree, have to depend upon judgment, with particular attention to various factors such as enthusiasm on the part of teachers or pupils as well as other variables that might influence the results. The research should attempt to show deficiencies as well as advantages. Even though the overall advantages may appear in some instances to be substantial, it is a recognition of deficiencies that will be most helpful in suggesting improved methods and better equipment designs.

Adult group

In an adult group, there should be some congenitally deaf individuals, preferably with rather firmly entrenched incorrect habits of speech, some recently deaf desirous of maintaining their speech proficiency, and a spread of intermediate conditions. Initially, the interest of the congenitally deaf would undoubtedly be in speech correction, and their secondary interest would be in the use of visible speech for communication. As in the intermediate group, the best procedure perhaps would be that of concentrating upon speech improvement, with learning to read the patterns incidental to using them for speech correction.

While the ultimate interest of the deaf adult might be in having an individual translator that would enable him to use the telephone and listen to the radio, he probably would have a continuing interest in attention to his speech. In order to retain acquired improvements it would be essential, of course, that he practice occasionally. A particular question to answer is that of how this practice might best be instrumented. No doubt the quality of speech during a long period would depend upon the time devoted to its maintenance. Some deaf adults would, no doubt, be sufficiently interested in perfection to spend time regularly upon practice. Others would tend to let such routines lapse until the effect becomes so noticeable in their social contacts that need for improvement is self-evident.

Work with an adult group might be organized to investigate, for example: how rapidly and to what extent speech may be improved by both assisted and unassisted practice, considering various conditions of speech improvement; how much attention must be given to maintenance of speech after improvements are once established, and how this might best be accomplished; how rapidly and accurately the average adult with defective hearing is able to learn to read the patterns.

The hard of hearing

In the discussion thus far, it has been implied that the initial research program would be restricted to the deaf, or those with too little residual hearing to be of any practical use. What part should be played by the partially deaf and deafened? This is another one of the points that needs to be determined by the suggested studies. Where hearing is normal, visible speech presumably might be of some assistance in perfecting average speech and also in correcting the many common varieties of slightly defective speech. If visible speech proves to be of value to the completely deaf in producing acceptable speech, it may also be of some service to the hard of hearing, but where its contribution might begin or end is impossible to say without much more experience than is now available.

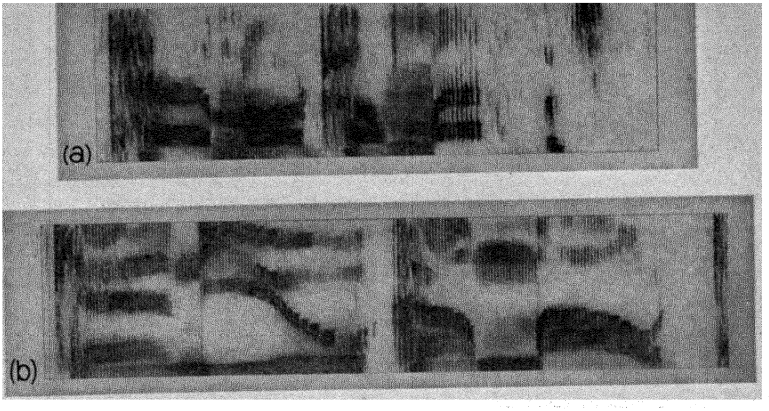


Fig. 1 — The sentence "Can you come out?" as recorded by: (a) a congenitally deaf adult; and (b) an adult with normal speech and hearing.

Patterns for Speech Education

In Fig. 1(a) is shown a sample of speech by a congenitally deaf adult, and below this in Fig. 1(b) are the same words spoken by an individual with normal speech and hearing. Essentially, the problem of speech education and correction is to make the pattern of the deaf individual correspond with the normal pattern. There are two general requirements for achieving such correspondence. One is to reproduce the articulatory movements associated with the resonation of voiced sounds and those concerned with the generation of unvoiced sounds. The other is to produce proper pitch changes or inflections in the voiced sounds.

Pitch may be displayed in two ways, one as an integral part of the speech pattern and the other by means of a separate display. In all the sound spectrograms of Part Two and also those in Fig. 1, pitch is shown as an integral part of the pattern. It appears in the disposition of the vertical striations or "pitch lines" of the voiced sounds. When these vertical striations are close together, the pitch is high; when they are far apart, it is low. The number of striations for a given time interval determines the frequency of the fundamental pitch, the frequency being high in high-pitched voices and low in those that are low-pitched. Variations in pitch appear in this integral form of display as changes in the spacing of the vertical lines.

A similar integral form of pitch display may be utilized in the visible speech translator by appropriate circuit changes. Also, the pitch-line spacings may be exaggerated or removed entirely if a separate pitch control circuit is used to produce the striations. If the pitch lines of the integral-type display had to be separated widely to make the

inflectional changes sufficiently evident for use in speech education, it might affect the pattern form adversely. Plosives and inconspicuous influences could be made less evident.

Other integral types of pitch display are possible. For example, pitch lines may be made generally horizontal rather than vertical. This might affect some pattern forms less, but it also could have a more adverse influence upon others. Much remains to be learned about the relative advantages and disadvantages of the integral-type pitch display. In its favor may be the incorporation of all the necessary information for speech instruction in one pattern. Against it may be the loss of pattern detail resulting from the combination. Again, this is one of the questions that will have to be answered by the visible speech research program, and the answer may be different for youngsters than it is for adults.

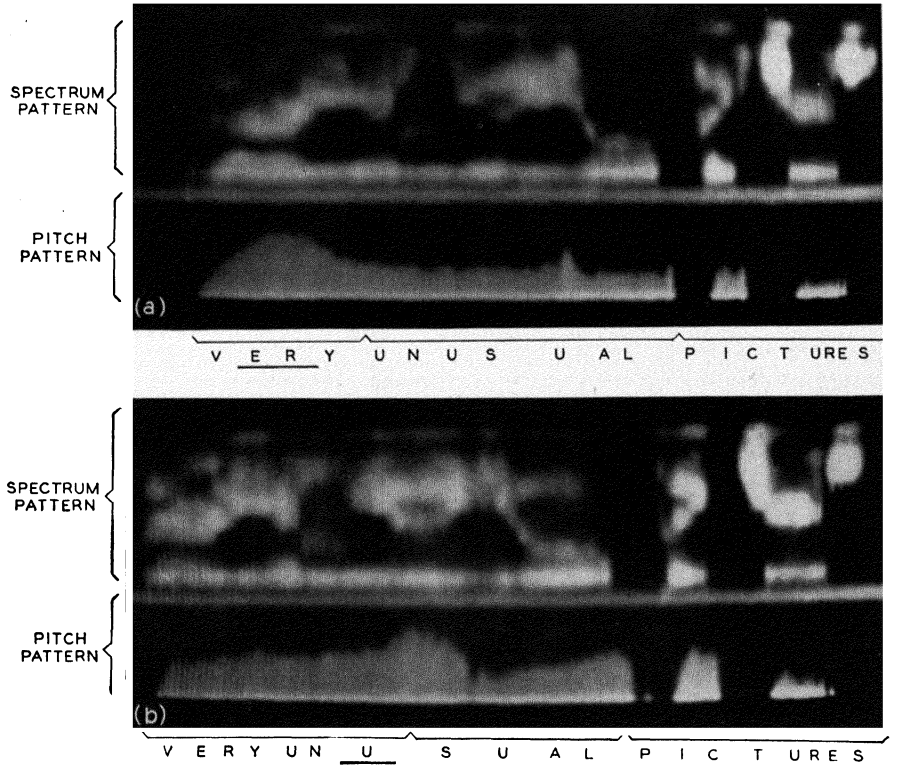


Fig. 2 — Speech pattern compared with pitch display. The stressed syllables are underlined in the legend below each illustration.

In Fig. 2 is shown a picture of a separate type of pitch display that has been applied to the experimental cathode-ray translator. The speech pattern appears at the top, and below it is another pattern showing the pitch change. The pitch is low where the vertical dimension of the pattern is low, and increases directly with that dimension. Other similar forms such as a single line that moves up and down with the pitch variations have also been considered.

Some of the pitch indicators that have been developed for use in deaf education in the past might also be thought of in connection with the visible speech patterns. Most of these are of a dynamic type in which a color changes or a light moves up and down with the pitch change.

Although we cannot be at all certain of the best ultimate arrangement, it would appear from present information that pitch change should be shown by a static rather than a dynamic display. Also, there would seem to be much in favor of an integral type to avoid the confusion arising in the use of two separate patterns. One might argue that the ear hears pitch as superimposed upon form, and that there should be some way found to picture these variables in combination without causing one to detract from the other.

Patterns of the deaf

Figure 3 shows the words that appeared earlier in Fig. 1(a), repeated by the same congenitally deaf individual after approximately fifty hours of rather casual speech training using the visible speech translator. Although this is admittedly sketchy evidence, it has seemed an encouraging indication in view of the fact that in mature pupils habits of poor speech are often rather fixed, and also considering that no pitch display was then available.

In Figs. 4 to 7, inclusive, sample spectrograms of the speech of deaf children are shown with comparison patterns of the same words spoken by an adult with normal hearing and speech. These youngsters had no training in visible speech, and their patterns are included merely as an illustration of the speech of the young deaf. Since

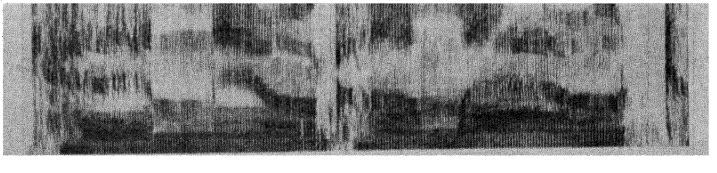


Fig. 3 — The sentence "Can you come out?" recorded by a congenitally deaf adult after approximately fifty hours of training with a visible speech translator.

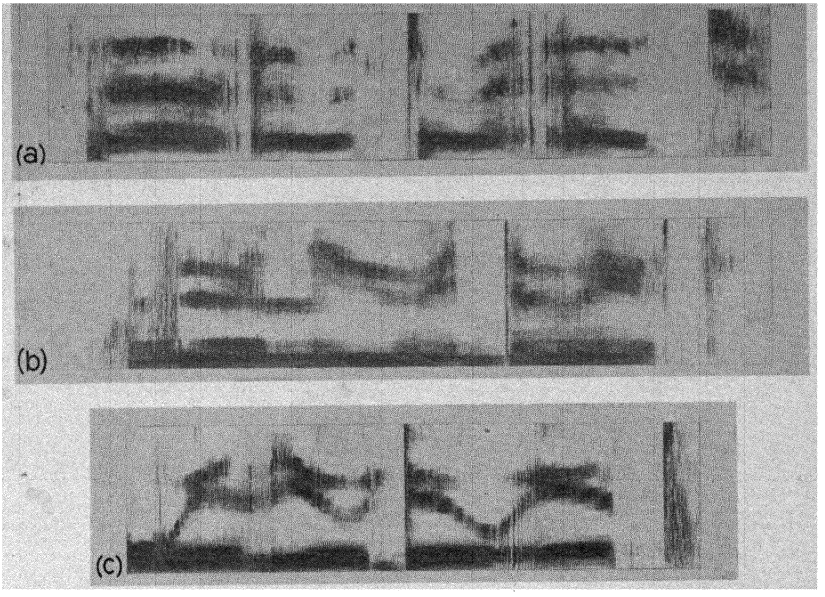


Fig. 4 — Broad-band spectrograms of the sentence "Will you do it?": (a) and (b) are the speech of deaf children; (c) is the speech of an adult with normal faculties.

differences between their patterns and the normal reference patterns are quite obvious, does it not seem very likely that the deaf will be able to use such visual comparisons as an aid in understanding and correcting their speech deficiencies? One of the first purposes of an educational research program should be to answer this question.

Specific Plans

At an early stage in the visible speech development, the conclusions of the present chapter, that an educational research program would be desirable and that this should be carried on in the educational field, were recognized. The question was how to initiate such a program in a way to make it most effective.

Following the first public demonstrations of visible speech, there were many discussions with experts in the fields of deaf education and speech correction. Interest in the experimental work was gratifying, and a number of competent groups expressed a desire to help evaluate the educational aspects of the development. If several translators of a suitable type had been available, the question of how to start the research

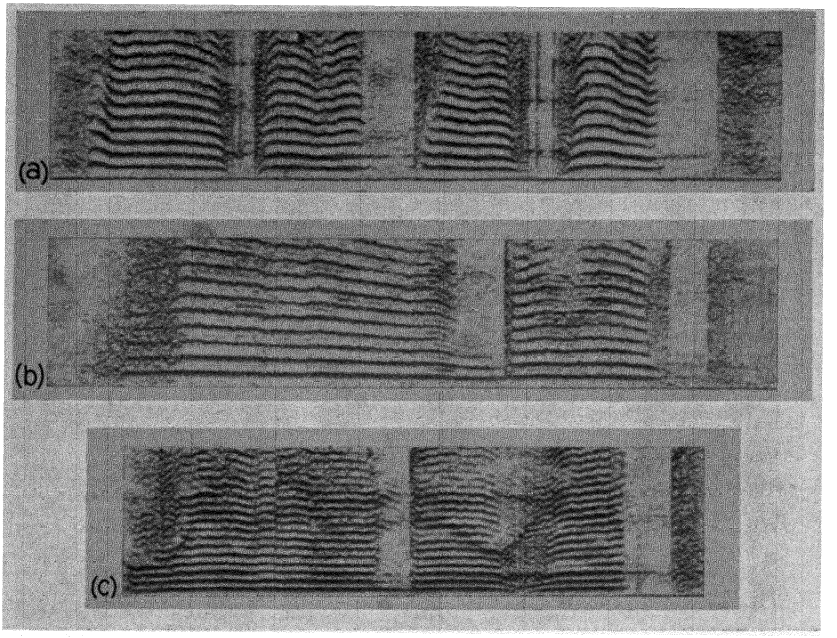


Fig. 5 — Narrow-band spectrograms of the sentence "Will you do it?": (a) and (b) are the speech of deaf children; (c) is the speech of an adult with normal faculties.

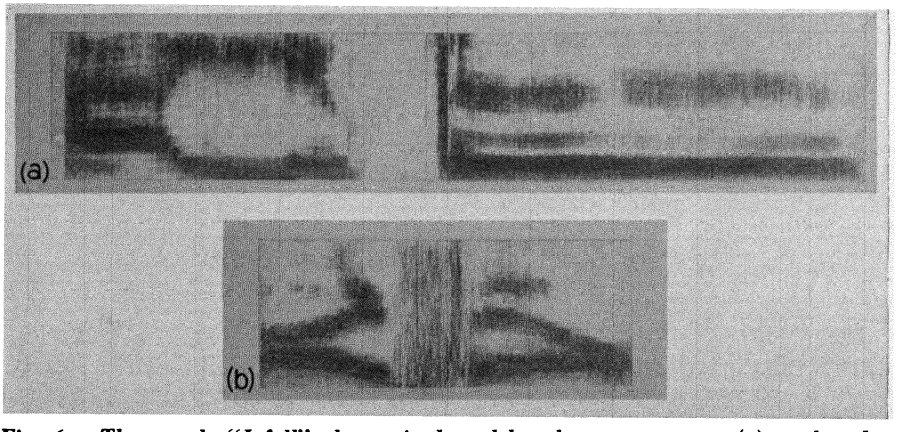


Fig. 6 — The words "I fell" shown in broad-band spectrograms: (a) spoken by a deaf child; (b) recorded by an adult with normal speech and hearing.

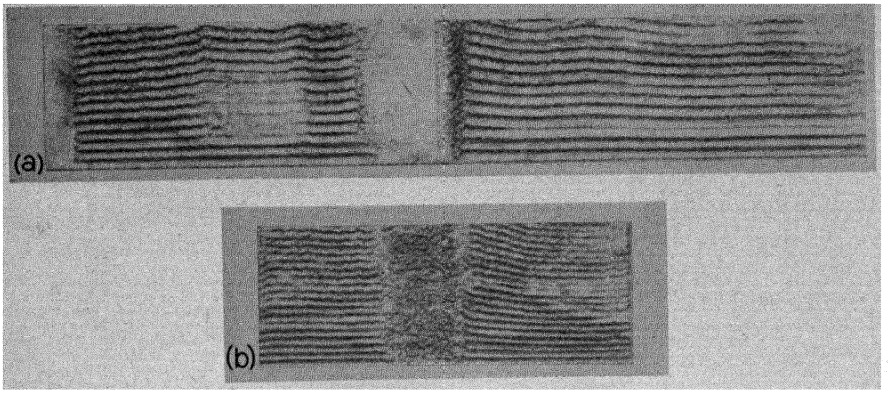


Fig. 7 — The words "I fell" shown in narrow-band spectrograms: (a) spoken by a deaf child; (b) recorded by an adult with normal speech and hearing.

would have been considerably simplified, but there was only one satisfactory unit, and even this was large and complicated.

Those in the Bell Telephone Laboratories responsible for the final decision as to where the one available unit should be placed, did so entirely on the basis of what arrangement might produce the best results with the least delay. The decision was announced * at the conclusion of a lecture-demonstration on "Visible Speech" before a meeting of the Convention of the American Association to Promote the Teaching of Speech to the Deaf, in Hunter College, New York City, on June 25, 1946. Following are quotations from this announcement.

"*We are all interested in seeing that possible deaf educational uses of visible speech are explored without delay. Such an exploration should obviously be carried on in the educational field, and so there is the question of where it might be established and how it might be implemented. It may be assumed that the Bell Telephone Laboratories would loan the equipment that has been used in the past experimental training; but the facts are that the best visible speech translator, and the only one suitable for experimental training, is a large, complicated, and very expensive unit — the one that the participants in the demonstration have been using this evening. The patterns of the large screen unit arranged for the audience to see are actually quite inferior when considered in detail, and the same is true of a very small unit that we built to*

* By Ralph K. Potter.

study the problems of size reduction. So the only equipment really available for implementing any progam at the present time, and probably for some time to come, is the large experimental cathode-ray translator.

"I feel strongly that the idea of evaluating this one visible speech translator only as a machine would be a serious mistake. It has demonstrated that speech can be turned into visible patterns that are readable, and it has helped establish the principles of visible speech, but it obviously is not ready for general use. It could, however, be used to determine how the principles *of visible speech may be applied to deaf education, taking into account the imperfections in design and, in fact, trying to determine what these deficiencies are — in other words, what the requirements are for translators of the future. What is needed is a research project rather than an apparatus trial — an experimental study of principles rather than the trial of a machine.*

"Also, if the possibilities of visible speech in deaf education are to be explored with the least delay, it has seemed to me very important that use be made of the accumulated experience available in those who have spent the past several years learning and teaching visible speech, investigating and organizing its principles, and otherwise cooperating in efforts to advance the development, and who are leaving the Bell Telephone Laboratories on September first of this year to return to the educational field. If the present translator were turned over to a group that had had no experience with visible speech, or with this particular equipment, the work would not only be delayed, but it would be necessary for the Bell Laboratories to divert effort from what, at this stage, is a very important part of the project, that of apparatus improvement.

"These points of view are explained in detail because they had a bearing upon the final decision, only recently made, as to the disposition of the experimental equipment. Dr. Kopp has been offered and has accepted a place on the faculty of the University of Michigan at Ann Arbor; and Miss Green has been offered and has accepted a position in the School of Special Education, Michigan State Normal College at Ypsilanti, nine miles from Ann Arbor, and she will also have a position on the staff of the University at Ann Arbor. In addition, arrangements have been made to loan to that University the equipment the Bell Telephone Laboratories has used to carry on experimental training in the past. The cathode-ray translator will be installed at the School of Special Education at Ypsilanti, where it will be used under Miss Green's supervision in experimental training of deaf and hard-of-hearing and other speech defectives; and a sound spectrograph and some other equipment will be located at the University of Michigan Speech Clinic in Ann Arbor, where it will be used by Dr. Kopp and others for various special studies. The work will be carried on as a cooperative project, with a committee rather than any one individual completely responsible.

"According to present plans, the first year would be largely exploratory, with the emphasis upon investigation of ways that visible speech might best be employed in

speech education of deaf children. Following the first year, but depending upon the results during this exploratory period, it is expected that the emphasis will shift to the question of how visible speech might be used in general education of the deaf.

"The classes at Ypsilanti will be small, but this should not be a handicap at the outset. If a wider range of experience seems desirable later on, there are several large schools for the deaf nearby in Michigan: one for example, at Flint, fifty miles from Ann Arbor, and another in Detroit.

"We are assured that the results of this work will be published regularly in University bulletins or elsewhere so that they will be available to all interested. There will, I feel certain, be a conscientious attempt to obtain early answers to the questions that we should all like to see answered, and I sincerely hope the project will have your confidence and support."

While studies of visible speech in relation to deaf education are going on in the educational field, investigation of the fundamental problems of sound portrayal will continue at the Bell Telephone Laboratories, both as a background for telephone transmission research and for the future development of visual telephony for the deaf. Before the latter will be possible, two major accomplishments will have to be realized: first, the establishment of visible speech in the deaf education field, and second, the production of very much more practical translators than any so far constructed.

Conclusions

No doubt the specific plans for research to be carried on in the educational field and for technical investigations to be continued in the Bell Telephone Laboratories, will be supplemented by additional activities. It is to be hoped that other groups will contribute both educational and technical ideas, even to the extent of developing new and better translating apparatus. One purpose of this book is to stimulate such contributions.

What can the deaf individual do to hasten the time when he may benefit by this development? At the moment it would be premature to suggest that he learn the patterns of visible speech, but it might be advantageous for him to become somewhat acquainted with them. He might even acquire an early and clear understanding of what is meant by visible speech and, in general, how the patterns are formed, so that he would be in a position to place his own evaluation upon their possibilities. Secondly, he can take an active part in encouraging investigations of visible speech wherever there is opportunity. He should bear in mind that a great deal of effort has been necessary to bring visible speech to its present stage, and that a great deal more will be necessary before its benefits will begin to be felt. However, the long-hoped-for conversion of speech to readable patterns is now known to be possible, and to know that a thing "can be done" is in itself a substantial step toward practical achievement.

Chapter 14
Phonetic Interests

Anyone with a knowledge of phonetics will find much that is familiar in Parts One and Two of this book. Some new terms have been introduced in Chapter 5, primarily because they are better adapted to the explanatory discussions of visible speech than are the usual phonetic terms, and in part because they serve to link the visual phonetic patterns with the underlying physical and physiological factors. In extending the aural and physiological classifications of sounds to include the visual, some reclassification was necessary, but a deliberate attempt was made to use customary phonetic categories and symbols wherever possible.

We have tried to present the available visible speech facts and information in a descriptive fashion so that hypothetical discussions of controversial issues might be avoided. As a matter of fact, any other treatment of the material would have been undesirable because of the limited amount of quantitative and qualitative data, limitations in the translation equipment, and, above all, restrictions in time due to the relative infancy of both the technical and the phonetic phases of the visible speech development.

Pattern Possibilities

Patterns for research in the various branches of phonetics may require different types of portrayal, depending upon the purpose of any particular investigation. Certain variables may need to be emphasized, and others suppressed or even deleted to meet specific needs. In all considerations of sound patterns, however, it would seem that the starting point ought to be the selection of the most suitable of three basic types of pattern. Conceivably, visible patterns may portray all that the ear can hear, or more than the ear can hear, or they may picture in detail one small segment of the range of audible or inaudible sound.

The ear as reference

For linguistic or phonetic purposes, the initial choice would appear logically to be in the pattern that attempts to depict all that the ear hears. Both the physicist and the phonetician have accumulated considerable information concerning the auditory characteristics of sound, and, in general, the mass of our knowledge of sound, particularly in the speech range, has been organized and interpreted from the point of view of the ear. But this does not imply that visible patterns of speech should necessarily have an emotional effect similar to that incurred in hearing the same passage. It does mean that, as a starting point, we should be able to see pitch changes that the ear can

distinguish, see as separate tones those that are heard as separate, see tones that are not masked by stronger tones adjacent in frequency and time, and even not see those that are masked in hearing.

If we are to see in sound patterns what is heard in the original sounds, the three basic dimensions of sound (frequency, intensity, and time) will need to be displayed in a manner that approximates the action of the ear. Obviously, such a task is not an easy one. Differences in the physical nature and function of the eye and ear cannot be resolved without a great deal of further research. In order to have useable visible patterns during such a research interval, we may quite reasonably start from present information on hearing, resulting from such measures as articulation tests and pure tone tests.

One of the general problems to be investigated concerns our intent in listening, or our interest, since it tends to govern what and how we hear. When listening critically for particular characteristics in voice and diction, we may hear differently than when listening as in ordinary conversation. Such auditory "gear shifting" may suggest different patterns for casual and for critical purposes. Or would the visual equivalent be that of looking at one "complete pattern", but concentrating upon different features to satisfy different interests?

Another complication results from the fact that frequency resolution varies in cycles per second over the hearing range. Therefore, if the characteristics of hearing are to be simulated in equipment for producing sound pictures, the variation of resolution with frequency needs to be considered in the display. Where resolution is greatest in cycles per second, near the low-frequency end, the pattern should be expanded to show this detail. Where resolution is smallest, at the high-frequency end, the pattern should be contracted.

Aural experience with octave steps suggests adoption of what might be termed an octave or logarithmic frequency scale. If resolution were a constant percentage through the audible range, the trace of a tone varied from low to high pitch would be of constant width. Over mid-range, the resolution of the ear does approach a constant percentage value, but at the low end the percentage value rises, so that the trace of a low tone on a display with a logarithmic frequency scale is widened. Does something analogous to this effect in aural perception account for the feeling that low tones are "heavier" than higher tones?

Logarithmic display

Although problems relating to logarithmic patterns have been discussed, as yet only a few experimental patterns have been produced. An early speech pattern, incorporating variable frequency resolution and a logarithmic frequency scale, is shown in Fig. 1. In this pattern, overtones of the voice are spaced logarithmically rather

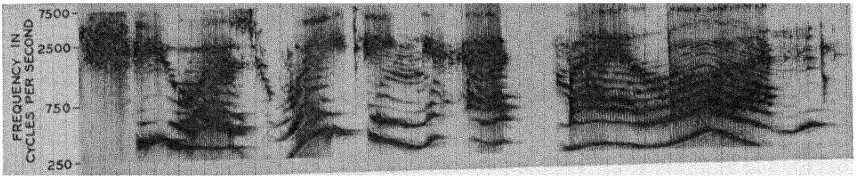


Fig. 1 — Logarithmic scale pattern of the words, "She was waiting on my lawn".

than linearly as in all other narrow-band patterns shown in this book. Another difference is in the response to pitch change. In the linear scale patterns, a rise in pitch causes an increased separation of the overtone traces, and as the pitch falls they appear to converge. When a logarithmic or octave scale is employed, variation in pitch results in a rise or fall of the whole overtone group without changing the amount of separation between adjacent harmonics. Incidentally, some peculiar transients appearing in these patterns were caused by a feedback type of filtering circuit.

Although the pattern of Fig. 1 is not an obvious improvement over the forms now in use, it is possible that greater emphasis of the resonance effects, and a modified feedback and volume control action, would emphasize important features, making their visual perception comparable with the aural. Note especially the emphatic treatment of plosives, the wide frequency range included, and the clear display of pitch inflection. This pattern appears to be crude and seemingly distorted after prolonged experience with linear types, but it may mark the future development trend. It is not really as strange as it appears to be on first examination. All of the same voiced and unvoiced sounds are still recognizable; only positions, shapes, and sizes are somewhat altered.

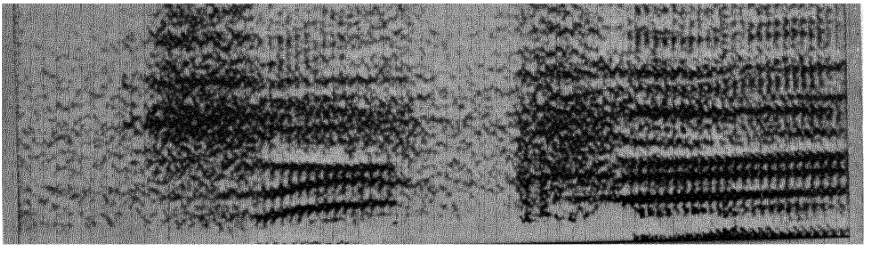


***Fig. 2 — Linear pattern (11,000 cycles, 0.8 second) of the syllables* si (*see*) *and* ki (*key*).**

Linear display

Patterns included in this text are, except where otherwise indicated, examples of a linear display. Such patterns may be varied in form by modifying the portrayal of any or all of the three basic dimensions, frequency, time, and amplitude. Samples of a few of the variations which may be of interest to the phonetician are included in this section, and reference is made to several used in other chapters. Most of these patterns may be produced by comparatively simple modifications of the present spectrographs. It is, of course, impractical to attempt to show all of the available linear pattern types, but perhaps these illustrations will suggest the possibilities of emphasizing or de-emphasizing specific factors in order to satisfy special requirements.

Fig. 2 illustrates a type of pattern especially useful for sound analysis, in which the frequency range is extended to approximately 11,000 cycles. The length of the spectrogram is about eight tenths of a second, and the analyzing band width is about 150 cycles. Considerable detail may be seen above the 5,000 cycle point, a little below the middle of the pattern. In the linear patterns used in this text, the cut-off point is a little above the fourth resonance bar in Fig. 2. At the left, the syllable is **si** (*see*); at the right, it is **ki** (*key*). The resonance bars of the vowel and the influence of the preceding sound are quite evident. It is interesting to note the predominance of the high frequency components in the **s**, and the energy concentration around 5,000 cycles for the **k**. The light random striations in the background result from noise.

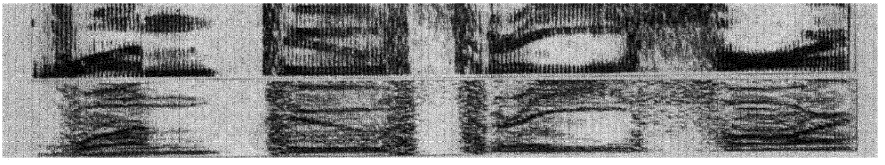


Fig. 3 — Compressed linear pattern (3,500 cycles, 2.4 seconds) of the words "one, two, three, four".

In Fig. 3, the frequency cut-off is at 3,500 cycles, but the vertical scale has been compressed in the original spectrograms from two inches to one inch. Since the time scale was not altered in these patterns, the aspect ratio (time-frequency ratio) is modified. However, the difference is not sufficient to make reading recognition difficult. These patterns may be compared with patterns of the same words, "*one*, *two*, *three*, *four*", as they appear in the lesson chapters. Convenience in recording, filing, and reproducing for texts or articles, may make the pattern with a vertical dimension of one inch or less quite valuable. Of course, reduction of the horizontal scale by a factor equal to that used in the vertical scale will permit of a constant aspect ratio.

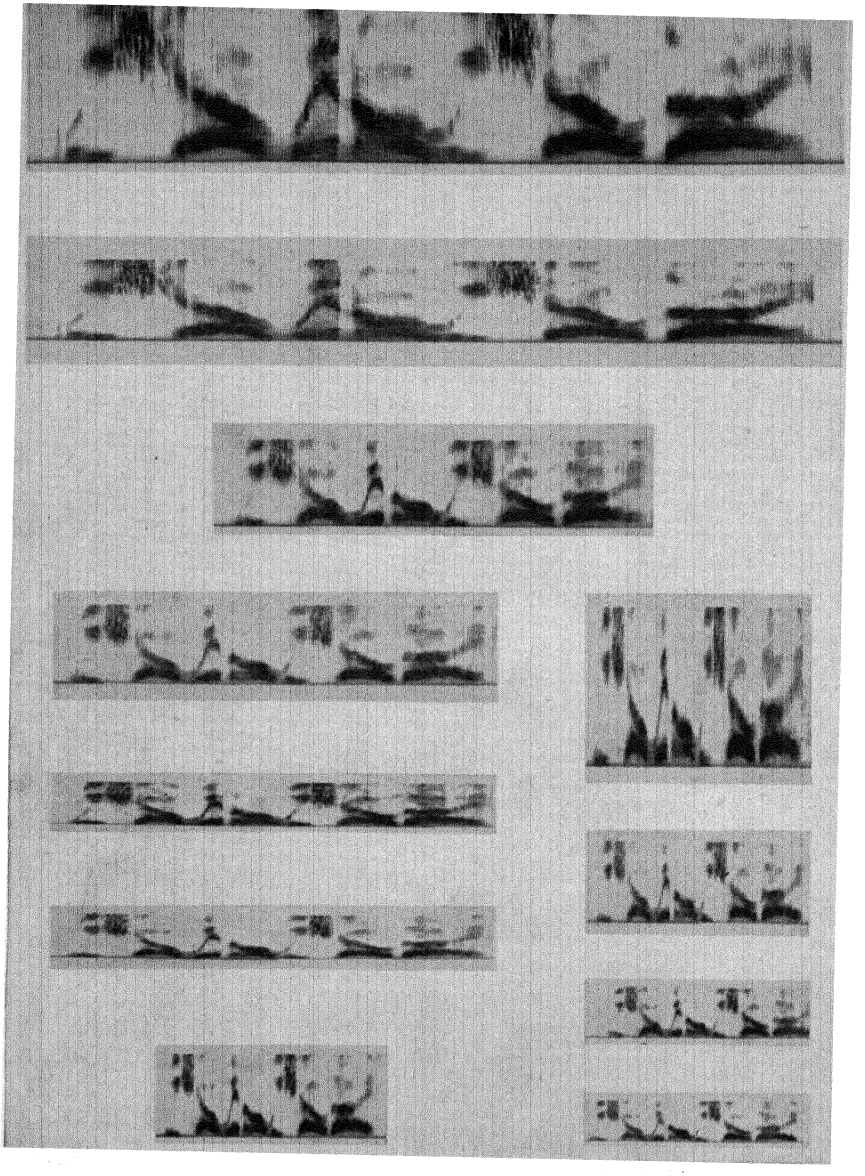


Fig. 4 — Variations in aspect ratio of linear pattern (3,500 cycles) of the words "We shall win or we shall die."

Varying the sample length is another possibility. If the frequency-time ratio is kept constant, samples of different lengths may be obtained by spectrograph modifications. Spectrograms of a long sample type are used in the chapter on Foreign Language Interests. Their value in phonetic research, as well as in many other fields of interest, is apparent. By varying both the time and frequency dimensions, the aspect ratio may be used to emphasize specific factors. For instance, it has been suggested that pattern legibility might be improved by elongating or varying the interval in which transitional influences occur. Fig. 4 illustrates some possibilities in varying the aspect ratio. In all cases in this figure the frequency cut-off is at 3,500 cycles, but both the vertical and horizontal scales have been varied in a number of different ways.

To the experienced phonetician, the inadequacies of the present translator for certain research purposes doubtless will be apparent. Substantial improvements and modifications are possible, but achievement of the ultimate flexibility and efficiency will be a continuing process, extending over years to come as more is learned about requirements and as new techniques for translation become available.

Phonetic Research Possibilities

Aside from the phonetic organization and information discussed in the lesson chapters, there are numerous problems of particular interest to the phonetician in which visible speech might be used to considerable advantage. There appears to be a large field of application in such stimulating issues as the extension of phonemic concepts; definition of the function of the syllable; and in the quantitative analyses of certain aspects of vocal quality, of speech sounds, of stressing and unstressing, and of rate and rhythm of speech. The field of comparative phonetics and linguistics also may find valuable evidence through studies of visible patterns. A few brief but thought-provoking illustrations pertaining to such investigations are presented in the chapter on Foreign Language Interests. There is no end to the number of potential applications that are suggested to the phonetician or to specialists in related fields who are concerned with phonetic notations, recordings, and analyses.

One additional application of visible sound patterns that might be mentioned is in the teaching of phonetics and in the training of phoneticians. Here, at last, we have a means of developing standard reference materials to be used in conjunction with ear training.

In general, the illustrative material in this chapter is presented as interesting but uninterpreted evidence of the potentialities of the present equipment so that the reader may evaluate them from his own viewpoint.

Relation to speech production

Occupying a prominent place among questions of widespread interest, is the relation of the speech sound to the mechanics and mechanisms of sound production.

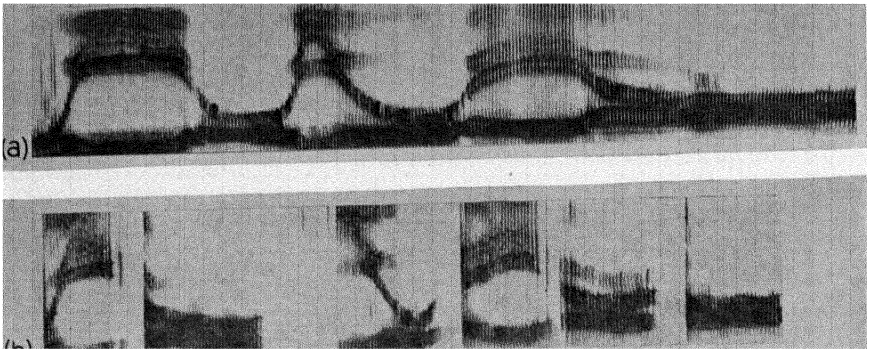


Fig. 5 — Resonance continuity in the words "We owe you aye our awe": (a) continuous speech; (b) voicing terminated after each word.

Consideration of the aural and visible sound patterns in connection with the articulatory process might be facilitated, in one phase of research, by comparing continuous speech patterns with patterns of sound units as illustrated in Chapter 5, page 38. It is not necessary to do more than mention the opportunity afforded for association of X-ray photos of moving and static articulation, with these visible patterns.

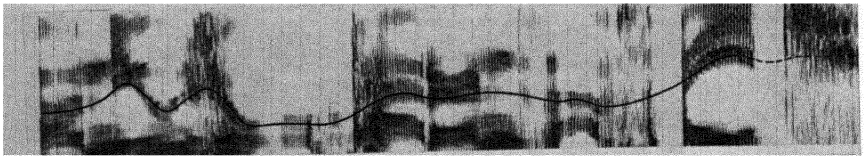


Fig. 6 — Movement of resonance bar 2 in the words "Unusual patterns of speech".

Resonance continuity might be studied by plotting continuously the potential resonance values for the various vocal cavities, and by studying the degree of coincidence of the potential plot with the resonance movements and positions as portrayed by visible speech. A simple form of illustration of resonance continuity may be seen in a comparison of Figs. 5(a) and 5(b), and also in Fig. 6. In Fig. 5(a), the words "*We owe you aye our awe*"* were spoken continuously in fairly rapid speech. Fig. 5(b) shows the same phrase with voicing terminated after each word. To trace the resonances through non-vowel sounds is more difficult. Movement of bar 2 in the words "*Unusual patterns of speech*" is traced approximately by the continuous line in Fig. 6.

* These words were used by Paget in his book "Human Speech", p. 46, to illustrate the recording of resonance positions by musical notation. It is interesting to compare the visible speech patterns with his melodious transcriptions.

Resonance sources

To correlate resonance movements in visible speech patterns with the concomitant variations in the vocal tract would require an identification of the associated cavities with the positions, shapes, and so-called "influences" of the resonance bars. Certainly opinions regarding such relations differ at the present time, and the matter cannot be settled satisfactorily without more facts.

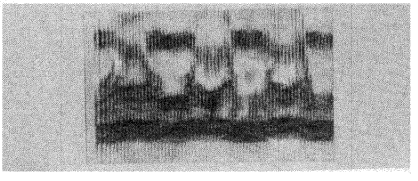


Fig. 7 — Effect on the vowel ɑ ***(father) of pulling the cheeks in and out.***

Disturbing the resonances can provide some clues. Fig. 7 shows the effects of pulling the cheeks in and out while saying ɑ (*father*). Although all the resonances are disturbed, the second and third bars appear to be changed most. Moreover, the changes are in opposite directions: as the frequency of the second bar decreases, that of the third bar increases.

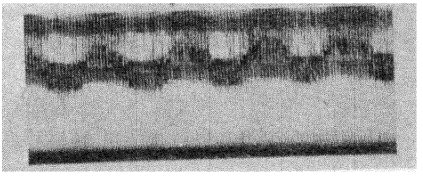


Fig. 8 — An attempt to say the vowel i ***(eve) with a small rod alternately in and out of the mouth opening.***

When a sound resembling the i (*eve*) was produced from an articulatory position that permitted the alternate pushing and pulling of a small (half-inch) rod in and out of the mouth opening about one-half to three-quarters of an inch, the result was as shown in Fig. 8. Bar 2 is moved up and down while there is no perceptible change in the position of bar 1. The influence upon bar 3 is indefinite, but it seems to be mainly one of change in intensity.

Fig. 9, (a) through (e), illustrates the effect of variations in resonation upon the vowel i (*eve*). Many variations are possible in the degree and nature of nasality and denasality, and in the manner of articulation and direction or obstruction of the breath flow. Such variations are reflected in the visible sound patterns in the appearance or

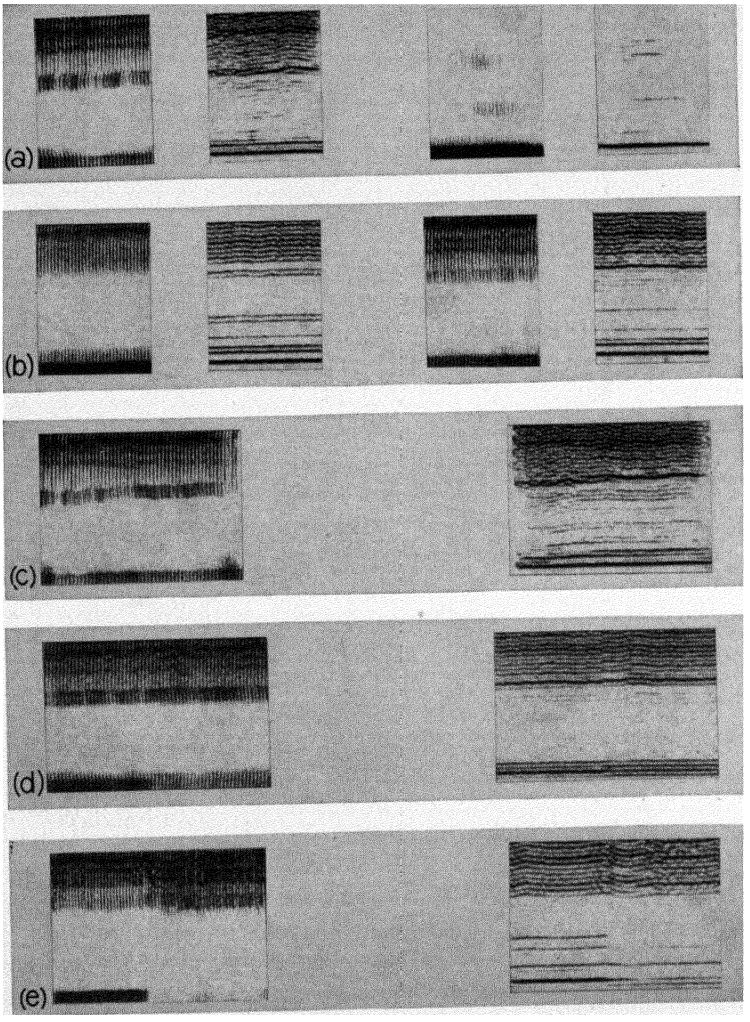


Fig. 9 — Effect on the vowel i ***(eve) of variations in resonation shown in broad (300-cycle) and narrow (45-cycle) band patterns: (a) left, normal; right, lips closed; (b) positive nasality: left, energy largely through nasal cavities; right, energy largely through oral cavities; (c) normal to nasal; (d) normal to blocked normal; (e) nasal to blocked nasal.***

disappearance of resonance bars, in shifts in resonance positions, and in changes in the relative predominance of resonance areas. All of the variations illustrated in Fig. 9 have been shown in both the broad and the narrow band patterns so that the effects may be seen clearly on harmonics as well as resonances.

An i (*eve*) produced normally is shown in Fig. 9(a). The pattern at the right end of Fig. 9(a) represents an attempt to say the i with the lips closed. Such a sound cannot be continued at will because of the pressure built up, and its pattern resembles the patterns of preliminary voicing for the voiced stops. Fig. 9(b) illustrates two types of positive nasality. At the left, energy was directed primarily through the nasal cavities, and here, the absence of the second resonance bar is quite striking, as is the strong first bar. At the right, energy was directed primarily through the pharyngeal and oral cavities, and secondarily through the nasal cavities. The upper resonance bars are not as marked as in the normal pattern, but bar 1 appears slightly stronger. Fig. 9(c) illustrates an i produced by starting normally and then gradually lowering the soft palate until a nasal i is produced. The shift in bar positions is evident in bars 2 and 3, and increasing emphasis on bar 1 is noticeable in the nasal portion of the vowel. Fig. 9(d) shows the effect, on a normally produced i, of restricting the nostrils by pinching. In Fig. 9(e), the change is shown in a nasal i when the nostrils are restricted by pinching. Restriction of the nasal cavity has conspicuously less effect on the normal vowel than on the nasal, but in both patterns the de-emphasis of bar 1 is apparent. A shift in the position of bar 3 may be observed during the brief transition immediately following the restriction. It is interesting to compare these patterns and to speculate on the cavity-movement-pattern relationships. Until further evidence is collected, however, it must remain just speculation.

Movements as revealed by these visible patterns of speech should be of particular interest. No doubt what Paget calls the "gestures of speech" are subject to the same limitations of inertia and continuity that tend to restrict speech and accuracy of movement in other parts of the body. It should be possible to study these restrictions by following the alterations of movement that accompany increase in speech rates. Possibly speech patterns synchronized with X-ray pictures would help materially to round out the present knowledge of speech production.

Non-linearities

That non-linearities exist in restricted air streams is demonstrated by Fig. 10, which illustrates a sound produced by simultaneously whistling and voicing. The whistle pitch was varied up and down in the region between 1,000 and 1,500 cycles a second. In the absence of voicing, it would appear as a single wavy line, but due to non-linearity, the low-pitched voicing tones modulate the whistle note to produce sideband

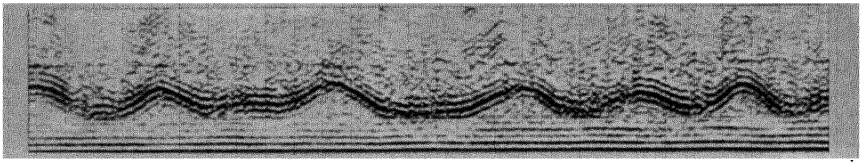


Fig. 10 — Simultaneous whistling and voicing.

products. The frequencies of these sideband products are equal to those of the whistle plus and minus those of the voice overtones. As a result, several parallel lines appear in a wavy band. The darkest is presumably the actual whistle note. The number of sideband components that appear is doubtless limited by the manner in which the voiced overtones fall off in intensity with frequency.

A resonance supports the oscillations of the whistled tone, and some indication of how resonance modulation affects sounds originating at the larynx and at the lips is afforded by Fig. 11. Voicing and a whistle are shown alternately instead of simultaneously as in Fig. 10. An effort was made to avoid any change in lip position or cavity shapes. The voicing shows one resonance that seems to conform with the whistle pitch.

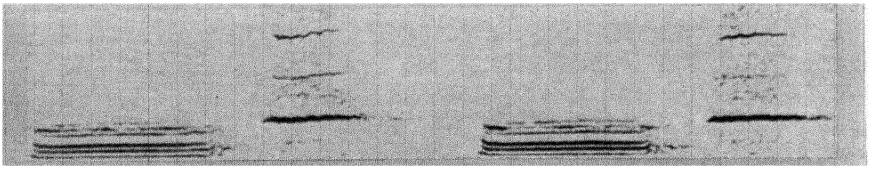


Fig. 11 — Alternate voicing and whistling.

Rhythm in speech

By recording speech in such a way that its energy envelope only is reproduced, it is possible to learn something about the effects of recurrences such as occur in the recital of rimes or poetry. In one form of portrayal, the rectified speech envelope wave is speeded up one hundred times and translated to sound pattern form as if it were an audible note. Syllabic components of a one-cycle-a-second rate thus would become oscillations at 100 cycles a second, while those at ten cycles would be boosted to 1,000 cycles.

Oral reading of ordinary page material, when reproduced in this way, results in a randomly mottled sound spectrogram such as appears in A of Fig. 12. The only evi-

dence of regularity is in the pauses for breath that appear as faint white lines forming a vertical grid. The length of this sample is about 40 seconds, and the same time scale is used in B and C. The vertical frequency scale is from zero to about 35 cycles.

When "*Hickory, dickory, dock*" was recited, the result is as shown in B. Notice that a semblance of order is beginning to appear in the low-frequency parts, although the upper components are still randomly related. In C, are shown four short samples that illustrate the effect of repeating words. To produce the first sample at the left, the word "*telephone*" was repeated over and over. For the second, two words "*telephone*" and "*spectrograph*" were repeated alternately. In the last sample, the words "*acoustic*" and "*spectrograph*" were used. The use of the same word extends the systematic relationship to a higher frequency.

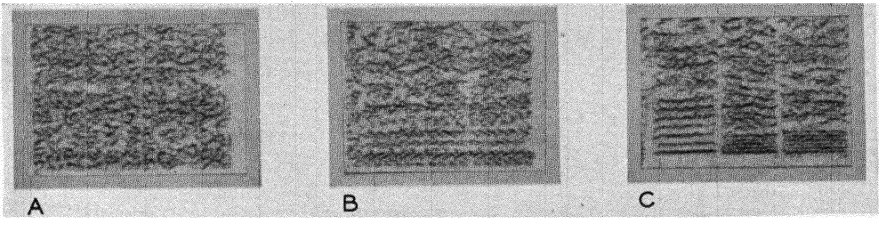


Fig. 12 — Variations in speech rhythm (35 cycles, 40 seconds): A, oral reading of prose; B, recitation of "Hickory, dickory, dock"; C, repetition of words.

Vocal quality

Vocal quality, generally, has received only limited consideration in this volume. Variables included under factors of quality contribute much to the basic patterns of speech in the way of color, individuality, emphasis, and supplementary meaning, and they are exceedingly important to the phonetician and also to specialists in other phases of speech. Adequate treatment of quality factors needs a type of systematic organization similar to that applied to those variables of visible speech discussed in the lesson chapters. While these quality factors are the pattern detail, to handle them adequately demands more knowledge of sound portrayal problems and greater control of the recording process than we now possess.

Figure 13 illustrates several common variations in vocal quality as imitated by a trained phonetician. Variations in quality are shown rather clearly even in the 300-cycle pattern. In all five patterns the words spoken were "*Can you come?*" Patterns (b) through (e) may be compared with pattern (a) and also with the patterns made by the deaf speaker in Chapter 13. Additional resonance bars are visible in the spectrogram illustrating positive nasality, especially towards the bottom and the top vowel

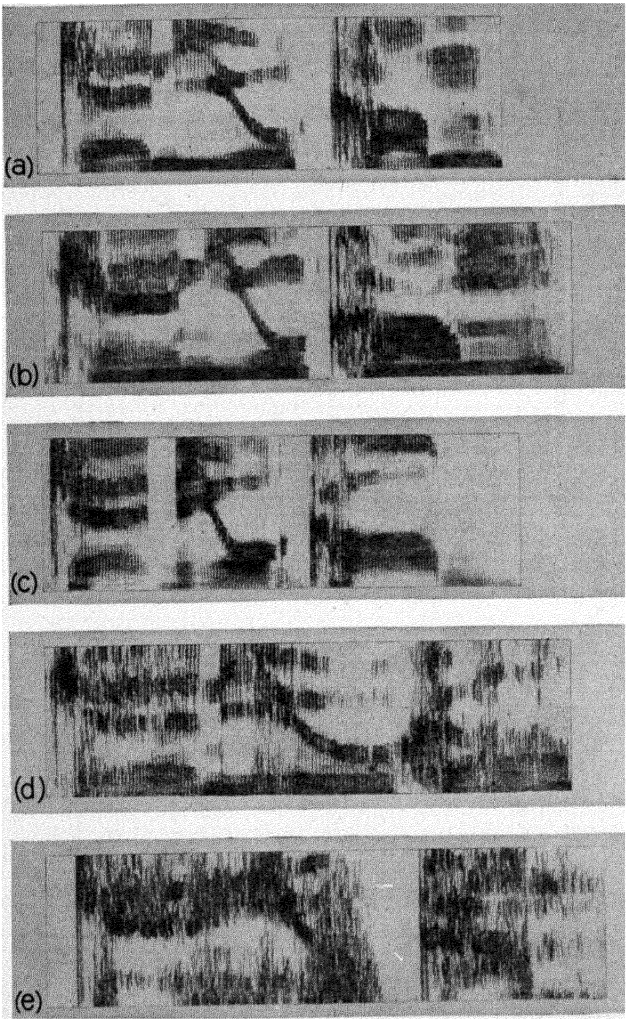


Fig. 13 — Variations in vocal quality in the words "Can you come?": (a) normal reference pattern; (b) positive nasality; (c) negative nasality; (d) aspirate; (e) whisper

resonances. Notice the reinforcement of the bar along the baseline. In the negative nasality or denasal pattern, the omission of the resonances for the nasals is apparent, and the absence of the baseline bar is quite striking. Although the nasal sounds are missing, the transitional influences are similar to those in both of the preceding illustrations. Comparison of the aspirate or breathy quality with the whispered and normally voiced patterns, illustrates the partial disintegration of the vowel resonances in the aspirate and the change in the appearance of the resonances in the whispered pattern. Some of the resonance forms are retained, however, in whispered speech and certainly such patterns provide an interesting means for investigation of the phonetic aspects of whispered speech.

There is no doubt that considerable detail can be seen in these sound patterns, but the situation is probably a little like the first experiments with a microscope. It is likely that many strange new forms were made visible by the first microscope, but to start organizing this new knowledge before comprehending fully what the microscope was doing would have been a very impractical approach. Similarly, proper organization of quality detail should await more complete understanding of the translation processes involved.

Miscellaneous Illustrations

Some miscellaneous items of popular as well as scientific interest are included in this section. They represent, in casual survey fashion, some of the avenues of interest to which visible speech may be applied.

Old and new techniques

Those who have studied Edmund Willer Scripture's "Elements of Experimental Phonetics" * will recall the extensive analysis of word sounds based upon his wave records. An example of one of these from Appendix II of Scripture's book (page 590, Fig. 357) is shown here as Fig. 14(a). This is a record of the words, "*Who'll be the parson?*" from the story of Cock Robin. The same words by another speaker are shown in visible speech form as Fig. 14(b). This latter pattern is analogous to the pattern of a rug, spread out so that it is clearly visible. The former is analogous to seeing all the threads of the rug unravelled and bundled together. The pattern material is there, but in no shape to be visualized as a meaningful whole.

* Charles Scribner's Sons (1904).

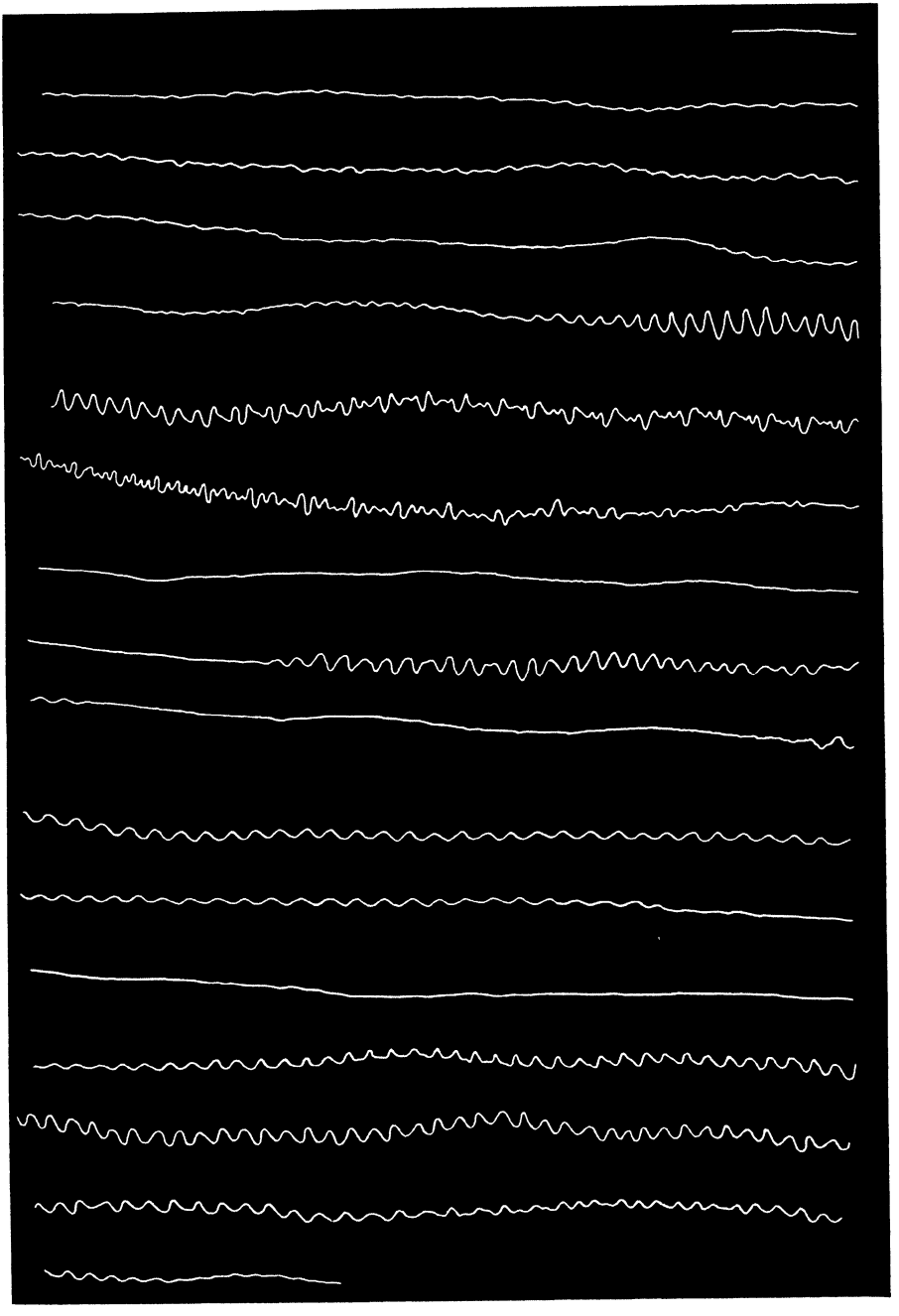


Fig. 14(a) — The words "Who'll be the parson?" from the story of Cock Robin, as shown in Scripture's wave records.

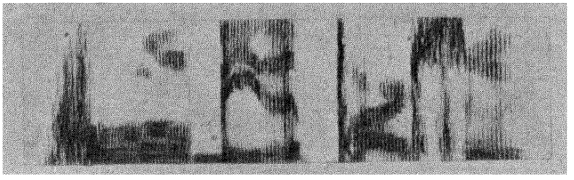


Fig. 14(b) — The words "Who'll be the parson?" as shown in visible speech

Tongue-twisters

So-called "tongue-twisters" might be investigated to advantage by patterns of visible speech. One of these difficult combinations of sounds, "*She sells sea shells*" is shown in Fig. 15. While the individual word units may be articulated easily, in rapid continuous speech the articulatory movements tend to become awkward and substitutions for one or more sounds or repetitions of sounds are common. Visible patterns may help in following the shifts in articulation and the telescoping of movement that take place during rapid repetitions of such phrases.

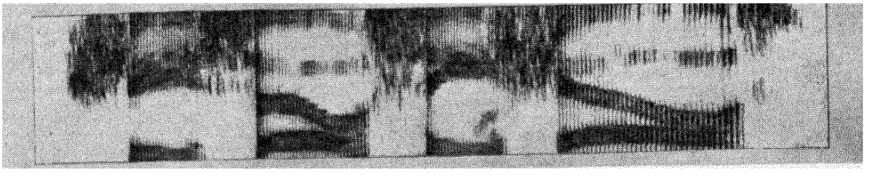


Fig. 15 — Pattern of the words, "She sells sea shells".

Vocoder speech

By means of an instrument known as the "vocoder" *, speech is converted to a series of fluctuating currents that may be transmitted to a distant point and there used to reconstruct the sounds artificially. Voicing is derived from a relaxation oscillator, and fricative sounds from an electronic noise source. Bands of these sounds are combined in automatically controlled amounts to produce the artificially formed speech. Such speech produced by one type of vocoder is illustrated by Fig. 16.

* "Remaking Speech", Homer Dudley, Journal of Acoustical Society, Vol. 11, p. 169 (1939).

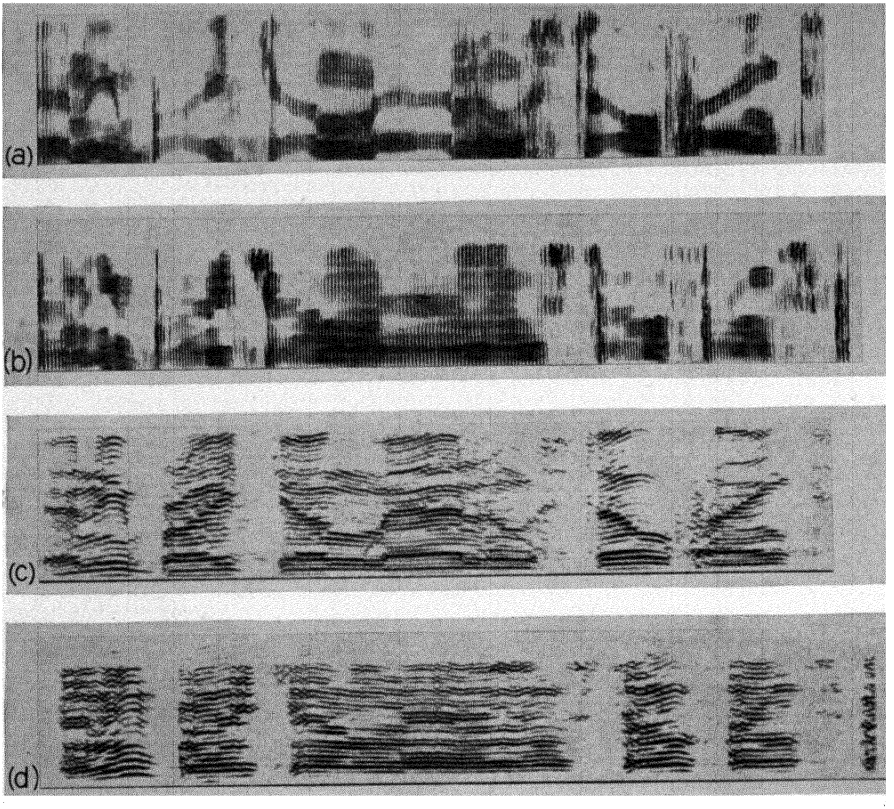


Fig. 16 — Patterns of vocoder speech of the words, "Any bright young man who still thinks": (a) normal reference pattern (wide band); (b) vocoder pattern (wide band); (c) normal reference pattern (narrow band); (d) vocoder pattern (narrow band).

In Fig. 16(a) appear broad-band patterns of the words, "*Any bright young man who still thinks*" spoken normally. The same sounds after going through the vocoder are shown in Fig. 16(b). In Figs. 16(c) and (d) are the narrow-band equivalents of patterns 16(a) and (b). It will be noticed that the frequency bands used in reconstruction of the speech show up distinctly in the broad-band patterns of 16(b), and that a good deal of the detail is lost in going from (c) to (d). It is fundamentally necessary that detail be lost since the number of analyzing and synthesizing bands is small in this particular instrument.

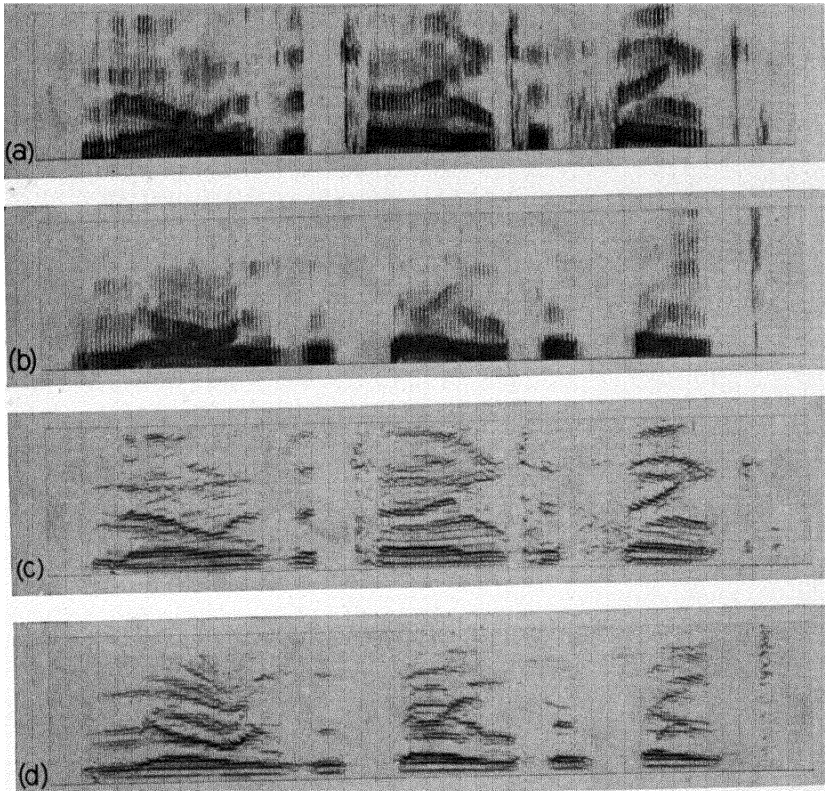


Fig. 17 — Speech with artificial larynx. The words, "Now is the time to think": (a) normal reference pattern (wide band); (b) artificial larynx pattern (wide band); (c) normal reference pattern (narrow band); (d) artificial larynx pattern (narrow band).

Speech with artificial larynx

A type of semi-artificial speech is illustrated in Fig. 17. Fig. 17(a) shows the words, "*Now is the time to think*" spoken normally, and (b) shows the same words produced by using an artificial larynx instead of the vocal cords. Both patterns were produced by the same speaker, so that the cavities coincide except insofar as they are modified by the use of this instrument * or by small differences in the successive pronunciations. The upper bars are not so much in evidence in (b), and the fricative sounds

* This was a Western Electric Artificial Larynx, Type 2-A.

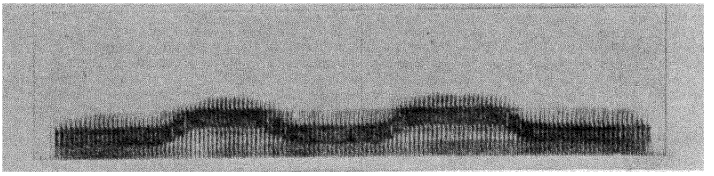


Fig. 18 — Patterns of vowels made by an artificial larynx with cupped hands as a cavity.

are missing. Although the artificial larynx only provides the voiced sounds, an experienced user can introduce his own fricatives to produce a combination that sounds fairly normal.

Loss of inflection in the artificially produced speech is evident in a comparison of the narrow band patterns, (c) and (d). In the early part of Fig. 17(d) are some resonance-like effects that have not been identified.

Cupped hands as cavity

Various experimenters in the past have tried the trick of producing speechlike sounds by forming a variable cavity in front of a sound source with the cupped hands. Fig. 18 shows the sounds "u–ɑ–u–ɑ–u" produced in this fashion with the artificial larynx as the voicing source. Only one cavity is in evidence, and its resonant frequency varies from approximately 600 to 1,000 cycles per second, which would coincide roughly with the lower half of the bar 2 swing. This suggests that the largest size

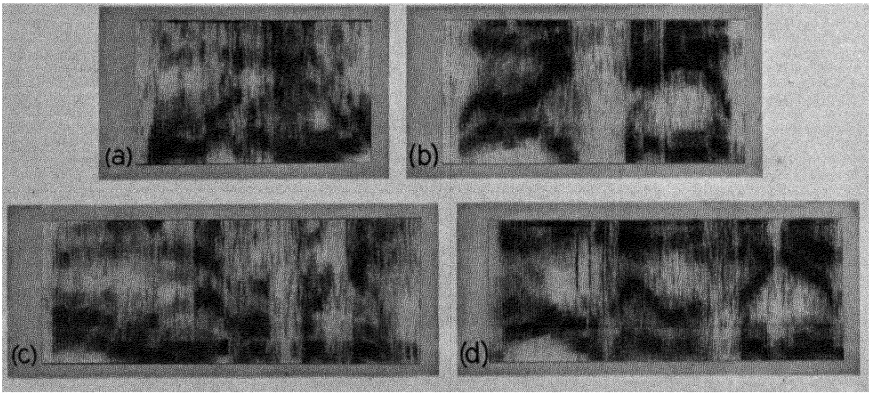


Fig. 19 — Patterns of the speech of a ventriloquist. (a) and (b) "Well, I tell you"; (c) and (d), "I'll never forgive you": (a) and (c) normal voice of ventriloquist; (b) and (d) voice of dummy.

of the bar 2 cavity might correspond to that of the cupped hands, and the smallest to half the size of the smallest cupping used to produce Fig. 18. Of course, this suggestion is open to suspicion on the basis that other variables may not be equivalent.

Ventriloquism

When a ventriloquist speaks in his natural and his affected voices it would seem there might be some consistent relation between the resonance conditions. Apparently, though, the bar positions are shifted appreciably. In Fig. 19 are shown four samples of speech by a ventriloquist, those of (a) and (c) being in his normal voice and (b) and (d) in the "projected" voice. The words in (a) are the same as in (b), and those in (c) are the same as in (d). Those in (a) and (b) are, "*Well, I tell you.*" In (c) and (d) they are, "*I'll never forgive you.*"

When the dummy is speaking, there is a substantial upward shift of both bar 1 and bar 2, averaging on the order of 50 per cent. In addition, there is a change in pitch as shown by the spacing of vertical striations, the higher pitch accompanying the higher resonance positions. Increased resonance frequency indicates a reduction in cavity size. Could it be that this range of cavity size change is accomplished by talking with the mouth fully open in one case and nearly closed in the other?

In a superficial examination of ventriloquized speech, a very unusual sound resembling a trill was discovered. This is shown to the left of center in the narrow-band pattern of Fig. 20(a). The first several words including the trill are ventriloquized and they are, "— *that stuff* (trill) *I love it!*" The last word on the record is "*Yeah!*" in the natural voice of the speaker.

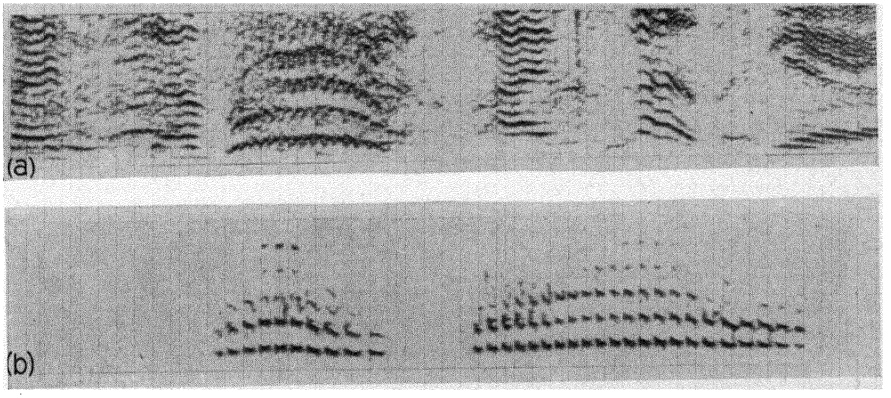


Fig. 20 — Pattern of a trill in ventriloquized speech: (a) trill produced by a ventriloquist; (b) attempt of another speaker to imitate the trill.

It will be noticed that the trill consists of unusually rapid pitch fluctuations. The rate is about 35 a second as compared with 5 to 10 cycles a second in the normal vibrato. An effort by another speaker to reproduce this kind of thing was not at all successful. Vibrating the tongue by a stream of air while voicing produces the effect shown in 20(b). To vary the pitch would seem to require a similar oscillation much closer to the vocal cords. Could this be produced by an oscillation of the soft palate or some action in the vocal cords themselves?

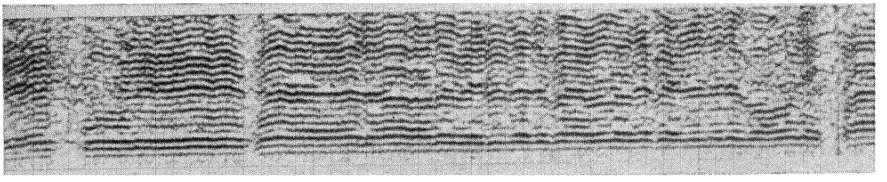


Fig. 21 — Pattern of the chant of a tobacco auctioneer.

Tobacco auctioneer's speech

Extremely rapid yodel-like sounds that have been publicized as the chant of the tobacco auctioneer are illustrated by Fig. 21. Apparently in a sample of this length, using this type of portrayal, not much is evident except small, sudden, and frequent variations in pitch. Closer examination reveals some recognizable sound formations between long intervals of illegible pattern. The resonance shifts identifying a sound pattern may, for example, be seen near the right end of this illustration.

Exaggerated vocal inflection

Extreme inflectional shifts in vocal pitch are used sometimes for emphasis or color. A sample of such pitch changes as produced by a radio comedienne is shown in the narrow band pattern of Fig. 22. The words are "*Sure, I argued with him!*" spoken rapidly. In this figure, the pitch may be observed to drop in a ratio of about three to one, within a small fraction of a second.

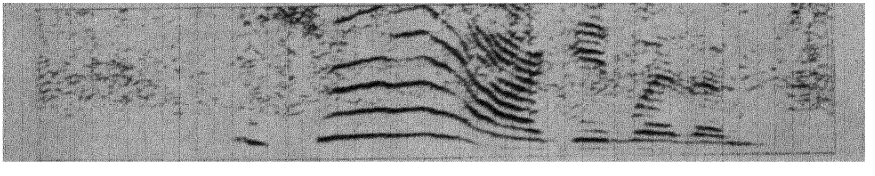


Fig. 22 — Pattern of an extreme inflectional shift.

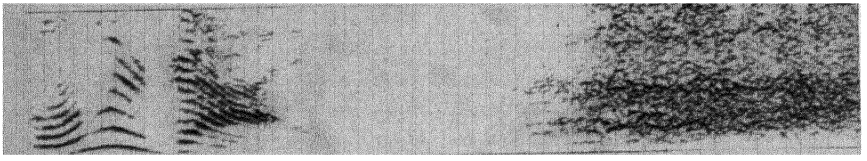


Fig. 23 — Pattern of audience reaction to a joke.

Audience reaction

Audience reaction in listening to a joke is illustrated by Fig. 23. The story finished with the words, "— *pull or drive!*" at the left, and three-quarters of a second afterwards the crowd laughed. Various voiced sounds or screeches seem to show against the mottled background of the laughing noises.

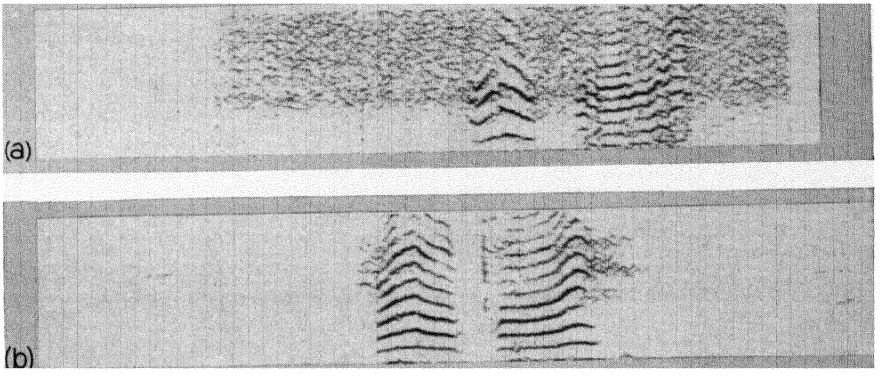


Fig. 24 — Pattern of a parrot saying "Hello, Polly": (a) parrot speech; (b) human attempt to imitate the parrot.

Parrot speech

In Fig. 24(a) are shown the words "*Hello, Polly!*" taken from a recording attributed to a parrot. An imitation of the recorded sound by a human being is shown in (b). There are curious yodel-like pitch breaks in the pattern of (a) that are not ordinarily present in the voices or speech of humans. A suggestion of resonance is also present in (a), starting at around 1,200 or 1,500 cycles and towards the end rising to possibly 2,500 cycles.

Need for Mutual Understanding

If the phonetician and engineer are to cooperate seriously in a development such as visible speech, they will have to become more cognizant of each other's problems and points of view than they have been in the past. Certainly, they will need to be well acquainted with both the problems of sound-to-sight translation and the problems of pattern form and interpretation. Apparatus that converts sounds to visible symbols is analogous to a mechanical ear capable of an endless variety of characteristics. If automatic transcription by visible speech methods is to be meaningful, the required characteristics will have to be specified; and to specify them, the phonetician will need an understanding of how they do, might, and should operate.

In the past, one of the major contributions of phonetics has been towards a universal form of phonetic symbolization of speech. It seems logical that this activity eventually might be broadened to include symbolization of all sounds in one standard pictographic language. To do justice to such a responsibility, and to understand fully the limitations and potentialities of various translation systems, the phonetician should know at least enough about electronics and circuitry in general to be conversant with the language and the basic problems of automatic transcription, especially if he is to associate and interrelate visible patterns, speech production, and the hearing process.

Similarly, the engineer concerned with the design and construction of translation systems needs to be equipped with at least a working knowledge of the fundamentals of speech, hearing, and phonetics. Certainly, his grasp of the requirements for sound patterns will be limited, if he is not aware of the problems, issues, vocabulary, and symbolization current in the various fields in which these patterns are to be used.

Developing in these separate fields a mutually intelligible language and a basic comprehension of the issues involved will not be a simple task, but the result in efficiency and flexibility of sound-to-sight translation will be more than worth the effort.

Chapter 15

Speech Correction Interests

This chapter is designed primarily to meet the interests of readers concerned with problems of speech correction and improvement, including the general field of organic and functional defects of voice or diction. We have discussed the deaf interests in a separate chapter because of the unique problems involved in adapting visible speech for the specific needs of the deaf and persons with severe auditory handicaps. Our concern here is with the potential use of visible speech as a supplement to the established aural, visual, and kinaesthetic methods used in correcting the speech of those with unimpaired hearing. As in the other chapters in Part Three, the treatment of the material is general and only suggestive, due to our restricted experience in applying visible speech to speech rehabilitation. The types of visible patterns needed for this purpose are likely to incorporate the requirements discussed in Chapters 13 and 14; and, since the limitations and possibilities of present translation equipment also have been presented in preceding chapters, it does not seem necessary to repeat either of these discussions here.

The work of the speech correctionist may be divided into several broad phases including: diagnosis, remedial therapy, records of progress and case histories, and periodic follow-up checks on proficiency. In addition, he frequently has the responsibility of training others in the field, and also he often undertakes research to increase his knowledge and to improve his techniques. In each of these categories, visible speech may play an important role.

Diagnosis

The speech correctionist is confronted with a special situation in each individual case. His first contact with a subject is analagous in many respects to that of a doctor meeting a new patient, for his initial task is to collect information as a basis for diagnosis. There is no doubt but that an experienced ear and eye can tell a great deal from the test situations that usually make up the diagnostic interview. It seems probable, however, that the techniques at present employed in diagnosis could be supplemented appreciably by the evidence contained in visible speech patterns. For the speech correctionist of the future, the sound spectrograms might serve in much the same way that motion picture X-rays of tissue movements serve the modern doctor.

By way of example, Fig. 1 shows a few patterns of common sound substitutions and omissions along with normal reference patterns. Each sound in question presents a rather striking difference in pattern and transition, but of course such simple variations in diction are quite apparent to the trained ear. Where, then, are the advantages

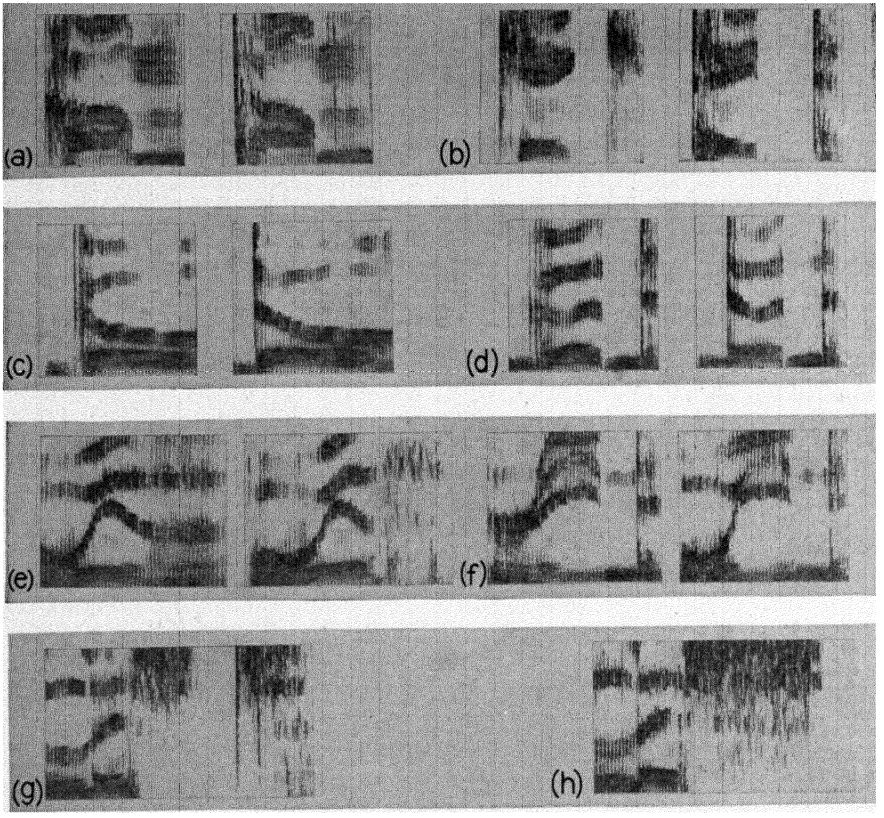


Fig. 1 — Illustrations of common sound substitutions and omissions: a, come (kʌm), ***tome*** (tʌm), ***b, cake*** (keɪk), ***tate*** (teɪt); ***c, go*** (goʊ), ***do*** (doʊ); ***d, good*** (gʊd), ***dood*** (dʊd); ***e, with*** (wɪð), ***with*** (wɪθ); ***f, read*** (rid), ***weed*** (wid); ***g, h, lists*** (lɪsts), ***lis*** (lɪs).

of visible speech? In each of these patterns, there is evidence of the nature of the articulatory movements, and of modulation shifts, that can be analyzed and interpreted on an objective basis. The value of such a technique is obvious to the experienced diagnostician and therapist, particularly in dealing with the complex speech disorders.

Remedial therapy

Drills and exercises with and without trained supervision are a large part of any remedial program. For the adult and child alike, visible speech provides a possibility of using visual reference patterns for comparison with defective speech patterns and for standards to supplement and reinforce the usual auditory — visual — kinaesthetic guides. In remedial drills, a combination of the direct translator with spectrographic reference patterns would probably be most desirable. However, either instrument certainly could be used to considerable advantage. Perhaps in the future, it may even be possible to transfer an increased share of the responsibility for his success to the trainee, and to reduce the amount of time expended by the supervising clinician on any one case.

Records of progress

It is customary for the case history file to contain records of progress of various kinds. Ideally, such files should be capable of interpretation and analysis by any trained clinician, and thus records should be kept in some permanent and standard form. In addition to phonograph records and phonetic notations, the visible speech patterns provide a permanent phonetic record that will not deteriorate with age, and that is economical of space. The advantages of objective analysis are obvious, and in most cases the actual techniques used may be quite simple. Some of the factors that are of interest to the correctionist and that can be computed readily or interpreted from the spectrograms are listed briefly for consideration: fundamental frequency or pitch; pitch range; harmonics missing or weak; resonances missing, weak, or emphasized; resonance bar positions; relationship of resonance shift to time; melodic pattern; inflectional changes; rate of articulation; duration of pause; precision of articulatory movement; completeness and speed of articulatory closure; and shifts in type of modulation. Many other variables will be suggested to the experienced clinician, and with pattern improvements the list may perhaps be expanded even further.

At the present time, the visible patterns have been reproduced by a playback machine only in experimental fashion. Perfection of that phase of the development would add immeasurably to the value of these patterns for all of the various fields of interest, and particularly in regard to the case history file.

Training clinicians

Possibilities of using visible speech patterns as a basis for reference libraries of normal speech and defective speech are of special interest to those concerned with instructing others in the field. Interchange and duplication of such materials would provide a large body of source materials for instruction in the basic principles of speech correction and pathology; in diagnosis; and in devising and evaluating clinical

therapy and tools. It may be of value also in testing and in making surveys, and in checking on students' comprehension of all of these factors. Duplication of the patterns is relatively inexpensive and should encourage interchange among speech correction centers throughout the world, thus enriching the experience of specialists, teachers, and students. It seems logical to suppose that visible patterns will aid in the teaching of physiological, comparative, and applied phonetics, and certainly these are basic subjects for the correctionist. These are important possibilities; but perhaps of utmost importance is the potential use of visible speech in helping to objectify and evaluate the concepts and methods of speech education and re-education.

Research

Of immediate interest, are the prospects for using speech spectrograms for research in the study of the speech characteristics of various types of speech defectives. The present patterns provide a source for information concerning speech and speech production that heretofore has been unattainable in practical form, and they offer unlimited potentialities of extending present knowledge in the different phases of speech correction.

It is, naturally, impossible in a text of this type to attempt to illustrate more than a few of the many variations in speech that are of special interest to the speech correctionist and pathologist. Any representative spectrogram of a speech defect or disorder may serve all of the major purposes discussed in this chapter. Certain modifications in recording and reproducing technique may be used to emphasize the variables of particular interest; but, in general, the broad- and narrow-band spectrograms should present a satisfactory starting point for diagnosis, permanent records, clinical drill, follow-up, and teacher-training. Spectrograms of speech variations included here were reproduced from phonograph recordings;* and consequently, in addition to speech variables, they portray some random noise and some frequency distortion. Direct recordings would have been desirable, and simultaneous spectrographic and phonograph recordings would have been ideal, but for our immediate purpose quite impractical. For each speech variation shown, a normal reference pattern of the same words is included to give some basis for comparison. All these reference patterns were spoken by an adult male with normal speech and hearing. The individual cases are classified in rather broad general categories since a detailed discussion of diagnostic information seemed undesirable for this presentation.

* Records obtained by the courtesy of the University of Michigan Speech Clinic, Ann Arbor, Michigan.

Spectrograms Illustrating Certain Speech Defects

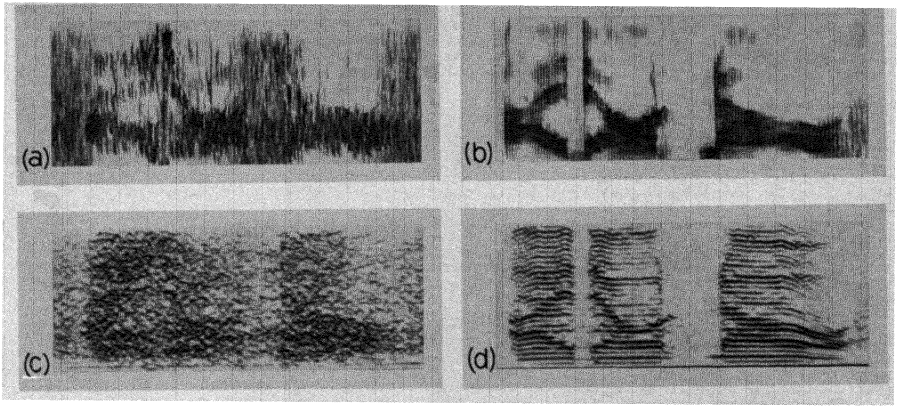


Fig. 2 — Broad and narrow band patterns of the words, "I don't know", spoken by an adult female with aphonia: a, aphonia; b, reference pattern; c, aphonia; d, reference pattern.

Aphonia (Fig. 2)

The aspirate nature of aphonic speech is illustrated by the partial disintegration of the resonance bars. See Fig. 2, for example. Some transitions are still apparent, but the strong frictional modulation and the excess breath combine to mask or distort many of the sounds and influences.

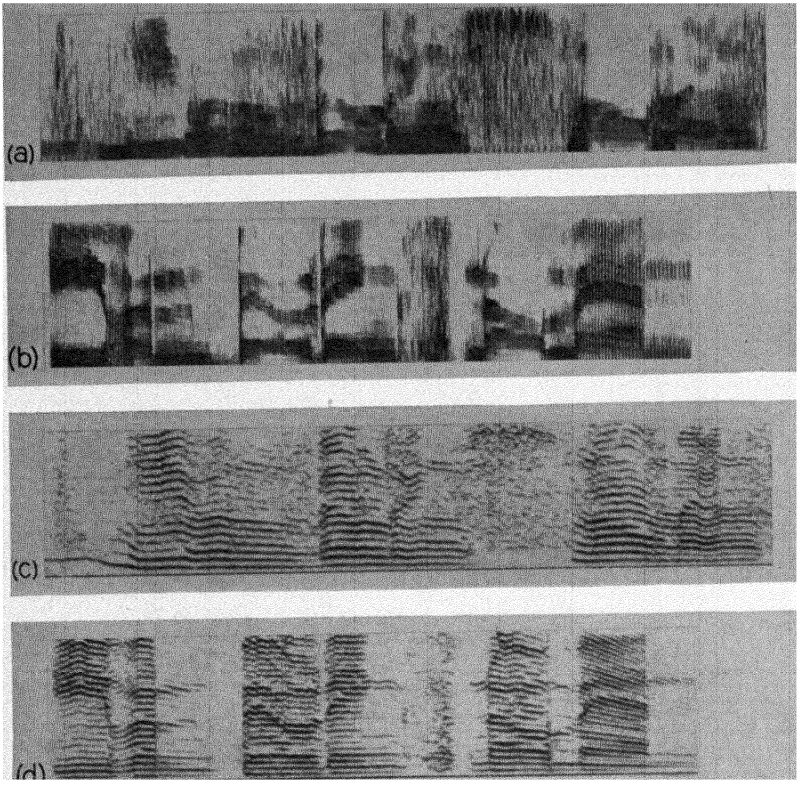


Fig. 3 — Broad and narrow band patterns of the words, "Even building snow men", spoken by an adult male with dysphonia: a, dysphonia; b, reference; c, dysphonia; d, reference.

Dysphonia (Fig. 3)

Weakness in the upper harmonics is evident in (a) and (c) of Fig. 3. It is interesting to see the gradual building up of the harmonic structure as shown at the beginning of 3(c). Stop gaps are not evident in the dysphonic patterns; this is due partially to continuation of resonation, as is evident in the narrow-band pattern. Some of the frica-tive fill may be due to aspirate quality, but a good deal of it probably results from record noise.

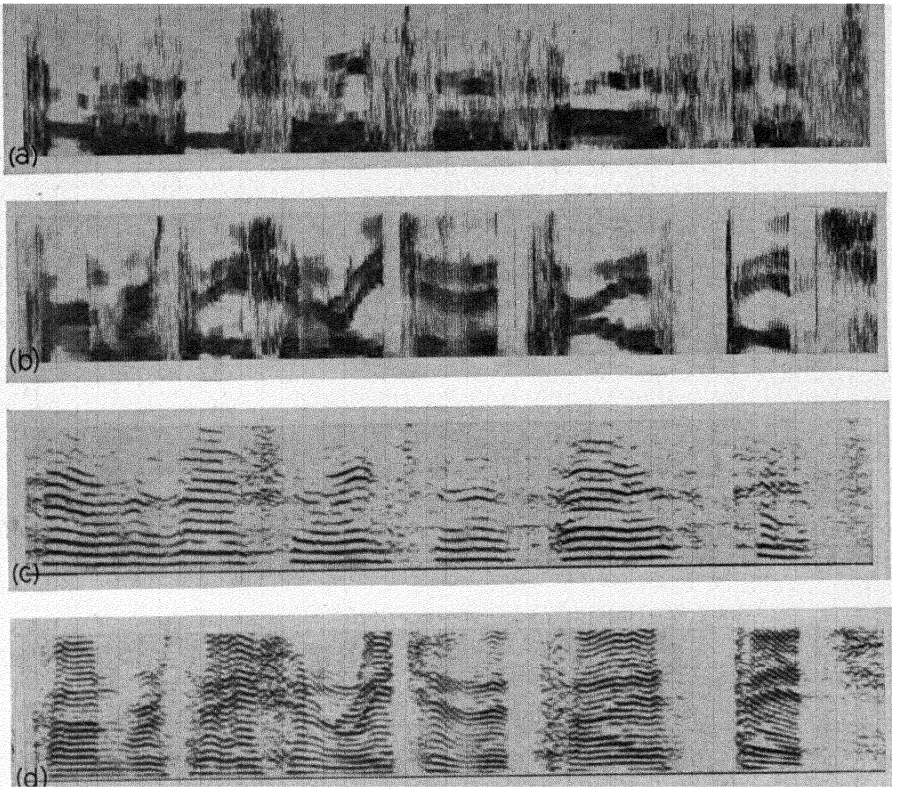


Fig. 4 — Broad and narrow band patterns of the words, "Conversation reaches its highest pitch", spoken by an adult female with dysphonia: a, dysphonia; b, reference; c, dysphonia; d, reference.

Dysphonia (Fig. 4)

Resonance and harmonic discontinuity are quite evident in (a) and (c) of Fig. 4. In several of the vowels, the upper resonance bars are partially or completely absent. Again the portrayal of aspirate quality is not readily separable from that of record noise. Many of the expected influences and transitions are not present, and it is difficult to identify syllable units and sounds. Only a slight inflectional shift may be observed in the patterns of dysphonic speech.

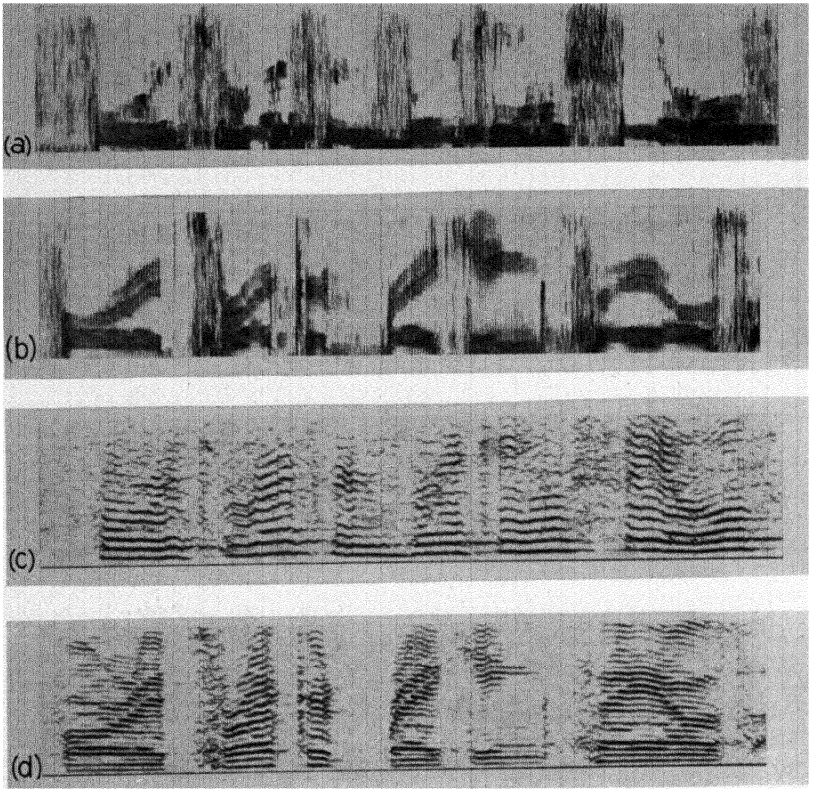


Fig. 5 — Broad and narrow band patterns of the words, "For its record breaking sales", spoken by an adult female with dysphonia: a, dysphonia; b, reference; c, dysphonia; d, reference.

Dysphonia (Fig. 5)

Dysphonic speech believed to be due to partial paralysis of the vocal cords is portrayed in (a) and (c) of Fig. 5. This was a quiet record, and a good part of the fricative fill overlaying the pattern may be attributed to the defective quality of the speech. Resonances and harmonics for the various vowels of this dysphonic speaker are quite similar in appearance, especially at the high-frequency end, to those shown in Figs. 3 and 4. It is difficult to follow the resonance areas in (c), or to identify the sound units and syllables in (a).

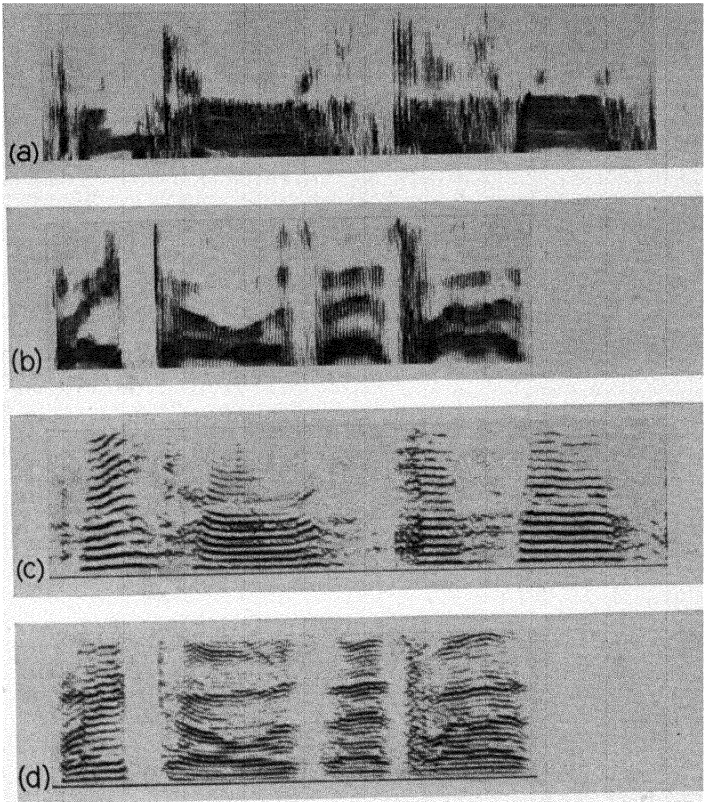


Fig. 6 — Broad and narrow band patterns of the words, "Plato's as well as", spoken by an adult female with cleft palate speech: a, cleft palate; b. reference; c, cleft palate; d, reference.

Cleft palate (Fig. 6)

Absence of the upper resonances is quite striking in (a) of Fig. 6. The lack of transitional influence and inflectional shift is pronounced in (a) and (c), and is similar to the patterns of speech of the deaf shown in Chapter 13. Strong resonance areas at the low-frequency end also may be compared with the patterns of nasal speech in Chapter 14. It is as difficult to identify the individual sounds visually as it is aurally. In many instances, the central neutral vowel has been substituted for the usual sound. An attempt was made by the speaker to produce a closure for the stops, resulting in a stop spike just before the t explosion in the word *Plato's*, although resonances do continue through the stop gap.

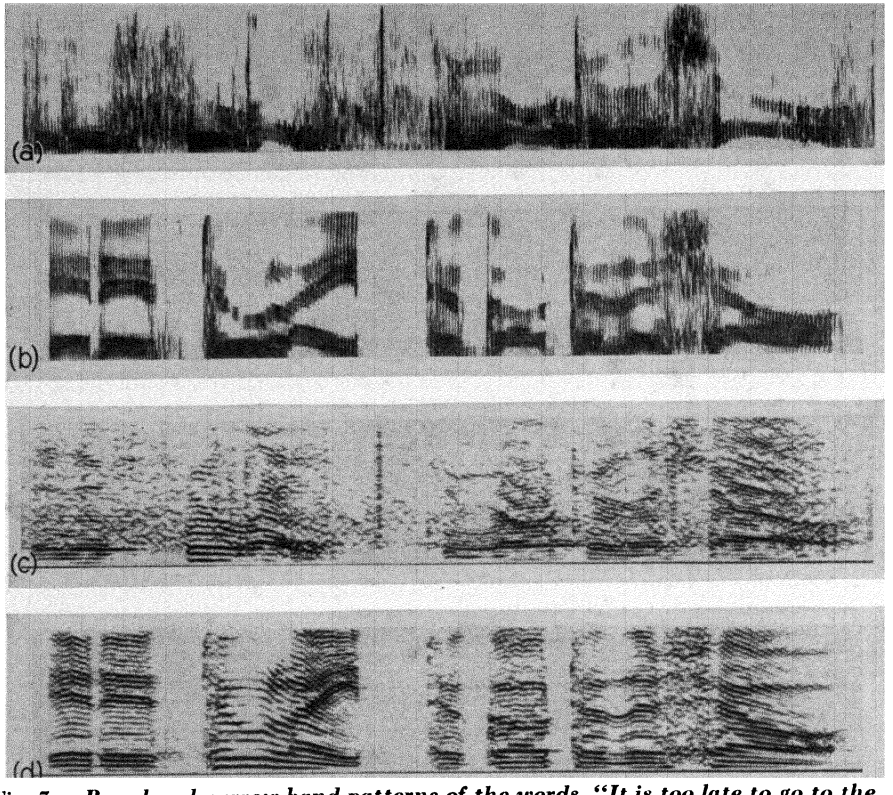


Fig. 7 — Broad and narrow band patterns of the words, "It is too late to go to the show", spoken by an adult male with cleft palate speech: a, cleft palate; b, reference; c, cleft palate; d, reference.

Cleft palate (Fig. 7)

More transitional influence may be seen in (a) and (c) of Fig. 7 than in the cleft palate samples in Fig. 6. Although the upper resonances are not yet sufficiently strong, the low-frequency end of both (a) and (c) is quite comparable to the reference patterns.

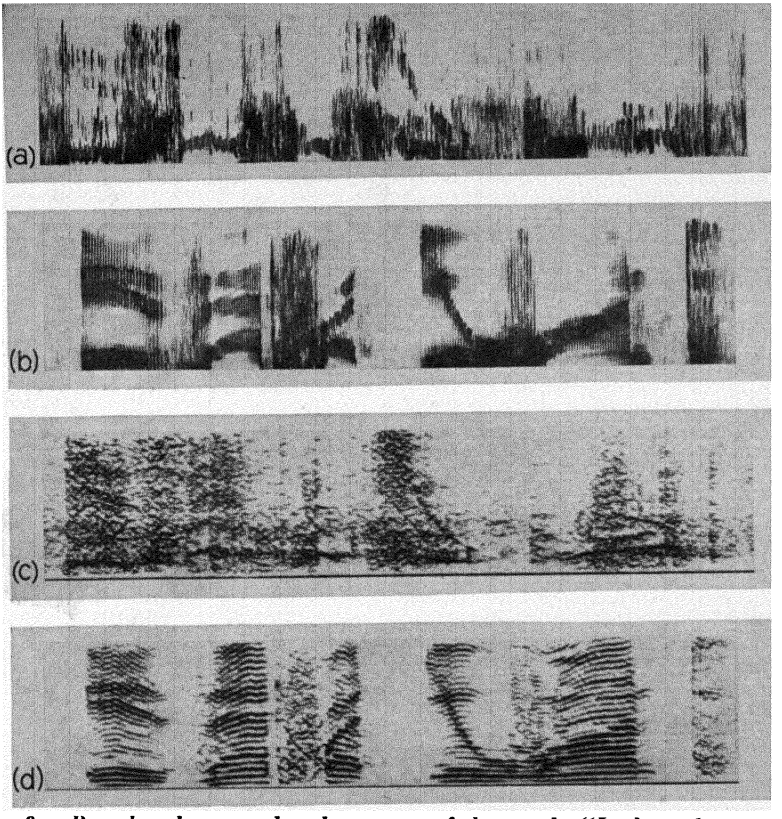


Fig. 8 — Broad and narrow band patterns of the words, "Is that what you want", spoken by an adult male using esophageal speech: a, esophageal; b, reference; c, esophageal; d, reference.

Esophageal speech (Fig. 8)

The narrow-band pattern of esophageal speech is especially interesting due to the broken and irregular nature of the harmonics. It is extremely difficult to follow the course of an individual harmonic through any portion longer than a syllabic unit. The irregular, slow, vibration rate may be seen in the individual beats, especially at the left end of (a). It is not possible to identify the visible patterns except for the "*is*" and the "*you*", and a stop spike for the t in "*want*" at the right end of the pattern. Upper resonance bars are weak or missing, and only one of the transitions, notably that in "*you*", is evident. Prolongation of individual sounds and of intervals between syllables is shown clearly. An added intermittent resonance area directly above the

first bar may be observed in both (a) and (c). Undoubtedly, the combination of cavity modulation with a vibratory source other than the vocal cords as in esophageal or artificial larynx speech should prove an interesting field for visible speech research.

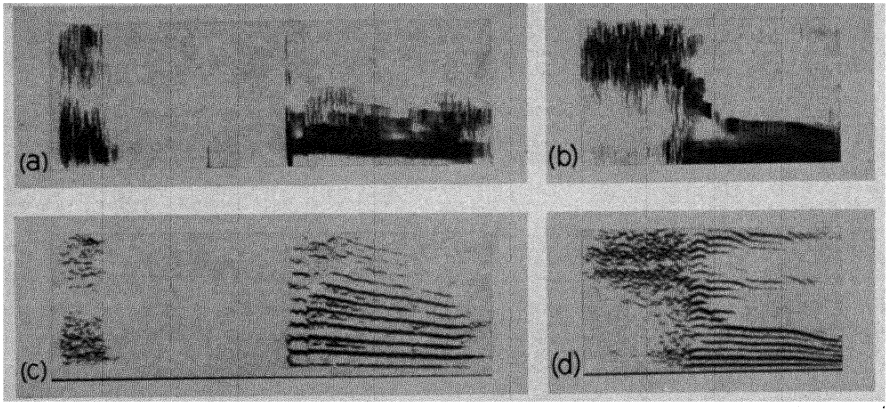


Fig. 9 — Broad and narrow band patterns of the word "shoe" spoken by a young boy with dyslalia: a, dyslalia; b, reference; c, dyslalia; d, reference.

Dyslalia (Fig. 9)

Separation of sound elements is evident in (a) and (c). Neither the transitional movements between these sounds nor the sounds themselves are produced correctly by this youngster. The fricative fill for the ʃ (*shoe*) appears darker towards the baseline rather than at the top of the pattern; and although some slight influence is observable on the vowel, it is not the usual influence of the ʃ, nor do the vowel resonances occupy their customary positions.

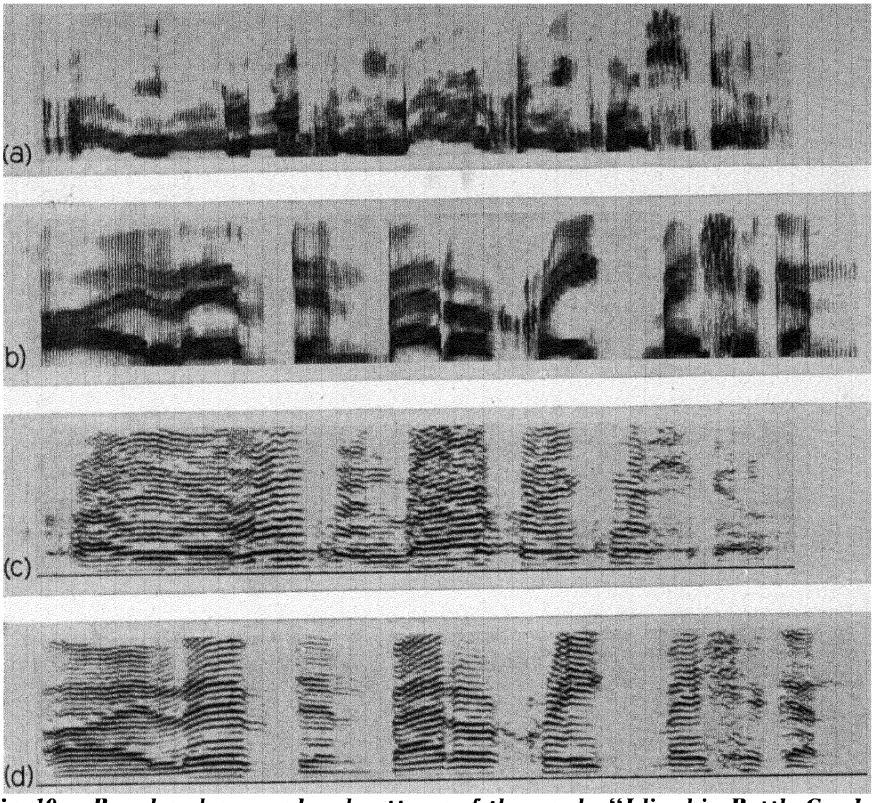


Fig. 10 — Broad and narrow band patterns of the words, "I lived in Battle Creek, Michigan", spoken by an adult male with dyslalia and some dysphemia: a, dyslalia; b, reference; c, dyslalia; d, reference.

Dyslalia and dysphemia (Fig. 10)

Influences and transitions are incomplete in many instances in patterns (a) and (c). Note the use of a back l with the front vowel in "*lived*", the indefinite nature of the stop gaps, and the substitution used for the t blend in "*Battle*". It is almost impossible to follow the resonance areas in Fig. 10(c) except for the last two syllables. Some dysphonia is evident in Fig. 10(c). Irregularity and considerable tremor may be observed in the harmonic structure.

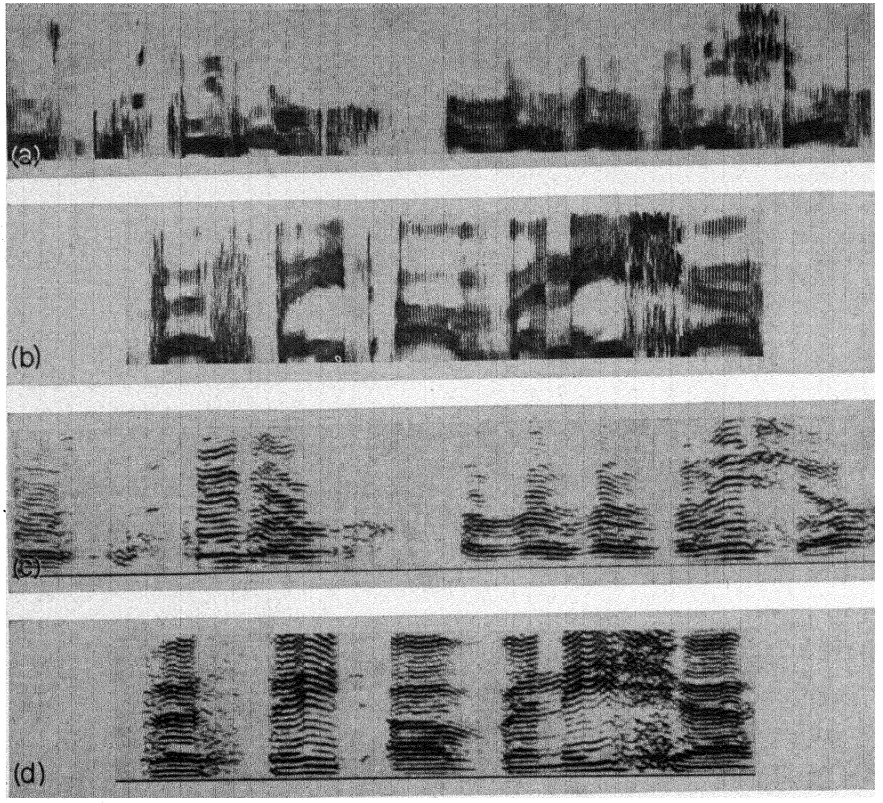


Fig. 11 — Broad and narrow band patterns of the words, "the state of Venitia", spoken by an adult male with dysphemia and some dyslalia: a, dysphemia; b, reference; c, dysphemia; d, reference.

Dysphemia and dyslalia (Fig. 11)

Syllabic repetition of the first syllable in "*Venitia*" is shown at about the middle to right portion of patterns (a) and (c). In general, the upper resonance areas in Fig. 11(c) are not clearly defined, nor are the influences and transitions readily identifiable. The reference pattern does not have an elision in "*state of*" similar to that in (a) and (c) which makes the initial parts of the patterns appear to be more different than they are. Irregularities may be observed in the harmonic structure in (c).

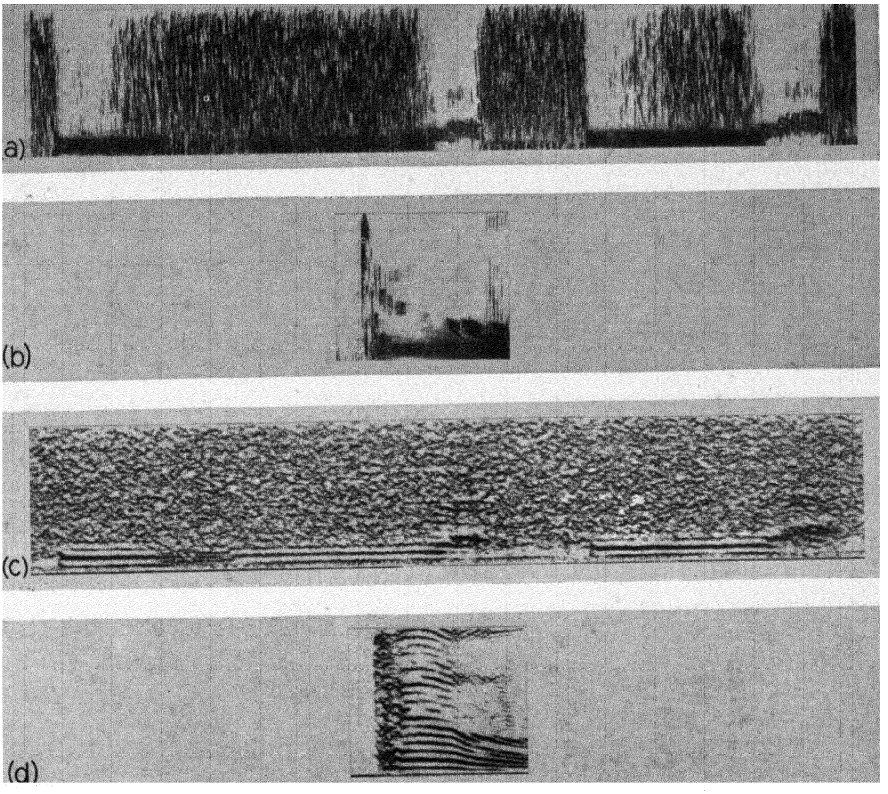


Fig. 12 — Broad and narrow band patterns of the word "to", spoken by an adult female with dysphemia: a, dysphemia; b, reference; c, dysphemia; d, reference.

Dysphemia (Fig. 12)

Rather severe dysphemia manifested by this individual is shown interestingly in the attempts to say the word "*to*". Patterns (a) and (c) read "*to*" at the left end, then ɑ (*father*), then "*to*", then ɑ (*father*). Some of the overlaying fricative striation is record noise, which unfortunately is not separable from breath. Unusual harmonic structure appears momentarily at two points in (c), apparently occurring during the attempt to terminate voicing for the initial "*to*", and the final ɑ. A somewhat similar pattern also may be observed during part of the first ɑ. Voicing was terminated just before the rather prolonged stop for the second t, and then was continued through the last two syllables. Absence of the upper resonance bars in (a) and (c) may be due partially to high noise level on the record and the consequent need to decrease the amount of amplification used in reproducing the sample.

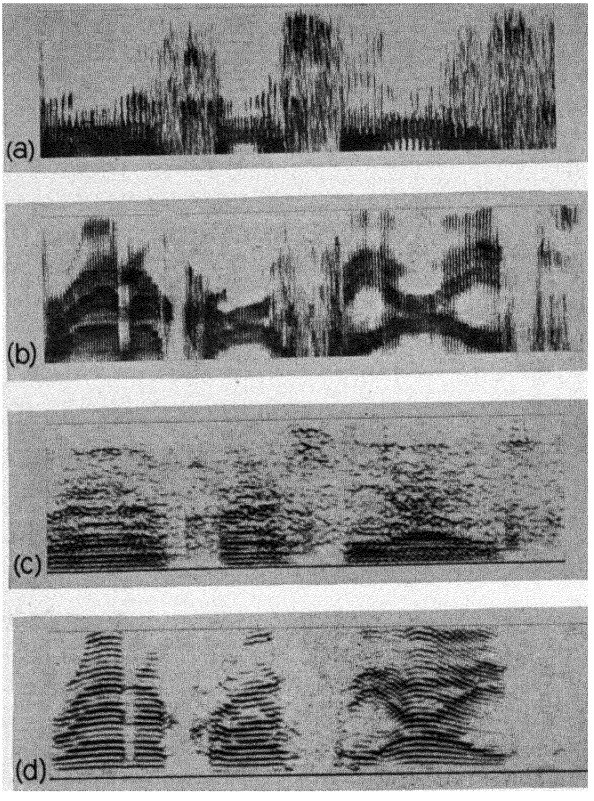


Fig. 13 — Broad and narrow band patterns of the words, "ran across the ice", spoken by an adult male with dysarthria: a, dysarthria; b, reference; c, dysarthria; d, reference.

Dysarthria (Fig. 13)

Patterns (a) and (c), manifesting dysarthria, were produced by an individual with cerebral palsy and Parkinsonian syndrome. In pattern (a), there is an absence of the upper resonance bars and an unusual shift in fundamental frequency or voice pitch, shown in the increased separation between adjacent vertical striations. Transitional influences are slight compared with the reference pattern, and also the individual sound units are prolonged. An interesting tremor is shown in the harmonics in pattern (c). The harmonic structure appears to be rather irregular, especially in the upper frequencies. Perhaps the most striking point is the extremely low pitch shown towards the end of this sample.

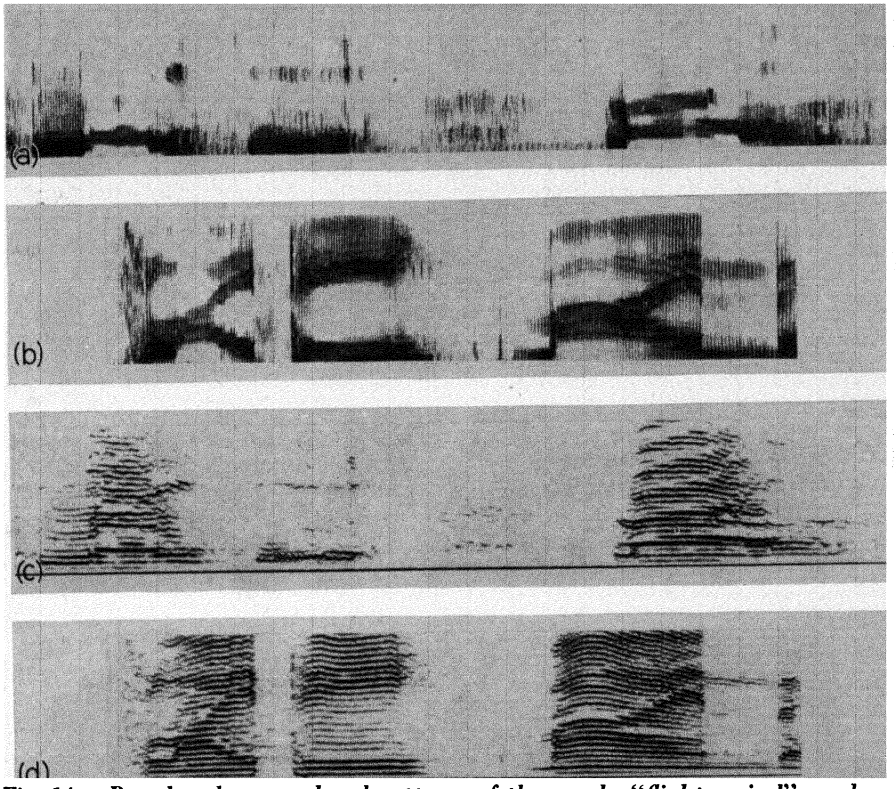


Fig. 14 — Broad and narrow band patterns of the words, "flighty mind", spoken by an adult male with dysarthria: a, dysarthria; b, reference; c, dysarthria; d, reference.

Dysarthria (Fig. 14)

An illustration of speech by another individual with cerebral palsy is shown in patterns (a) and (c). It is similar to the preceding sample in the absence of upper resonance bars, but the shift in fundamental frequency is not as marked as in Fig. 13(a) and (c). Some transitional influences are noticeable on the resonance bars present. In Fig. 14(c), some tremor may be observed in the harmonics, but it is most marked in the low-frequency end of the second syllable. Duration of individual words compares favorably with the stressed speech shown in the reference sample, but pauses between words in the sample of defective speech are approximately twice as long as in the normal reference pattern. It would be difficult to identify individual sounds or syllabic combinations in either Fig. 13 or 14, due to the unusual nature of the physiology involved and the resultant combinations of modulation.

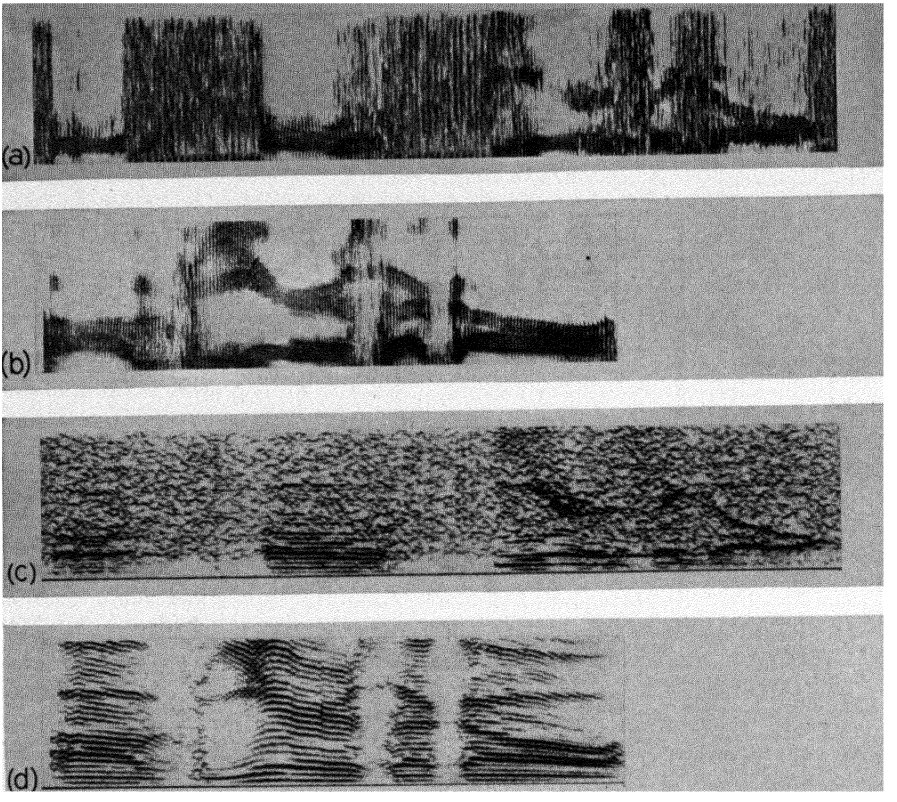


Fig. 15 — Broad and narrow band patterns of the words, "of years ago", spoken by an adult male with aphasia: a, aphasia; b, reference; c, aphasia; d, reference.

Aphasia (Fig. 15)

Patterns (a) and (c) illustrate the speech of an asphasic. This particular sample was selected because it shows the characteristic groping for words, evidenced by the repetition of the word "*of*" at the left end of the pattern. Further discussion of this pattern is inadvisable because of the interference of record noise, as seen in the fills between words. Unfortunately, there was no other less noisy record available.

Although the survey of speech defects and pathologies is brief, the spectrograms limited, and discussions necessarily cursory, it is hoped that they may indicate some of the potential applications of visible speech in the different areas of speech correction, including both teacher training and research.

Chapter 16
Foreign Language Interests

Foreign language applications of visible speech are perhaps evident in the discussion of phonetic organization in Part Two, and in the summary of Phonetic Interests in Chapter 15. However, these discussions were devoted primarily to English, and it has seemed desirable to add to them a few illustrations of sounds and sound combinations in other languages.

Samples of speech in various languages are shown in visible speech form on the following pages. The list is not complete, nor is it suggested that the samples are truly representative of the languages illustrated. Where possible, the speakers were chosen from the personnel of the Bell Telephone Laboratories. Selections were not critical, although most of the speakers were of foreign birth and acquainted with the language through regular use. No doubt an American accent is present in some of the patterns.

Recording procedure

Male voices were used in all cases. After arrangements had been made with individuals for visible speech recording, they were given the statement shown opposite the English sample on page 352, and were asked to translate it into the foreign language they would use. At the subsequent recording session, the speaker read his translation into a microphone; and the speech was recorded simultaneously on the magnetic tape of a long-sample spectrograph and on a phonograph record blank. For each case, then, there was a written statement of what was said, a phonograph record of the speech sounds for possible future reference, and a visible speech spectrogram.

In the following illustrations, the visible speech transcriptions are shown on pages to the left, and facing each on the right are the words of the foreign language as translated and written by the speaker. Each language sample was recorded as one long spectrogram that was later divided into sections to fit the page. Samples were cut to a standard length regardless of where divisions occurred, a procedure that considerably simplified the layout arrangements, but in some cases at the expense of pattern continuity.

Background of speakers

Following is a brief summary of the background of each speaker with his own classification of the language used.

Arabic — Born in Cairo and brought up in that area. Describes the language he speaks as Classic Arabic. Learned English when about 15 years of age.

Chinese — Born in Liaoning and brought up in that area. Had spoken English for about two years, and classifies his Chinese as "Mandarin".

Danish — Born in Copenhagen and brought up in that area. Learned English shortly before coming to the United States about 25 years ago. Classifies his speech as "middle-class Danish".

English — Born and brought up in New York City. Speech is classified general American.

French —Born in Paris and moved to Dieppe when about 8 years of age. Went to school in England when about 9 or 10 years old. Has spoken English since early childhood. Classifies his speech as "Parisienne — parents of Norman background".

Gaelic — Born in Tipperary and brought up in that area. Learned both Gaelic and English at an early age. Classifies his speech as "Western".

German — Born in Cologne and went to Metz when a few years old. Speech classified as "high German". Came to the United States about 20 years ago.

Icelandic — Born in Reykjavik and brought up in that area. Speech classified as "characteristic of southern part of Iceland". Started to learn English at an early age, and came to the United States about 30 years ago.

Italian — Born in Milan and went to Genoa when about 15. Classifies his speech as "that of Northern Italy". Talked English from an early age.

Norwegian — Born near Stavanger in the western part of Norway and brought up in that area. Started learning English in early grades at school, and came to the United States about 20 years ago. Classifies his speech as "book language — Western influence".

Ojibway (American Indian) —Born on Onondaga Reservation in New York State. Learned both Ojibway and English in early childhood. Considers his speech to be "typical of eastern woodlands people in the Great Lakes Region".

Russian — Born in Geneva, Switzerland, of Russian parents. Returned to Moscow at about one year of age and was brought up in that area. Has spoken English for about 20 years. Classifies his speech as "Great Russian or Central European Russian".

Spanish — Born in Manila and brought up in that area. Had spoken English from early childhood. Classifies his Spanish as "Castilian".

Swedish — Born in Helsingsborg, and came to the United States as a young man about 20 years ago. Talked English prior to arrival. Classifies his Swedish as "pure" rather than of the dialect spoken in Skane.

Yiddish— Born in Brooklyn, New York, and brought up in that area. Learned both Yiddish and English at an early age. Classifies his speech as "mainly Lithuanian Yiddish".

Comments

When first looking over illustrations of visible speech in different languages, there is apparently a tendency to expect that strange languages will show strange patterns. It may, therefore, be something of a surprise to find that there is no easily recognizable distinction between the patterns of Chinese speech and those of English. And no doubt the distinction would be even less apparent if individual characteristics

could be eliminated. Apparently the impression of strangeness one experiences when hearing a completely unfamiliar language is less a matter of basic difference in the sounds involved than in the way these sounds are used. The fact that an International Phonetic Alphabet is at all possible supports this statement.

While articulators are used in different ways in different languages we are, after all, built much alike regardless of language. Nature imposes certain limitations upon the articulatory mechanism in shapes, sizes, and rates of movement. These are common to all races, and languages are developed within these limitations.

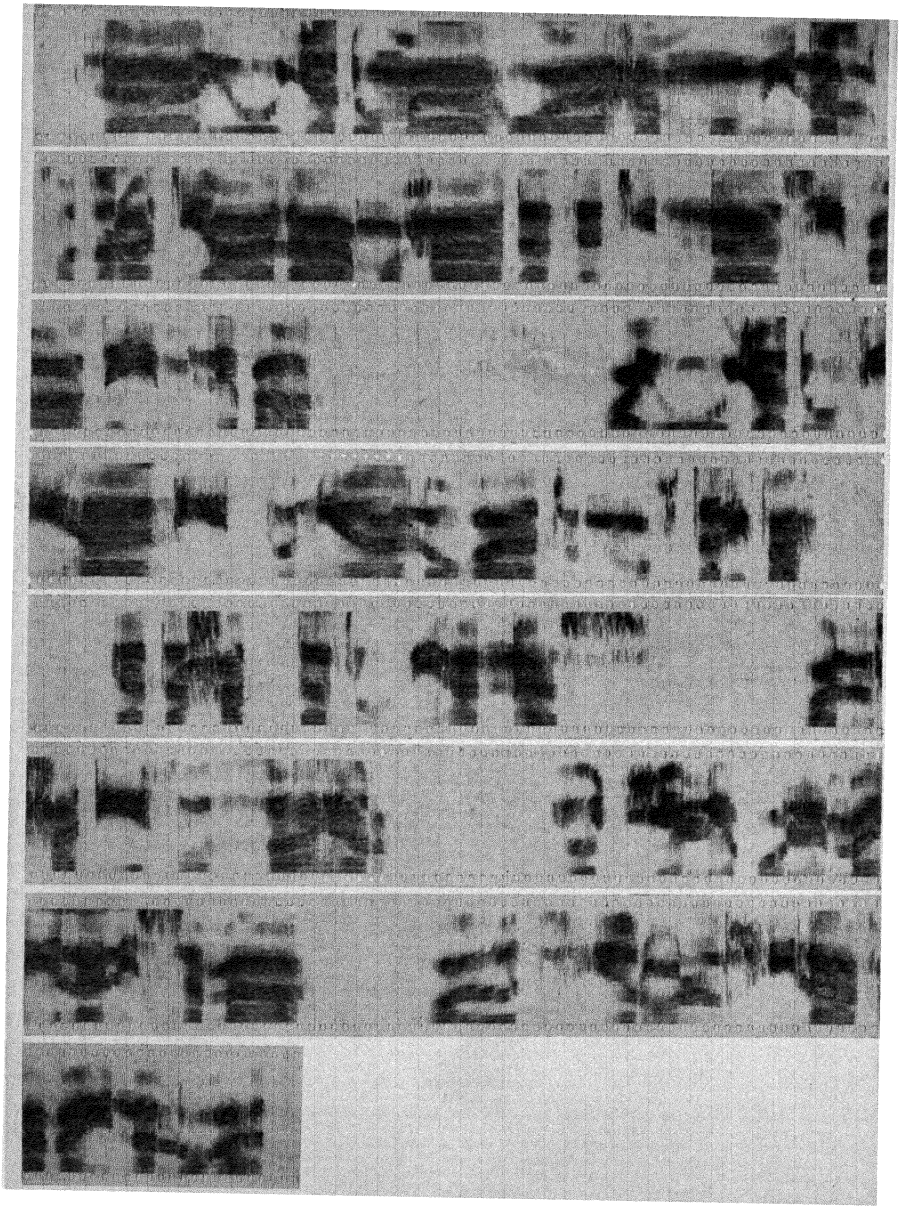
However, the student of language with a knowledge of visible speech will undoubtedly be able to see in these pictures details that the average observer would miss, movements that are associated with particular languages or language groups. Some of these movements become readily evident when the patterns are examined in detail. In the first half of the third line in the Spanish transcription there is the heavily rolled "*r*" of "*otro*". Another appears near the middle of the top line in the Icelandic pattern. There are several interesting bar shifts as, for example, in the Italian sample where there is a sharp descent and rise of Bar 4 in the word "*nulla*" at the beginning of the record. Again, at the end of the first quarter in the fifth line accompanying the word "*visione*", Bar 4 forms a downward loop that appears to overlap Bar 3 in an unusual way. At the left of the middle in the third line of the French transcription is an upward loop of Bar 2 followed by a trill, both associated with "*aux autres*".

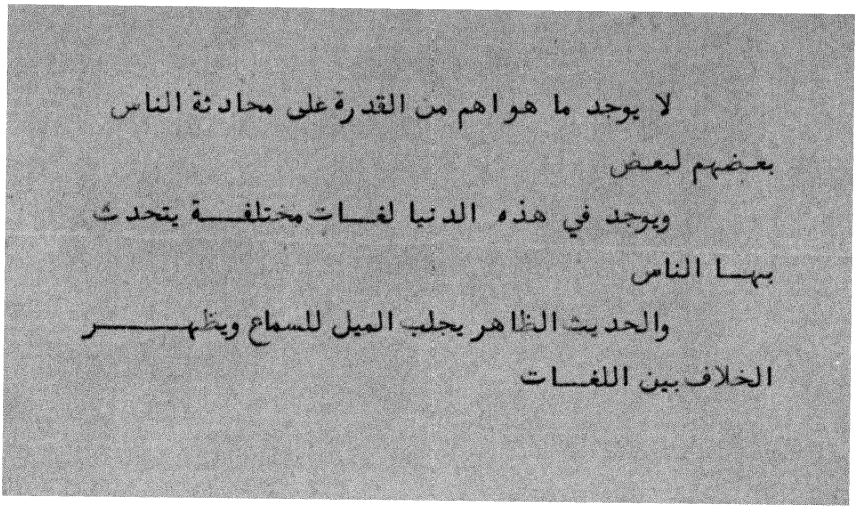
Change in resonance activity is apparently most extreme in the Arabic sample. During the first two lines, the bar movement seems almost completely restrained, but rapid variations do appear occasionally. There is considerable resonance movement in the Gaelic patterns, and it will be noted that the pattern of the word in the center of line five is repeated accurately at the end.

In the Russian sample, the most evident effect is the frequent termination in a down-glide of Bar 2. The German patterns show much less of Bar 1 than most of the others, but this may be an individual characteristic.

Of all the patterns, those of the Ojibway (American Indian) appear to be most distinctive, due apparently to an infrequent use of unvoiced sounds and almost continual shift of the resonance positions. The former is obvious upon comparison with the German sample, and the latter when considered in relation to the Arabic.

Other detailed differences, such as clarity of resonances, speech rates, and tendency to run words together or separate them completely, are, no doubt, associated with individual habits of speech rather than with the language. In studies of these patterns, the student of languages would become accustomed to neglecting such effects and concentrating upon fundamental forms and configurations.

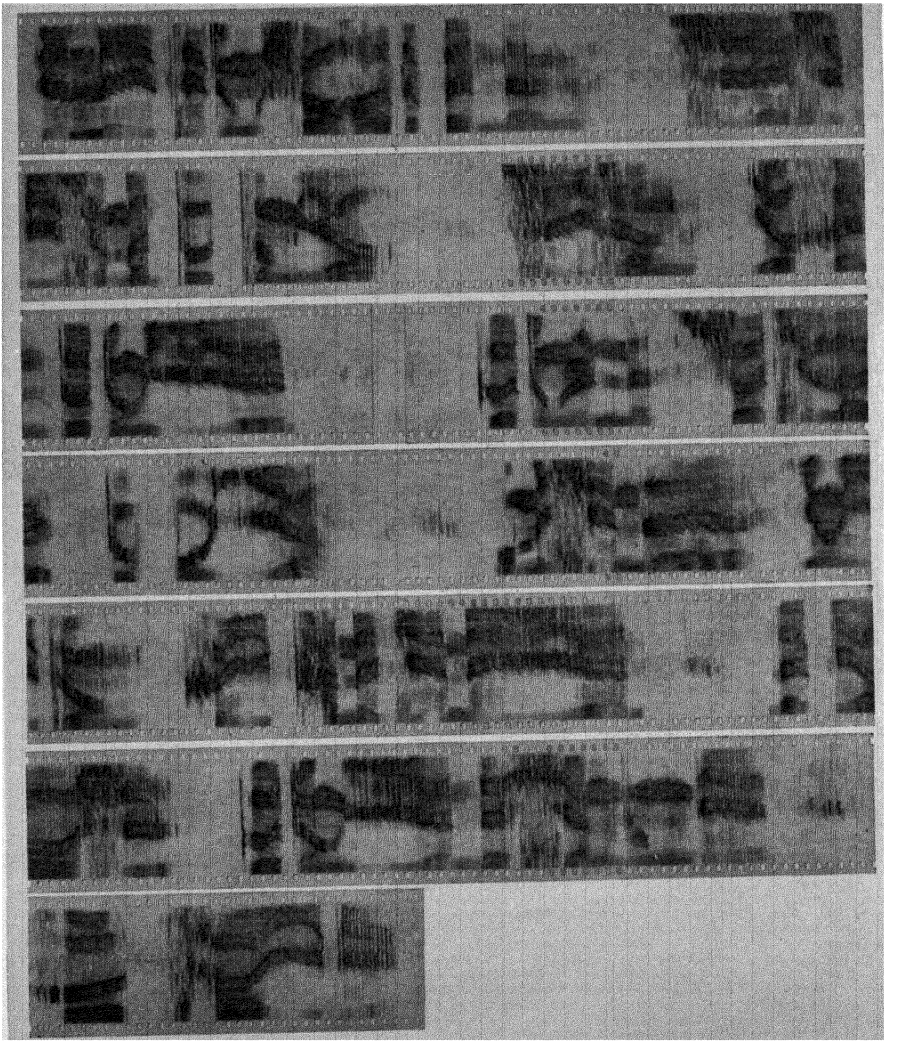


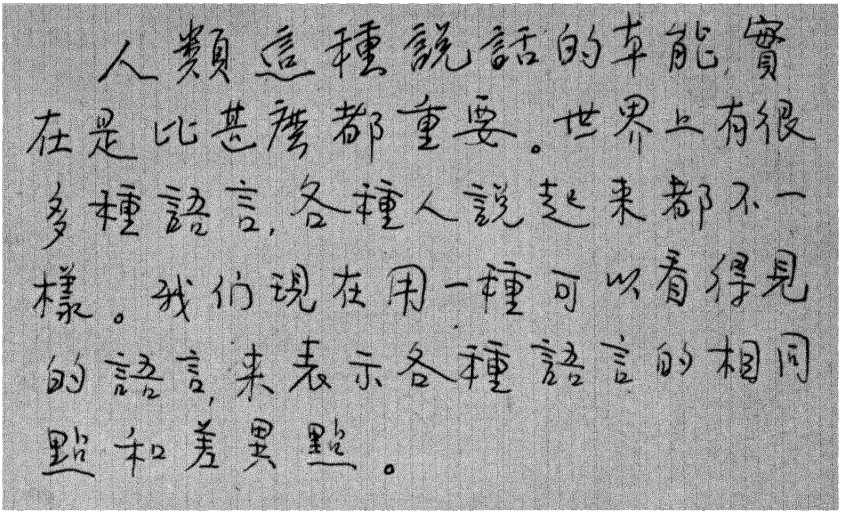


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Arabic

This speaker was born in Cairo and was brought up in that area. He describes the language he speaks as "Classic Arabic." He learned English when about fifteen years of age.

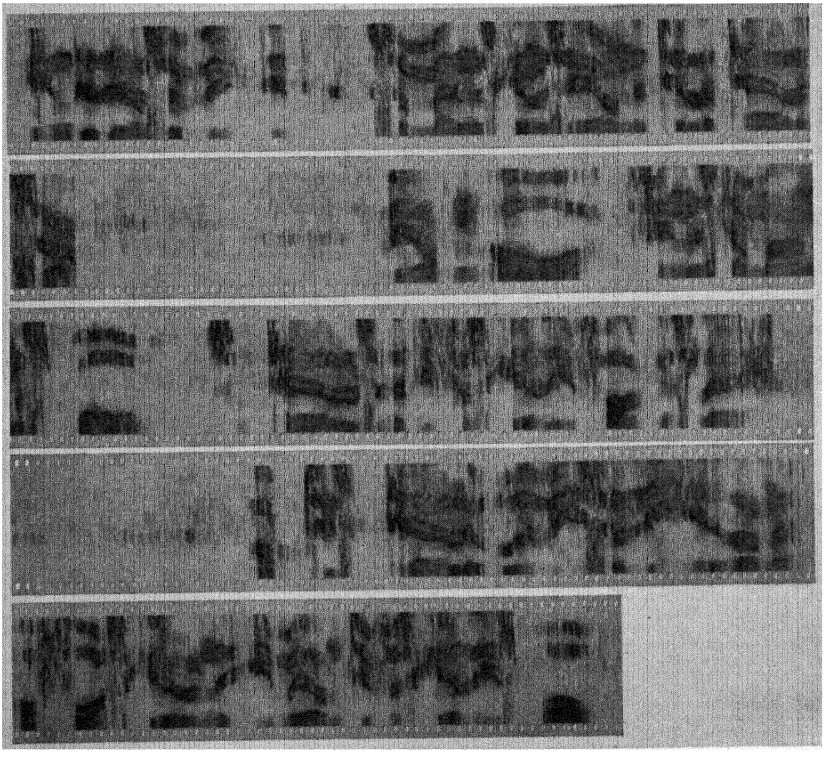


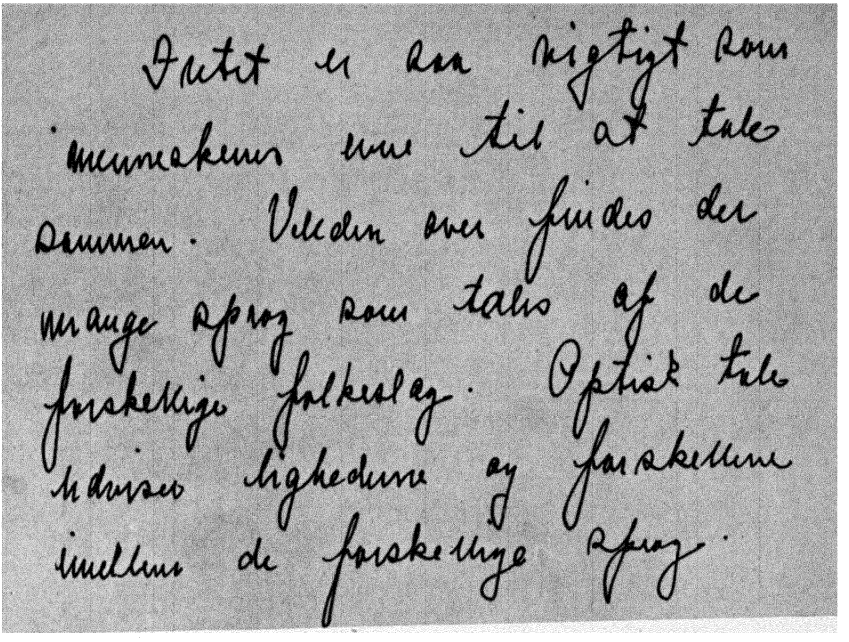

人類這種說話的本能實在是比甚麼都重要。世界上有很多種語言，各種人說起來都不一樣。我們現在用一種可以看得見的語言來表示各種語言的相同點和差異點。

The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Chinese

This speaker was born in Liaoning and was brought up in that area. He classifies his speech as "Mandarin." He had been speaking English for about two years at the time this sample was recorded.

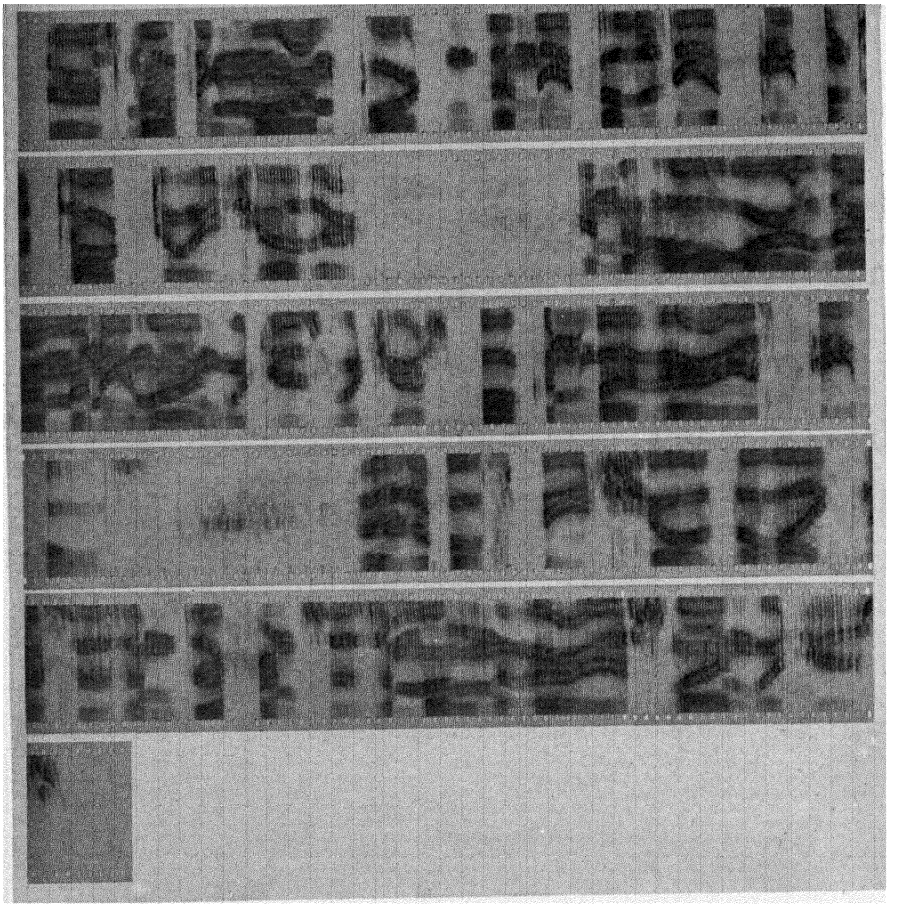


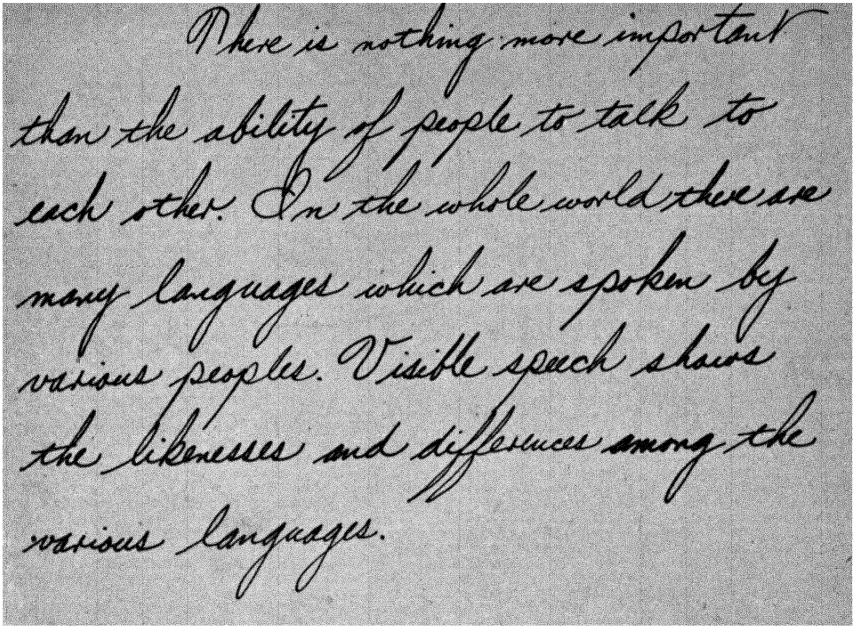


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Danish

This speaker was born in Copenhagen and was brought up in that area. He classifies his speech as "Middle-Class Danish." He learned English shortly before coming to the United States about twenty-five years ago.

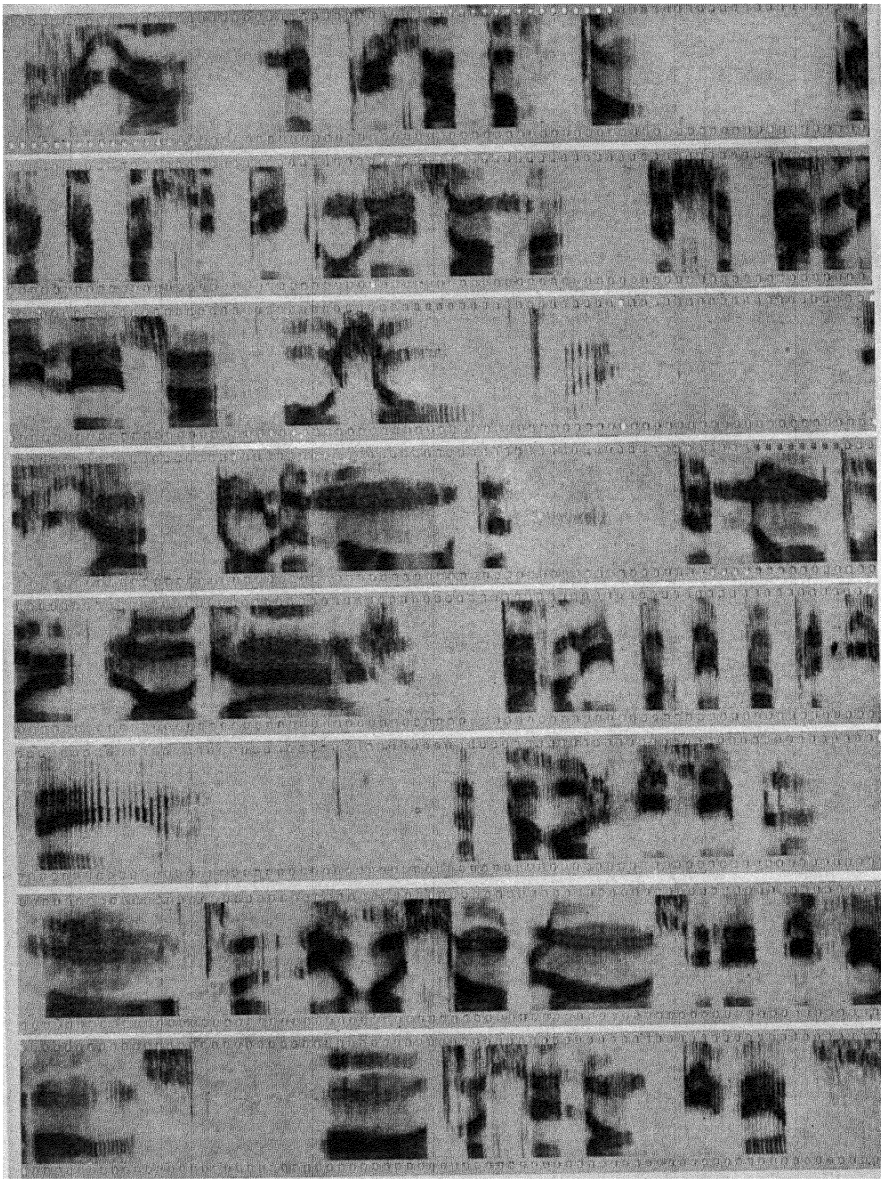


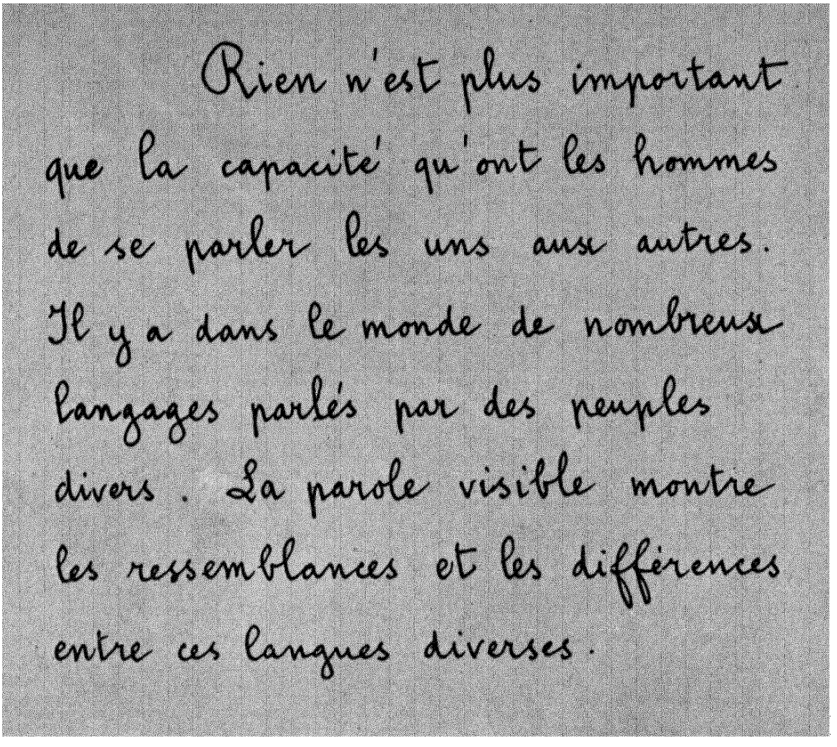


The English text used for the foreign language samples is shown above in the speaker's manuscript; on the facing page appears the visible speech transcription of his oral rendition.

English

This speaker was born and brought up in New York City. His speech is classified as "General American."

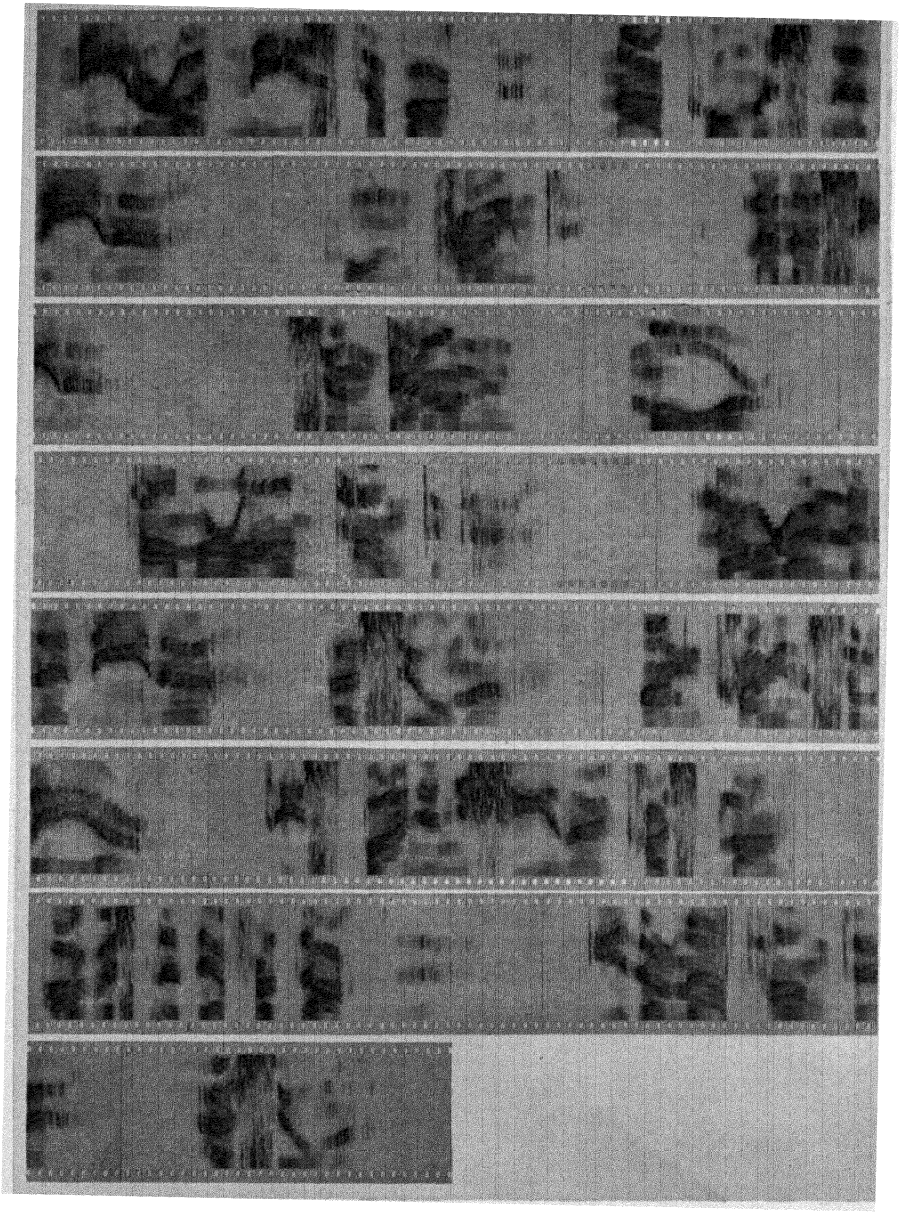


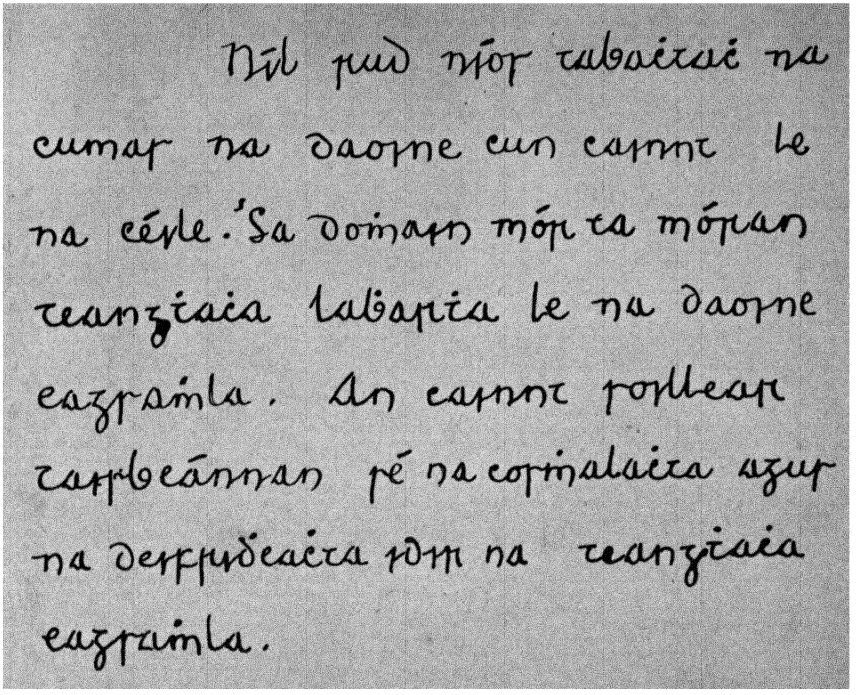


The speaker's translation of the English text is shown above in his manuscript: the visible speech transcription of his oral rendition appears on the facing page.

French

This speaker was born in Paris and moved to Dieppe when about eight years of age. He classifies his speech as "Parisienne — parents of Norman background." He went to school in England when about nine or ten years old, and has spoken English since early childhood.

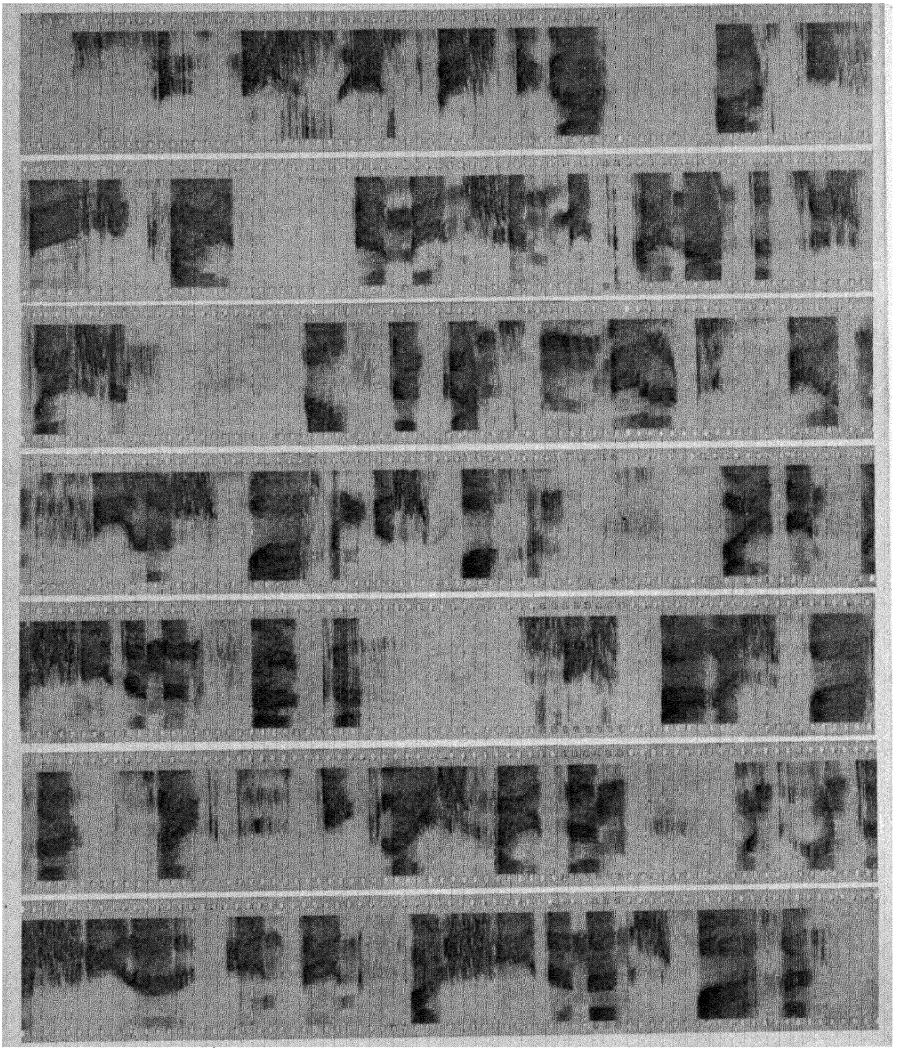


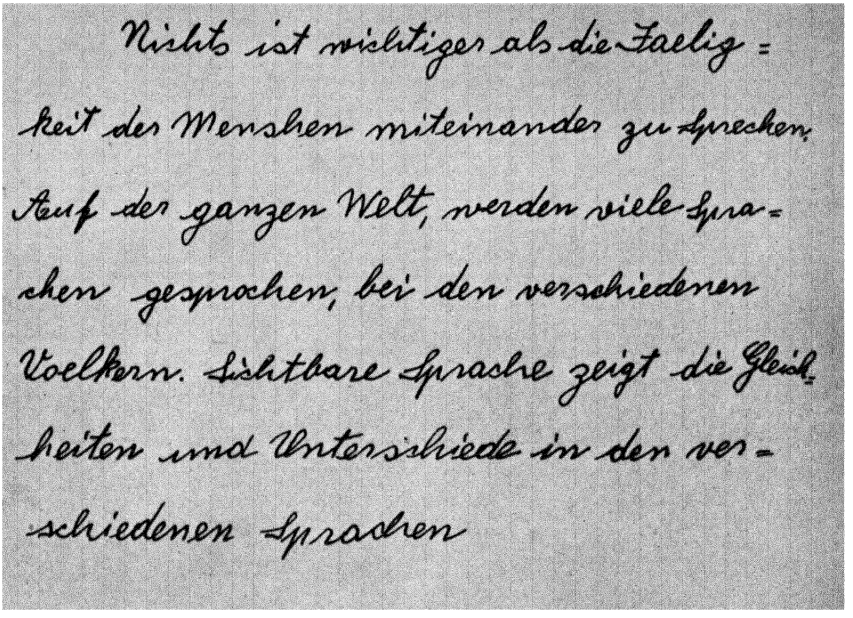


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Gaelic

This speaker was born in Tipperary and was brought up in that area. He classifies his Gaelic speech as "Western." He learned both Gaelic and English at an early age.

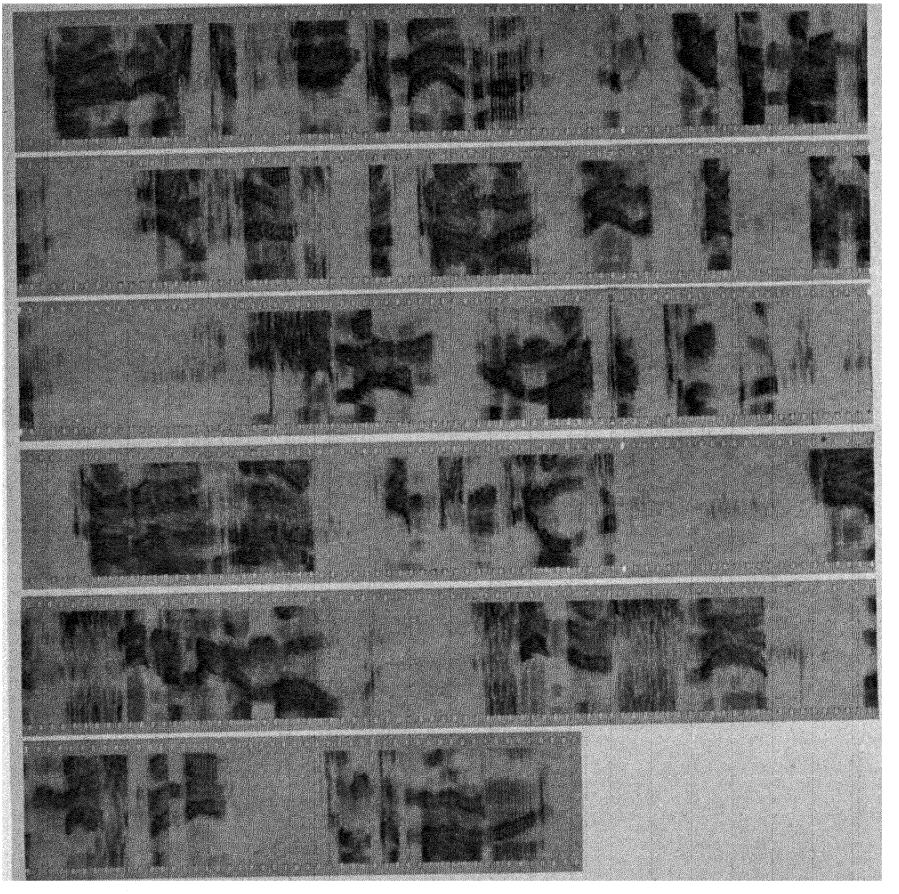


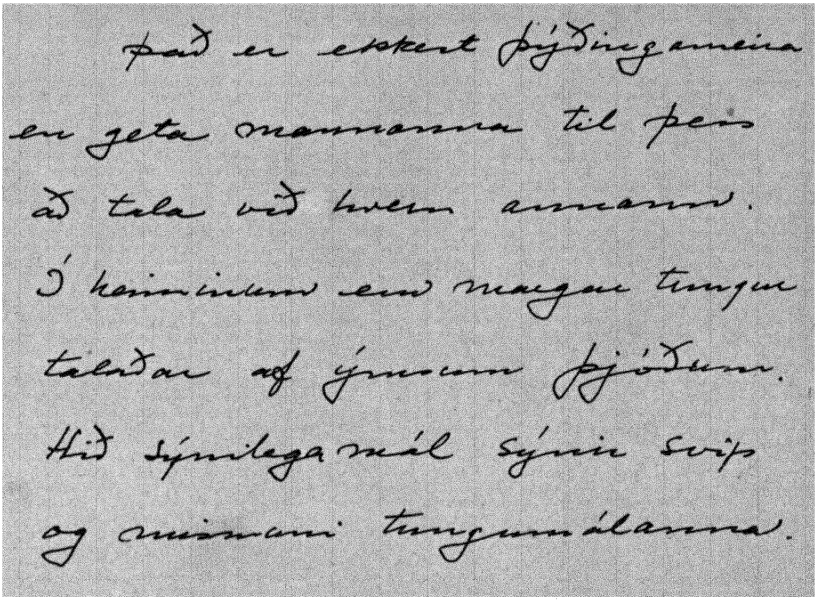

Nichts ist wichtiger als die Faehig-
keit der Menschen miteinander zu sprechen.
Auf der ganzen Welt, werden viele Spra-
chen gesprochen, bei den verschiedenen
Voelkern. Sichtbare Sprache zeigt die Gleich-
heiten und Unterschiede in den ver-
schiedenen Sprachen

The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

German

This speaker was born in Cologne and went to Metz when a few years old. His speech is classified as "High German." He came to the United States about twenty years ago.

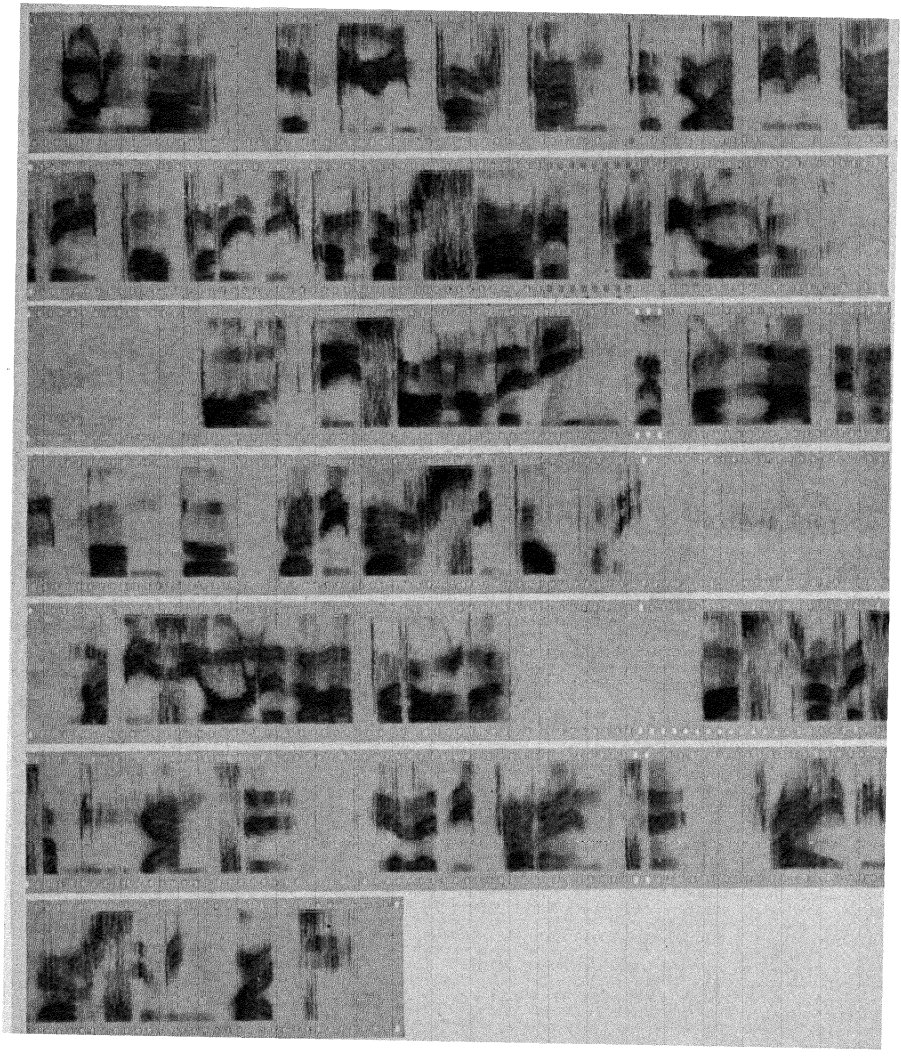


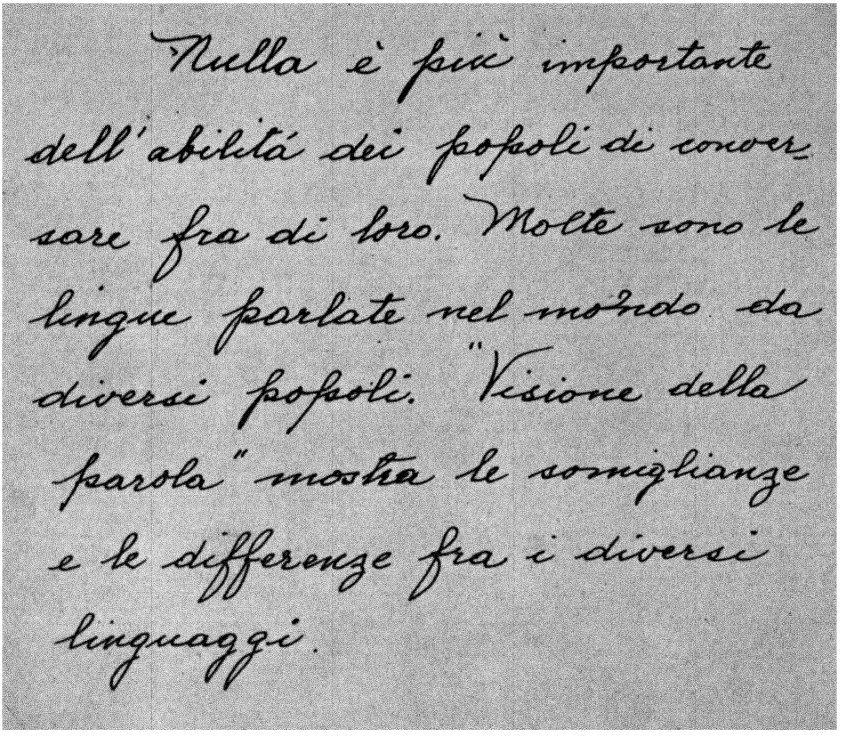


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Icelandic

This speaker was born in Reykjavik and was brought up in that area. His speech is classified as "characteristic of the southern part of Iceland". He started to learn English at an early age, and came to the United States about thirty years ago.

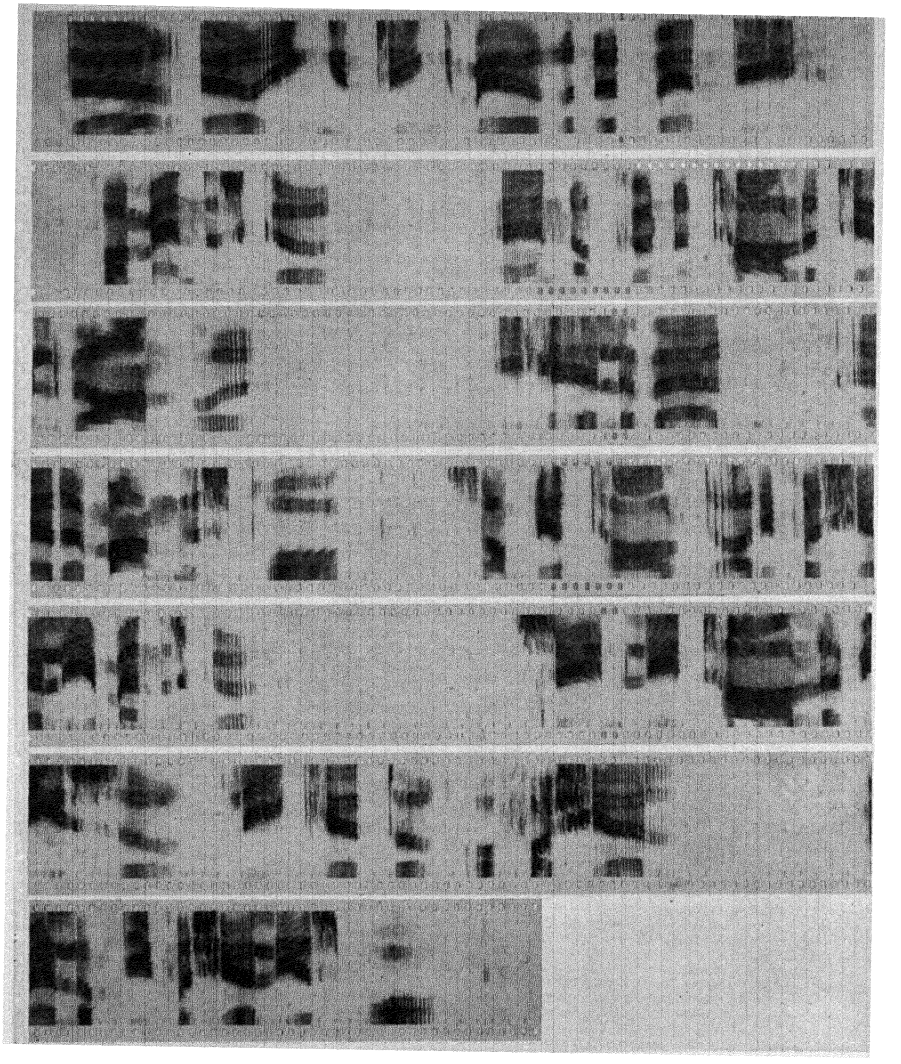


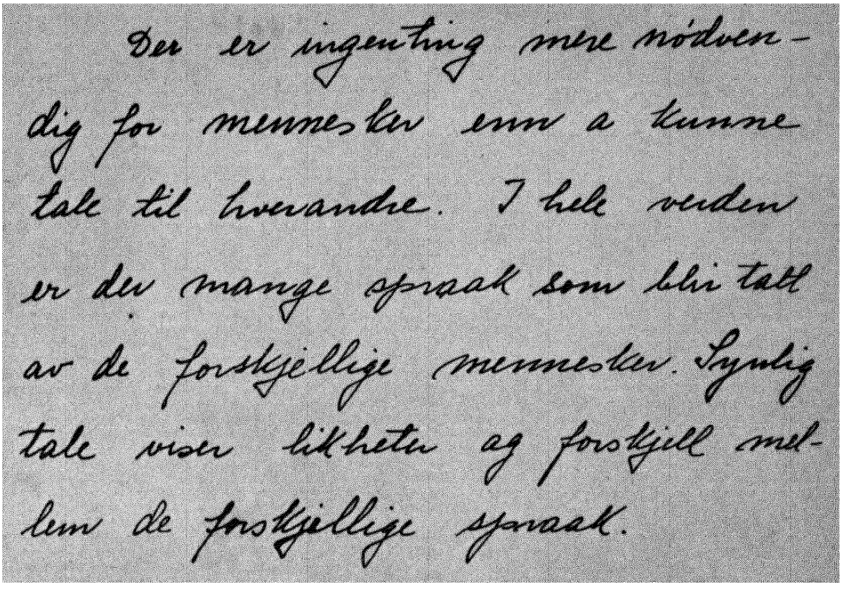

Nulla è più importante
dell'abilità dei popoli di conver-
sare fra di loro. Molte sono le
lingue parlate nel mondo da
diversi popoli. "Visione della
parola" mostra le somiglianze
e le differenze fra i diversi
linguaggi.

The speaker's translation of the English text is shown above in his manuscript: the visible speech transcription of his oral rendition appears on the facing page.

Italian

This speaker was born in Milan, and went to Genoa when about fifteen years old. He classifies his speech as "that of Northern Italy." He has talked English from an early age.

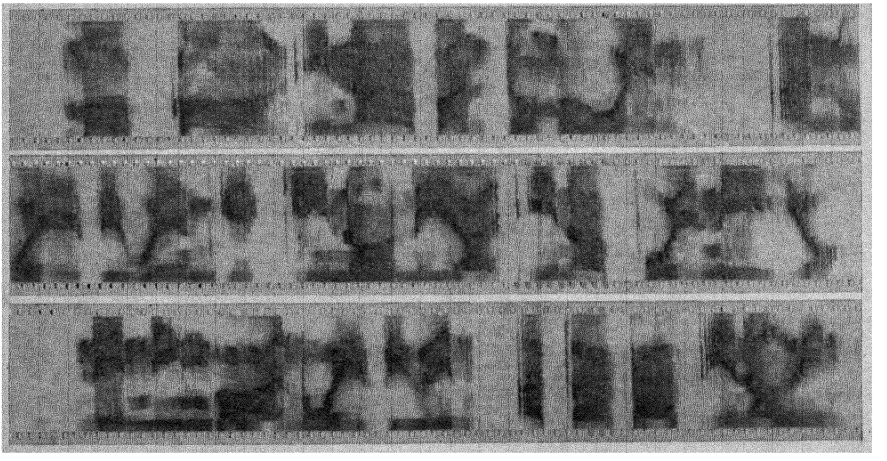


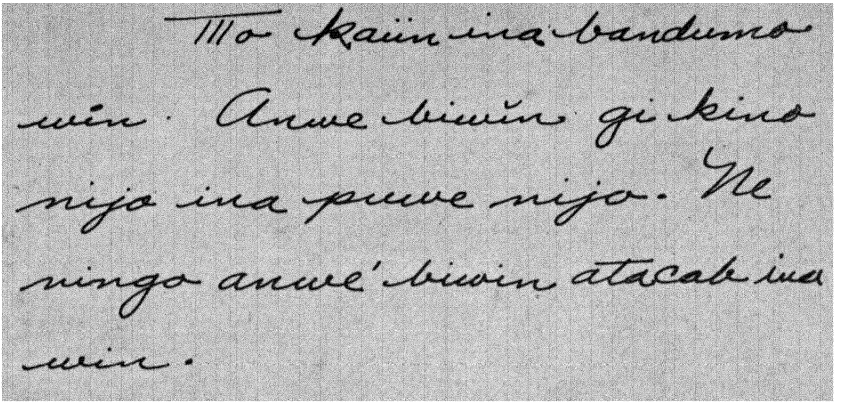

Der er ingenting mere nødven-
dig for mennesker enn a kunne
tale til hverandre. I hele verden
er der mange sprаak som blir talt
av de forskjellige mennesker. Lydig
tale viser likheten og forskjell mel-
lem de forskjellige sprаak.

The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Norwegian

This speaker was born near Stavanger, in the western part of Norway, and was brought up in that area. He classifies his speech as "book language — Western influence." He started learning English in early grades at school, and came to the United States about twenty years ago.

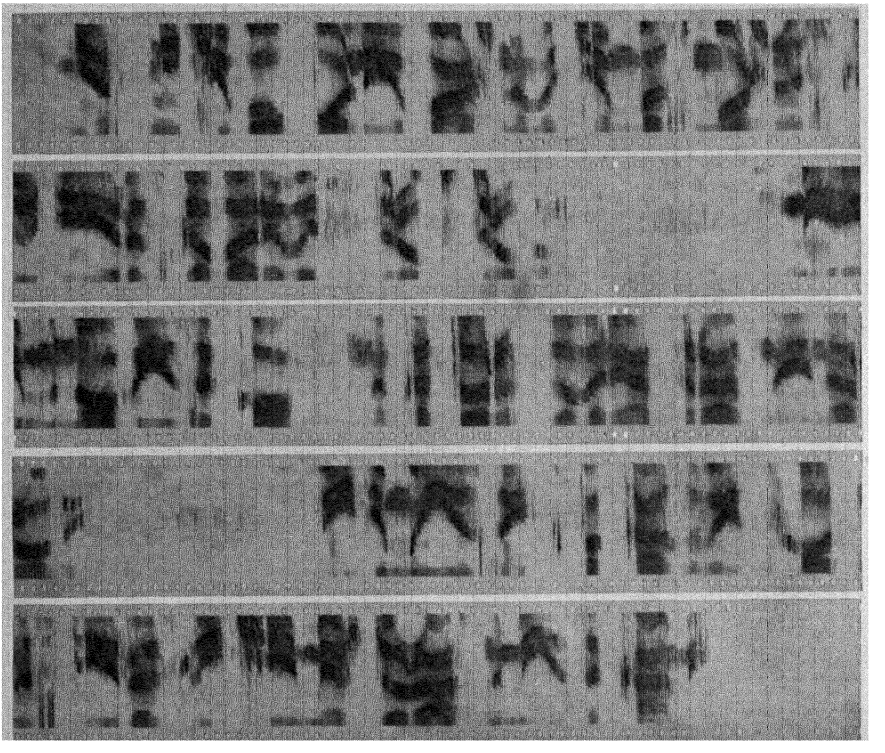


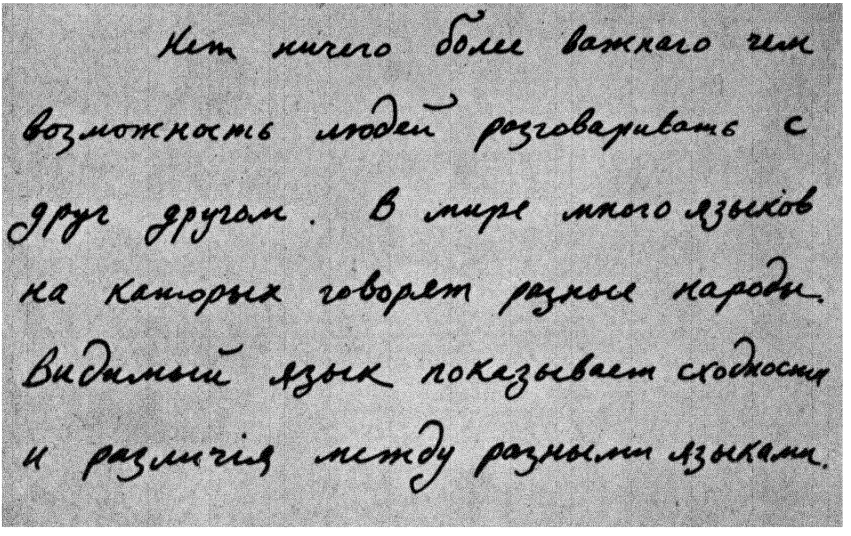


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Ojibway (American Indian)

This speaker was born on the Onondaga Reservation in New York State. He considers his speech to be "typical of the eastern woodlands people in the Great Lakes region." He learned both Ojibway and English in early childhood.

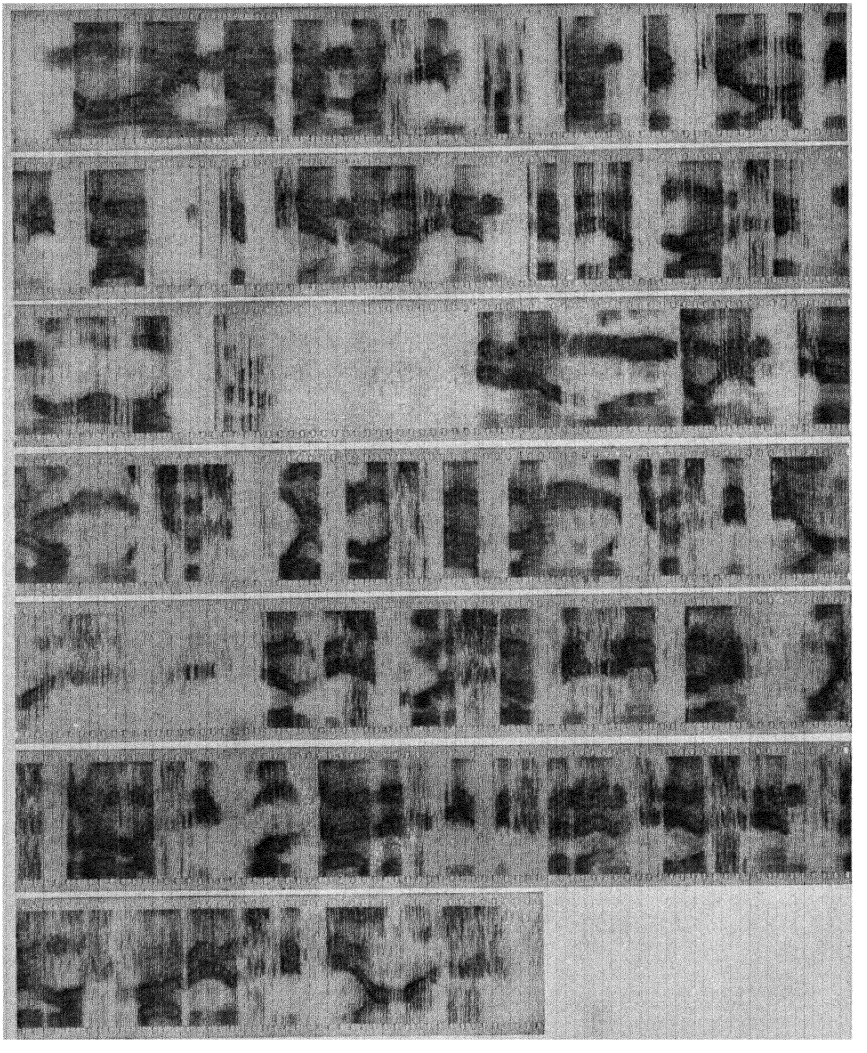


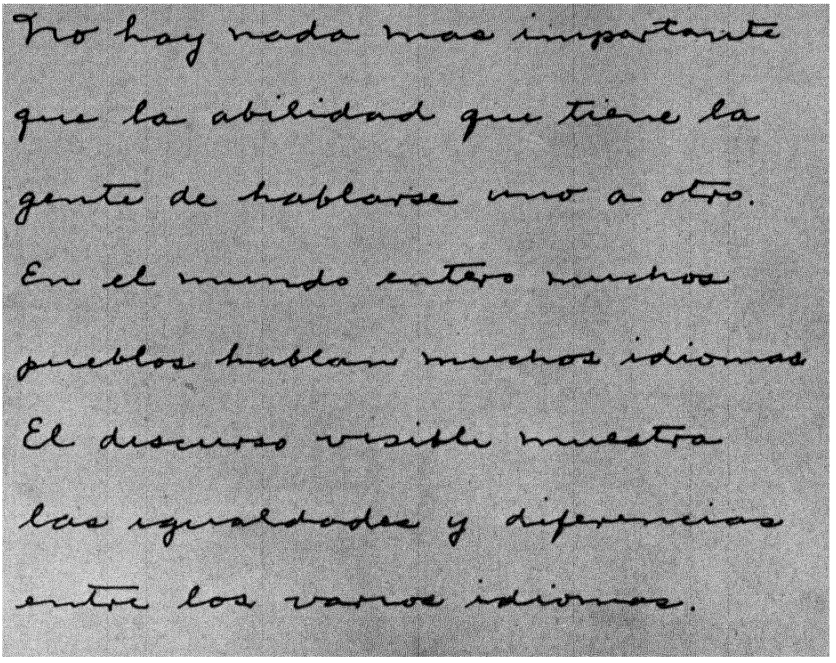

Нет ничего более важного чем
возможность людей разговаривать с
друг другом. В мире много языков
на которых говорят разные народы.
Видимый язык показывает сходности
и различия между разными языками.

The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Russian

This speaker was born in Geneva, Switzerland, of Russian parents. He was taken to Moscow when about a year old, and was brought up in that area. He classifies his speech as "Great Russian, or Central European Russian." He has spoken English for about twenty years.

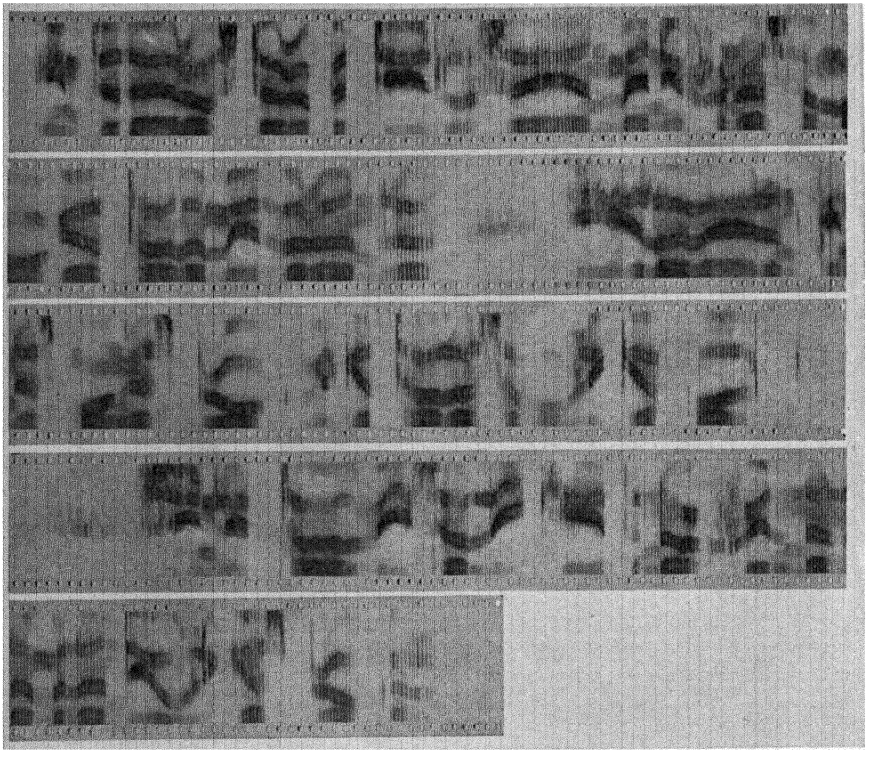


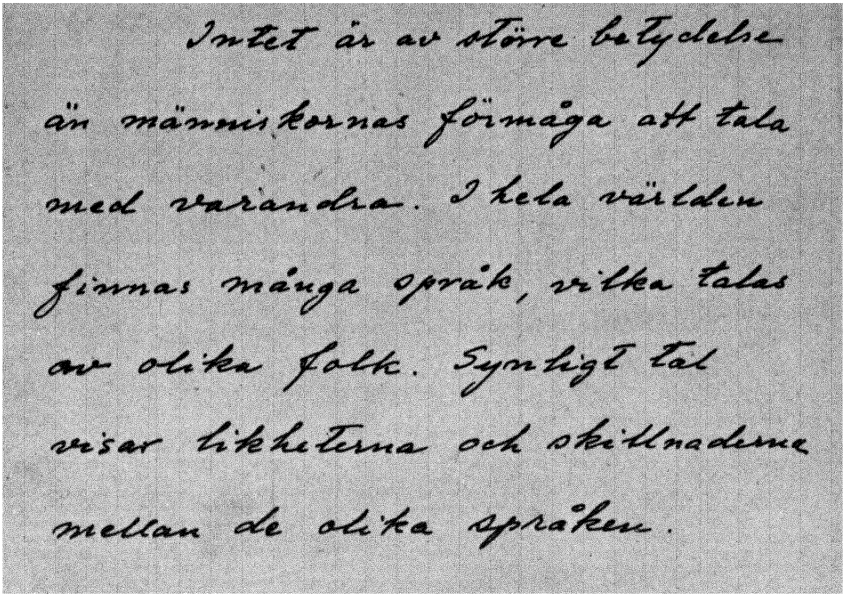


The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Spanish

This speaker was born in Manila, Philippine Islands, and was brought up in that area. He classifies his Spanish as "Castilian." He has spoken English from early childhood.

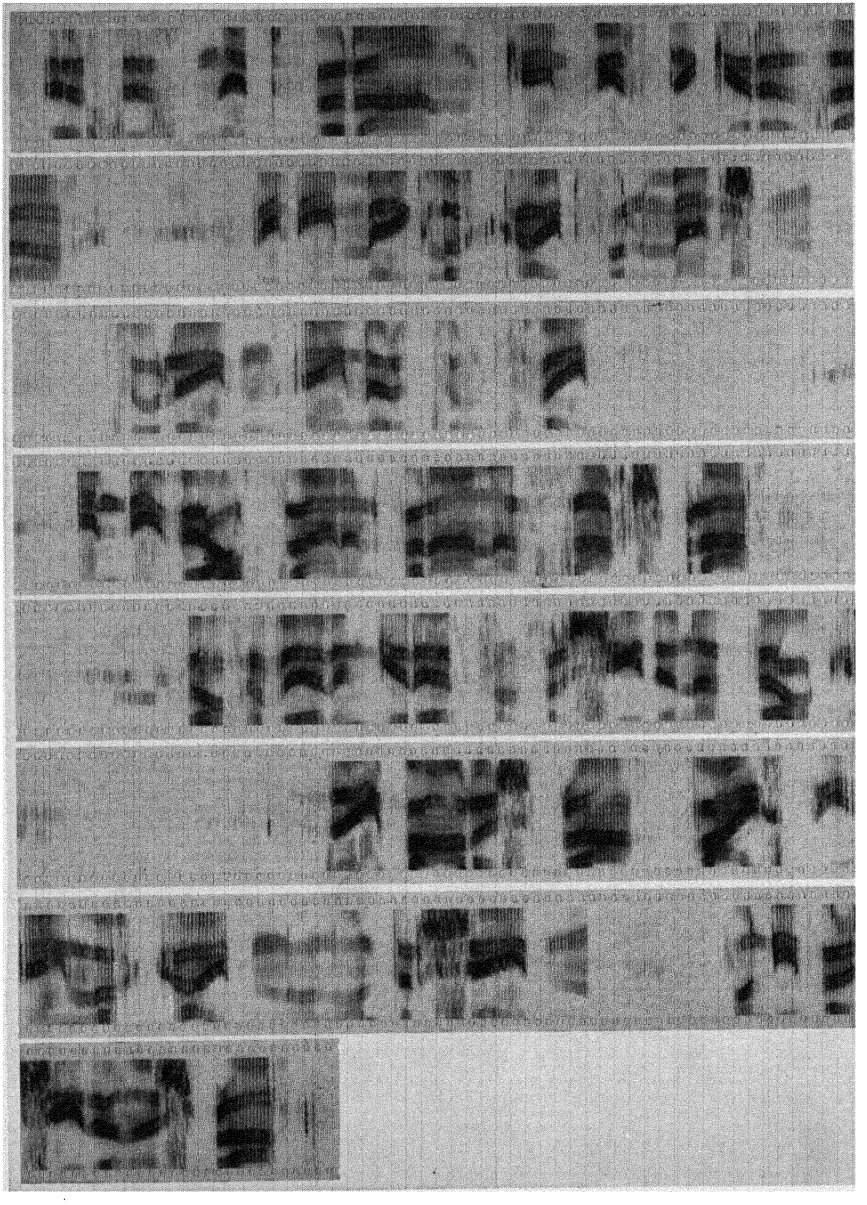


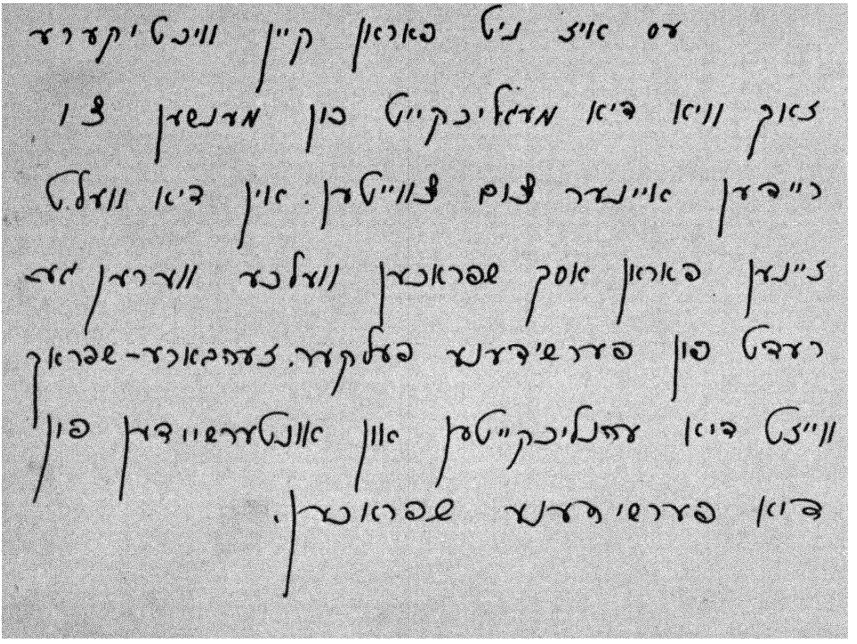

Intet är av större betydelse
än människornas förmåga att tala
med varandra. I hela världen
finnas många språk, vilka talas
av olika folk. Synligt tal
visar likheterna och skillnaderna
mellan de olika språken.

The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Swedish

This speaker was born in Helsingfors. He classifies his speech as "pure Swedish," rather than the dialect spoken in Skane. He spoke English before he came to the United States as a young man about twenty years ago.





The speaker's translation of the English text is shown above in his manuscript; the visible speech transcription of his oral rendition appears on the facing page.

Yiddish

This speaker was born and brought up in Brooklyn, New York. He classifies his speech as "mainly Lithuanian Yiddish." He learned both Yiddish and English at an early age.

Chapter 17

Vocal Music Interests

Techniques of sound portrayal such as have been discussed through the preceding chapters will doubtless find application in studies of vocal music. A proper presentation of these possibilities should, of course, be based upon an analytical study and organization of data such as was carried on in the visible speech case, but this would be a major undertaking and one requiring specialized talent and time not available. As a matter of fact, comparatively few sound spectrograms of vocal music had been made prior to the accumulation of material for this chapter, and attempts to analyze such patterns were limited to rather casual inspection. Accordingly, there was little background upon which to base a plan for either collection or presentation of material. The procedure finally adopted was simply that of obtaining phonograph recordings of a number of voices, listening to these recordings for passages that might provide interesting patterns, and then making spectrograms. Approximately a dozen of the voices selected might be classed as operatic, and as many more are well known as singers of popular music. In the operatic group are included a few samples taken from old recordings.

Examples were not chosen with any thought of displaying representative qualities in the different voices, nor has any attempt been made to associate the effects made visible with their audible equivalents, except in a very superficial way. Perhaps the most general impression one gains in a first examination of such a collection of patterns as those following is that the voices of all singers show many conspicuous similarities, so it may be the inconspicuous differences that account for variety in voice quality.

Visible versus audible effects

Some of the variations that appear in the following illustrations are readily identifiable with the aural equivalent. An example is the vibrato. Others are not so recognizable for the reason that the aural pattern is less well defined. The regular pitch variations in a vibrato are clearly distinguishable by ear as pitch changes; but when there is a small deviation in pitch superimposed upon a regular pattern, the aural and visual equivalents are not so evident. There is probably nothing in these visible patterns that is without its audible equivalent; but the association may, in some cases, be obscured, either because we have never before properly visualized the sound movement in these cases or because the translation apparatus used in obtaining the sound pictures does not adequately approximate the characteristics of the ear in frequency and time resolution.

Much remains to be learned about the relation between visible and audible equivalents, and it is a subject that should be of considerable interest to those in the musical field because it offers the possibility of a consistent system of visualization, and a common ground for musical thought and discussion. It would be possible to talk about musical sound effects in terms of these patterns without encountering the difficulties that arise in the use of any artificial method of symbolization.

Data on illustrations

Below each of the following illustrations is given the name of the singer, data concerning the phonograph record, and the approximate point on this record from which the sample was taken. For example, the caption of Fig. 1 reading "*Enrico Caruso (tenor); Vesti la Giubba (I Pagliacci, Act I). RCA Victor Record No. 6001–A (1-5/16 inches),*" means that the record is of Caruso's voice in "*Vesti la Giubba*" as recorded on RCA Victor record No. 6001–A, and that the sample was taken at a point approximately one and five-sixteenths inches in from the first outside groove. With this much information it should be possible to find the phonograph record and the transcribed sample.

All the vocal music patterns are shown here to the same scale, except where otherwise indicated. With exceptions, as noted, the time intervals occupied by the samples are always 2.25 seconds, and the vertical frequency scale extends from zero at the thin black line along the lower edge to 3,400 cycles. The frequency and time scales are shown in Fig. 1.

Wider frequency bands would probably be desirable in future studies, but for the purposes of a preliminary discussion it was decided to confine the portrayal to this rather narrow but important range. Also, a logarithmic or octave frequency scale would afford certain advantages; but the patterns here shown are all of the linear form because this form of display is in a more advanced stage of development.

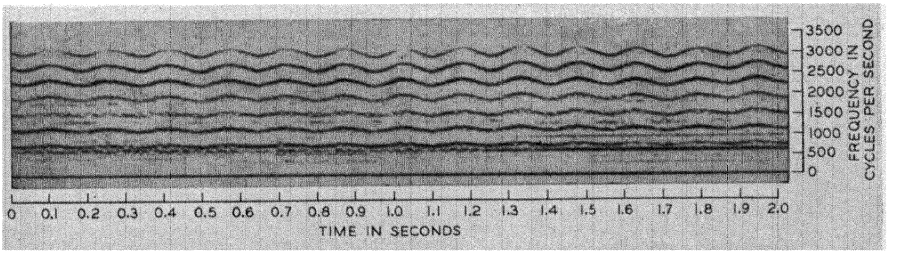


Fig. 1 — Enrico Caruso (tenor); "Vesti la Giubba" (I Pagliacci, Act I). RCA Victor Record No. 6001-A (1-5/16 inches).

Pattern components

Although the structure of speech patterns has been discussed in considerable detail throughout the previous chapters, it may be worth while to repeat a few points in relation to these vocal music patterns.

All of the patterns shown in this chapter are of the so-called narrow-band type, which means that the scanning filter width is 45 cycles as compared with the 300-cycle bandwidth utilized in the case of visible speech. This 45-cycle bandwidth approximates the average frequency resolution of the ear. It would separate tones spaced by more than 45 cycles, but would show as one variable intensity line any beat-note combination of tones that falls inside of the 45-cycle interval.

In Fig. 1, showing the pattern of a sustained note in G flat, on the equal tempered scale, by Caruso, the overtones appear as horizontal wavy lines. Since these overtones are in multiple frequency relation they are spaced at regular intervals up the frequency scale of the picture. The fundamental, or first horizontal bar in the series starting at the bottom, is almost missing, presumably due to the position of the first resonance. The faint background of substantially straight horizontal full and broken lines is due to instrumental music, the broken lines being beats and the full lines steady tones.

Voiced sounds are, of course, far more in evidence than unvoiced sounds throughout vocal music, so that most of the patterns shown here comprise orderly arrangements of overtones that appear and disappear selectively with change in resonation, and more or less in unison when the voicing is started or stopped by initiation or termination of the breath flow.

The vibrato

A conspicuous effect in most of the patterns is the vibrato. It appears in the shape of waves in the overtone lines, and is well illustrated by Fig. 1. The number of waves per second, or the vibrato rate, is 6.3 in this case; and the total frequency swing from top to bottom of the vibrato wave is 7.7 per cent. (A semi-tone is approximately 6 per cent.)

In Fig. 2, taken from a little later point in the same Caruso recording, there appears a gradually reduced vibrato swing accompanying a diminuendo. In the earlier part to the

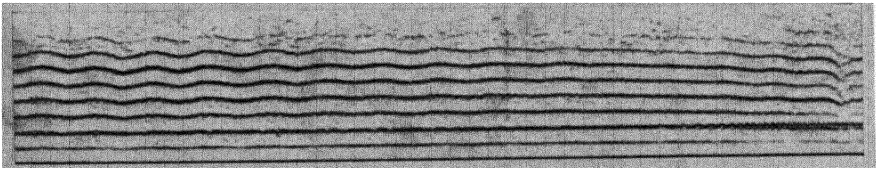


Fig. 2 — Enrico Caruso (tenor); "Vesti la Giubba" (I Pagliacci, Act I). RCA Victor Record No. 6001-A (1-9/16 inches).

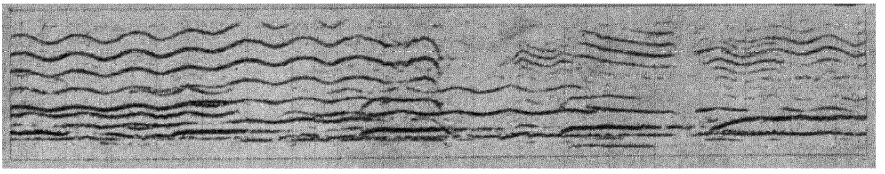


Fig. 3 — Enrico Caruso (tenor), Marcella Sembrich (soprano), Antonio Scotti (baritone), Gina Severina (contralto); "Bella Figlia dell 'amore" (Rigoletto, Act III). RCA Victor Record No. 16-5001-A (2-10/16 inches).

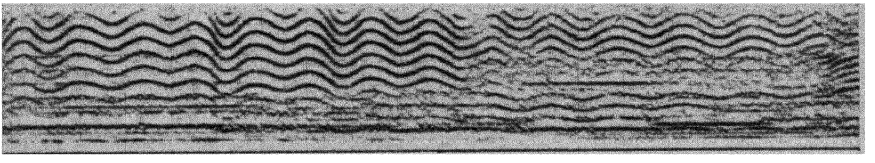


Fig. 4 — Louise Homer (contralto); "Je Viens Célébrer la Victoire" (Samson et Dalila, Act I). RCA Victor Record No. 15-5003-B (1-1/4 inches).

left, the vibrato rate is 6.7 cycles and the swing about 6.8 per cent. At the right, the vibrato practically disappears.

In Fig. 3, from a different record, is shown Caruso's voice at the left, and intermingled with it and following it are fragments of the voices of Marcella Sembrich, Scotti, and Gina Severina.

The dark recording that builds up from the left toward the middle of Fig. 4, means that the volume in this pattern is increasing to a peak slightly to the right of center, where it suddenly drops. Preceding the drop, the vibrato has been somewhat irregular, but in the low volume interval that follows it is quite steady. In the left half, the vibrato rate averages about 6.3, while in the low-volume region it falls to 5.5 cycles. The swing ranges between 12 and 14.5 per cent throughout the sample. The horizontal lines are due to a background of instrumental music.

In Fig. 5 appears the vibrato pattern of Frances Alda's voice with the voice of Caruso added at the extreme right. Alda's vibrato rate is about 7.7 cycles per second in these passages.

Fig. 6 shows a sustained vibrato of 6.6 cycles at a pitch of 990 cycles (B^1).*

* In the system of notation used here to indicate the approximate location of a tone in the musical scale, C designates middle C. The notes in the octave below and above C are designated by subscripts and superscripts, as C_1 D_1 E_1 F_1 G_1 A_1 B_1 C^1 D^1 E^1 F^1, etc. The notes two octaves below and above middle C are written C_2 and C^2 respectively.

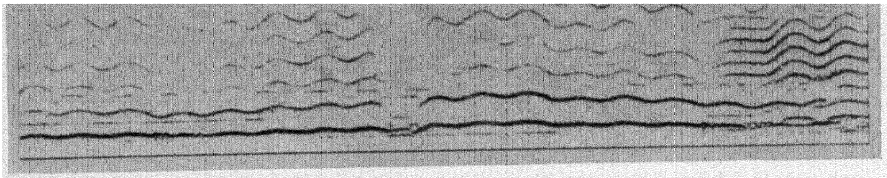


Fig. 5 — Frances Alda (soprano) followed by Enrico Caruso (tenor); "Qual Volutta Trascorrere" (I Lombardi, Act III). RCA Victor Record No. 16-5002-B (1-15/16 inches).

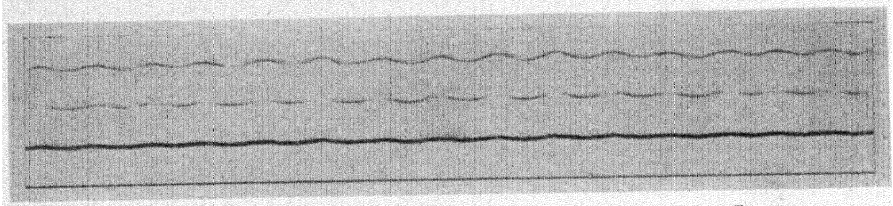


Fig. 6 — Luisa Tetrazzini (soprano); "Chi mi Frena" (Lucia di Lammermoor, Act II). RCA Victor Record No. 16-5000-A (1-5/16 inches).

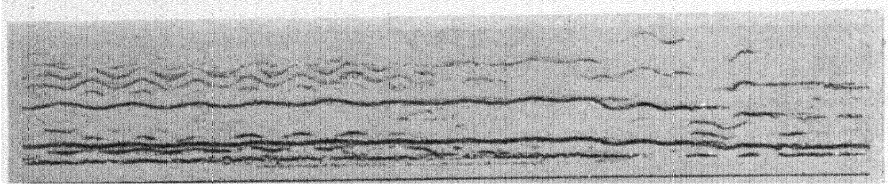


Fig. 7 — Enrico Caruso (tenor) and Frieda Hempel (soprano); "E Scherzo, od E Follia" (Un Ballo in Maschera, Act I). RCA Victor Record No. 16-5000-B (1-1/16 inches).

Fig. 7 shows Frieda Hempel's voice, identified by the widely spaced and darker lines running through the pattern, and with what was tentatively identified as Caruso's voice appearing in the steady vibrato toward the top of the record. The vibrato rates in the two cases are about 5.6 and 7.5 cycles per second, respectively.

A regular vibrato appearing after a smoothly executed down glide appears in the pattern of Fig. 8 by Sigrid Onégin. The vibrato rate is 5.9 cycles and the swing about 6.8 per cent. The pitch at the left is 970 cycles (near B^1), and at the end of half a second it drops to an average 550 cycles (near D^1 flat). Again there is a faint background of instrumental tone.

Another pattern by Onégin in Fig. 9 shows a down glide followed by a vibrato in a pitch range more than an octave below that of Fig. 8. The vibrato rate near the middle of this sample is 6.6 cycles and the swing is about 9 per cent. At the top on the right,

following a rise in volume, the voicing terminates on the down swing of the vibrato and is resumed again on the up swing.

The Martinelli pattern of Fig. 10 shows a sustained vibrato followed by a rapid down glide in which some evidence of vibrato is retained. In the first half, the vibrato rate is about 7.0 cycles and the swing 5.7 per cent. Figures 11 to 13 inclusive show various sustained passages from the same recording.

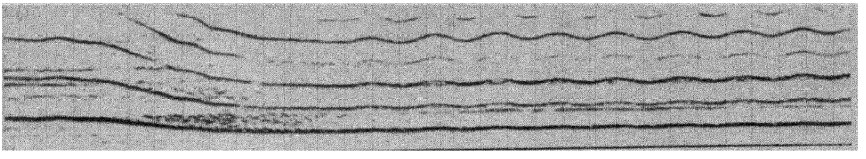


Fig. 8 — Sigrid Onégin (contralto); "Mon Coeur S'ouvre à Ta Voix" (Samson et Dalila, Act II). RCA Victor Record No. 7320-A (3-1/4 inches).

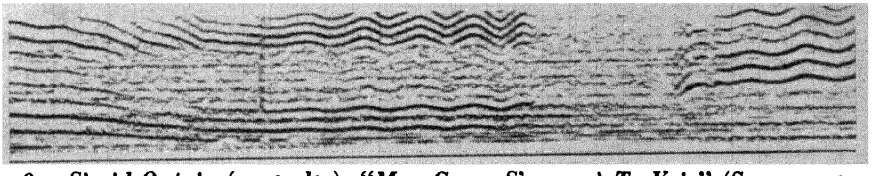


Fig. 9 — Sigrid Onégin (contralto); "Mon Coeur S'ouvre à Ta Voix" (Samson et Dalila, Act II). RCA Victor Record No. 7320-A (1-7/16 inches).

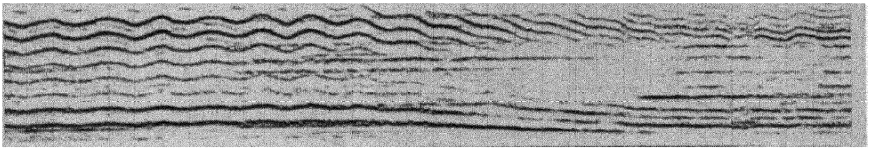


Fig. 10 — Giovanni Martinelli (tenor); "Céleste Aida" (Aida, Act I). RCA Victor Record No. 14206-B (2-2/16 inches).

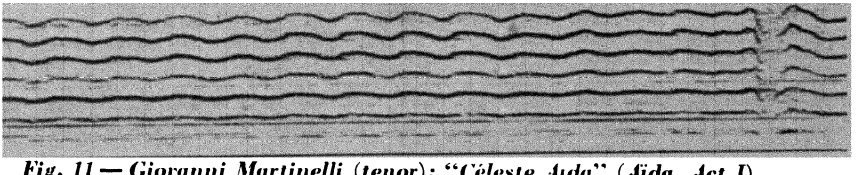


Fig. 11 — Giovanni Martinelli (tenor); "Céleste Aida" (Aida, Act I). RCA Victor Record No. 14206-B (3-1/16 inches).

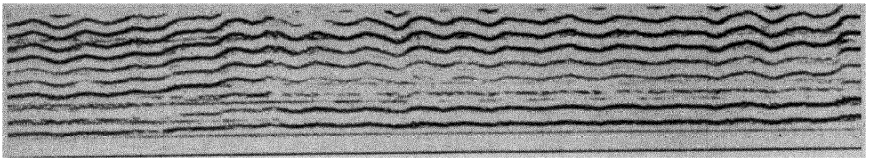


Fig. 12 — Giovanni Martinelli (tenor); "Cèleste Aïda" (Aïda, Act I). RCA Victor Record No. 14206-B (1-7/16 inches).

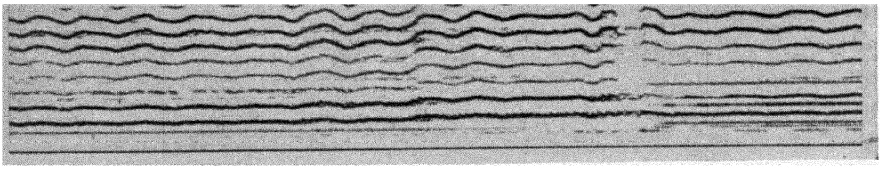


Fig. 13 — Giovanni Martinelli (tenor); "Cèleste Aïda" (Aïda, Act I). RCA Victor Record No. 14206-B (1-5/16 inches).

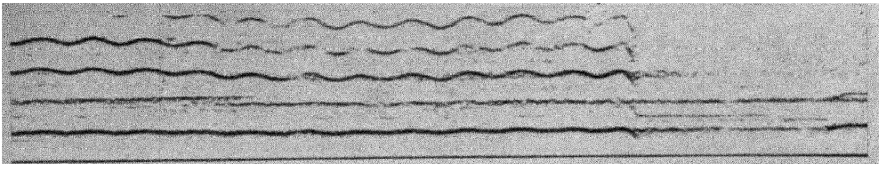


Fig. 14 — Lucrezia Bori (soprano); "Mi Chiamano Mimì" (La Bohème, Act I). RCA Victor Record No. 14206-A (2-5/16 inches).

The pattern of a vocal music sample from a pattern by Lucrezia Bori in Fig. 14 shows a vibrato rate of 7.5 cycles and a swing of 7.0 per cent, while the average pitch drops from 735 cycles (F^1 sharp) to about 590 cycles (D^1) and then rises slightly beyond. A little to the right of center appears a slight accentuation of swing during one vibrato cycle, the characteristics of which will be significant in relation to effects discussed in the following section. At the right is a stop followed by resumption of tone with the fundamental accentuated.

Vibrato shape variations

When a steady tone is shifted very gradually over a small pitch range, the aural impression is quite different from that experienced when the same small pitch change occurs momentarily. Presumably there is an analogous situation in the vibrato case.

When a vibrato of one form is changed gradually to another that is slightly different, the change may not be noticeable; but if it occurs from one vibrato cycle to the next it may be very obvious.

In a later section on "Pitch Wavers", it is shown that small momentary changes of pitch in the absence of vibrato are responsible for a very conspicuous and well known effect in singing. It may be that similarly small variations superimposed upon the vibrato are also responsible for effects that are obvious to the ear. That is, further study of vocal music patterns may attach considerable significance to the wave shapes of vibrato cycles and particularly the cycle-to-cycle variations in these shapes.

Irregularities in vibratos that sometimes appear in sustained passages are illustrated by Figs. 10 to 13 inclusive. Others that accompany modulation may be seen in Fig.15. These changes in vibrato shape appear frequently and could probably form the basis of a vibrato classification. Many vibrato shapes appear throughout the samples of Figs. 9 to 16. A good example of a slightly angular, but symmetrical, vibrato wave form appears near the center of Fig. 9, while an asymmetrical shape is shown toward the extreme right of Fig. 16.

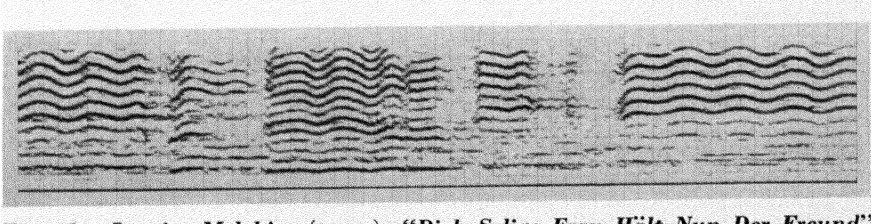


Fig. 15 — Lauritz Melchior (tenor); "Dich Selige Frau Hält Nun Der Freund" (Die Walküre, Act I). RCA Victor Record No. 14204-B (1-7/16 inches).

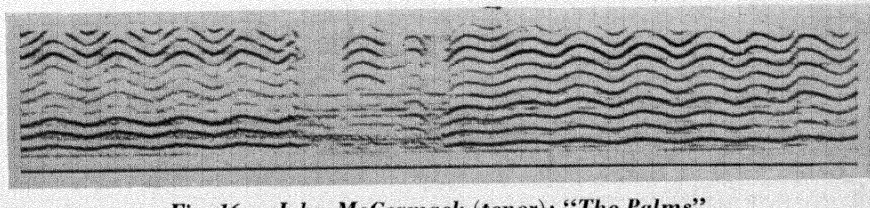


Fig. 16 — John McCormack (tenor); "The Palms". RCA Victor Record No. 6607-B (1-10/16 inches).

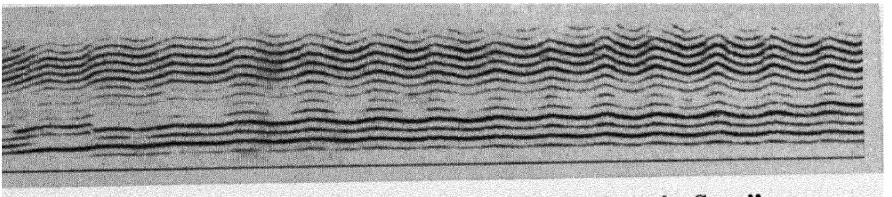
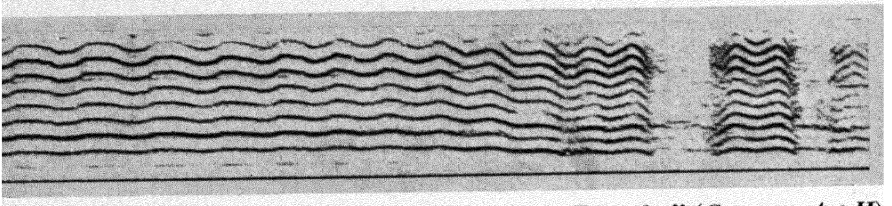


Fig. 17 — Lawrence Tibbett (baritone); "Pilgrim's Song". RCA Victor Record No. 7779-B (3 inches).



g. 18 — Lawrence Tibbet (baritone); "Chanson du Toréador" (Carmen, Act II) RCA Victor Record No. 14202-A (2-3/16 inches).

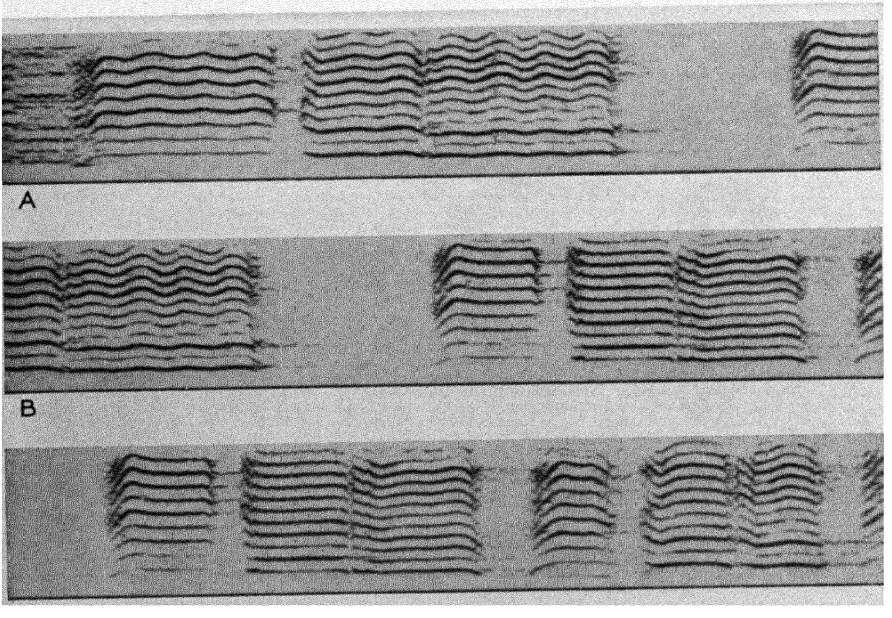


Fig. 19 — Lawrence Tibbett (baritone); "Largo al Factotum" (Barbiere di Siviglia). RCA Victor Record No. 14202-B (2-7/16 inches).

It is interesting to trace the treatment of vibrato through the samples of Figs. 17 to 19 inclusive. Note the change in shape during the two and a fraction seconds included in Figs. 17 and 18. Also notice that A, B, and C of Fig. 19 are all part of the same passage, and the repeated sequences indicate the exactness with which the same parts may be reproduced in different spectrograms of the same sounds.

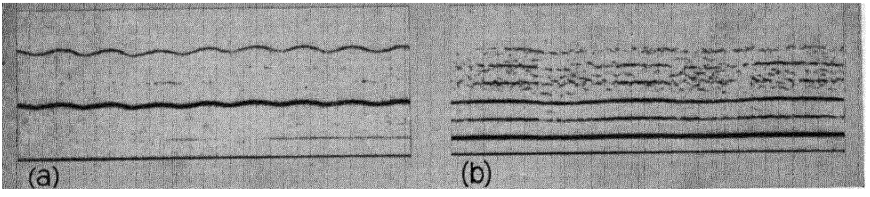


Fig. 20 — Lily Pons (soprano); "Caro Nome" (Rigoletto, Act I). RCA Victor Record No. 14203-A (3-11/16 inches). In (b), the frequency scale is 3.07 times that regularly employed.

Patterns of the vibrato under the numerous conditions of use by one singer are illustrated in another case by Figs. 20 to 25, inclusive. These show samples that give some indication of what might be learned from a careful analysis of such variations. In Fig. 20(a), the pitch is about 1325 cycles (E^2), the vibrato rate is sustained at around 7.5, and the swing about 7.0 per cent. The straight horizontal line that appears about 500 cycles above the base line is an instrumental note. The very dark wavy line is the fundamental, and double the distance of the fundamental from the base line is the second harmonic. The third and any higher harmonics would fall outside the frequency range shown in this type of pattern. In Fig. 20(b) is shown the same sustained vibrato in a wide-frequency-range type of record. The frequency scale used here is 3.07 times that of Fig. 20(a) and the time scale is divided by 3.07. Note that the third harmonic, which is beyond the range of Fig. 20(a), shows up clearly in Fig. 20(b), and that in Fig. 20(b) there are also traces of the fourth, fifth, and sixth harmonics.

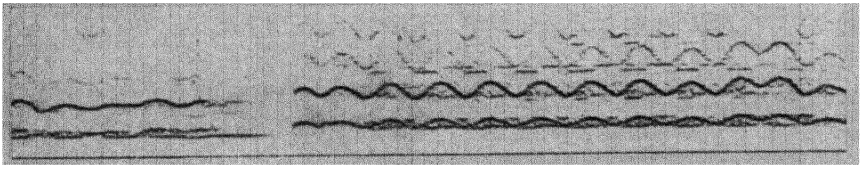


Fig. 21 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (3-1/4 inches).

The sound sample in Fig. 21 would probably be classed as a "trill", and it is remarkable in that the vibrato is sustained at 7.5 cycles while the average swing is about 23.5

per cent! The pitch is about 790 cycles. In this case, the resonance is low; and it would be interesting to know whether such a wide vibrato swing could be sustained in a higher resonance, and also whether it would be as acceptable under those conditions.

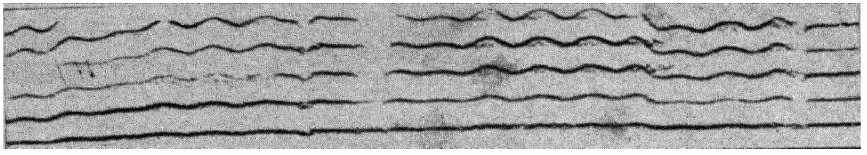


Fig. 22 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (11/16 inches).

Fig. 22 shows variations in the vibrato accompanying a rapid rise in pitch at the left; an extremely brief stop left of the middle; followed, toward the right, by a sudden drop from 650 cycles (E^1) to 580 (D^1); and followed finally by another brief stop about 0.4 seconds later.

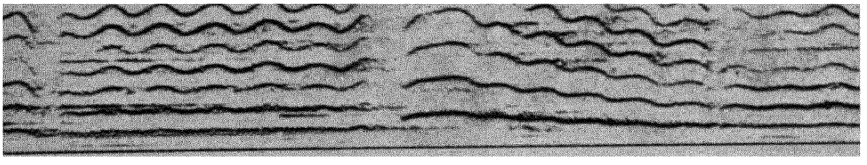


Fig. 23 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (1 inch).

At the left of Fig. 23 there is a brief stop followed by a sustained note, then another break and a down glide. The vibrato terminates on the up-swing and picks up again, also on the up-swing, in going to the higher pitch.

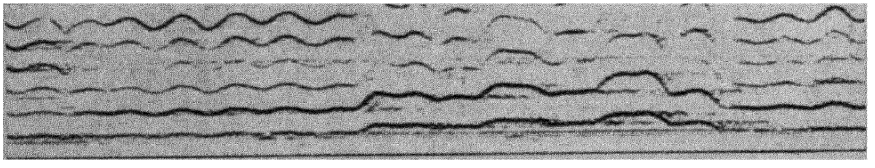


Fig. 24 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (15/16 inch.)

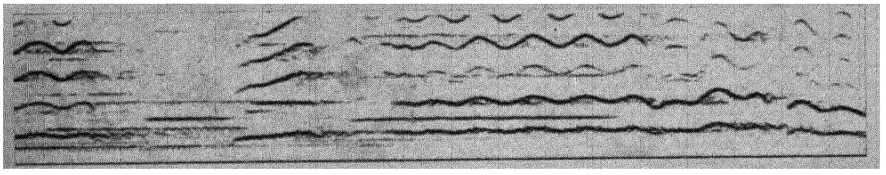


Fig. 25 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (1-6/16 inches).

The tendency to maintain vibrato throughout passages that are modulated very irregularly is illustrated by Figs. 24 and 25. Except in the sudden up-glide at the beginning of the sample in Fig. 25, there is evidence of pronounced vibrato in all parts of these pictures.

Repetition of vibrato

The ability of a trained singer to repeat a vibrato run is shown rather clearly in Fig. 26. This is a similar down-glide sung to two different pitch levels and taken from different parts of the same phonograph record. As shown by traces of the third harmonics from the two accompanying spectrograms, one is nearly a duplicate of the other except in average pitch. The slight irregularity in the vibrato that anticipates the down glide is much the same in each case, and the vibrato rate and extent of swing is nearly equal over the remainder of the sample.

Close duplication of the vibrato relationship seems to occur in many instances, and these patterns of sound can show very clearly both the manner of handling of the same passages in different parts of one rendition, and their handling in different renditions of the same song.

Some interesting examples of the latter appeared while these records were being made. In one case, a phonograph record was accidentally scratched, causing vertical shading on the spectrograms at the record rotation interval; so a new record of the same make, number and title was obtained. Spectrograms remade with this new record did not show exactly the same detail, and after comparing several passages it was apparent that the two were made from different renditions of the same song. But despite the differences that appeared in the patterns, the parts were repeated in both recordings with rather remarkable accuracy, even to a correspondence of the number of vibrato swings. Often the differences were in minor variations of shape in particular vibrato cycles.

Variations in vibrato rate

Changes in the vibrato rate, or shape, or swing, or combinations of these, appear to anticipate modulations of volume or pitch, as if preparation for these modulations

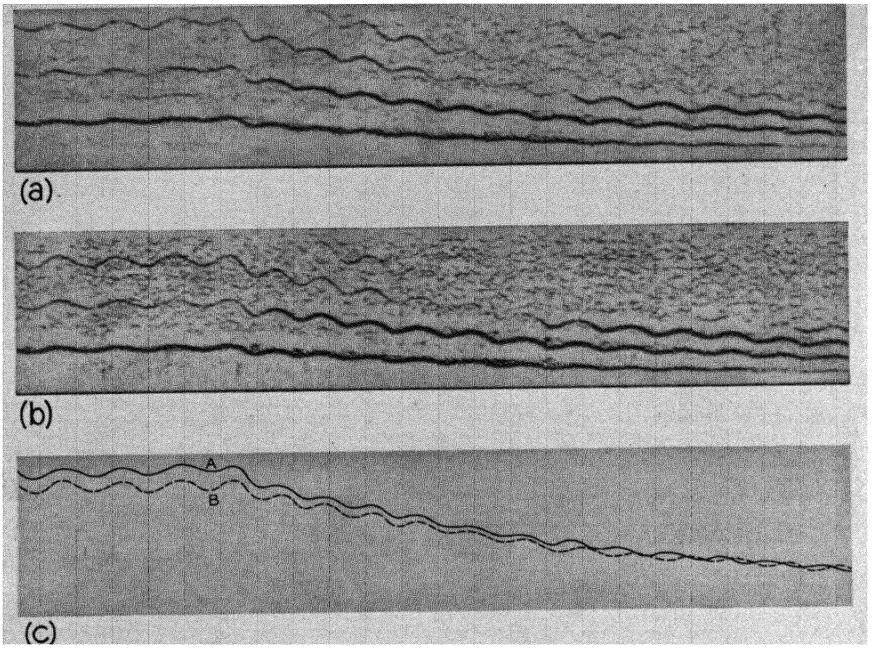


Fig. 26 — Lily Pons (soprano); "Caro Nome" (Rigoletto, Act I). RCA Victor Record No. 14203-A: (a), (2-3/8 inches); (b), (3-1/4 inches); (c), tracings of the third harmonic from (a) and (b).

somehow affects the vibrato mechanism. Also, an after-effect of the modulation upon the vibrato is usually evident, although in some cases it may be small. In Fig. 27, for example, the sustained vibrato rate at the left is 6.5 cycles a second and the swing 10.0 per cent. Preceding the stop and immediately ahead of the down glide into this stop, the vibrato rate increases to over 7.0 cycles. Following the stop is an appreciable interval during which the vibrato seems to be building up.

At the left of Fig. 28, the vibrato follows a slight rise in pitch and then terminates in a down sweep at the stop approaching the middle of the sample. The conditions following this brief stop are similar to those pictured for the longer partial stop of Fig. 27. The vibrato rate rises to over 7 cycles a second and then falls back to about 6.7 cycles with a swing of about 10.0 per cent.

During the fraction of a second between stops left of center in Fig. 29, the vibrato rate is 6.7 cycles and the swing 10.0 per cent. Following the centrally located stop, the vibrato rate rises to 7.3 per cent without a change in the swing. With the abrupt pitch rise from 550 cycles (C^1 sharp) to 970 cycles (B^1), the vibrato rate drops suddenly to

5.8 cycles; and later, beyond the limits of this sample, speeds up to 6.4 cycles with a swing of about 4.0 per cent.

A considerable change in vibrato shape and rate often accompanies the enunciation of words in vocal music. Fig. 30 represents the words "*the thrill it knew*". The darker regions of resonance may be seen superimposed upon the overtones. The passage begins with the short voiced "*the*". The voiceless "*th*" in "*thrill*" appears as the weak sound that follows, and continues into the remainder of the word "*thrill*", blending into "*it*" which terminates in the stop portion of the "*t*". In the left middle range, the vibrato rate averages about 6.0 cycles and the swing averages about 8.5 per cent. Note that the change in resonance just ahead of the stop at middle right causes an irregularity in the vibrato. Following the stop, the vibrato rate falls to 5.8 and the swing again builds up to 8.5 per cent.

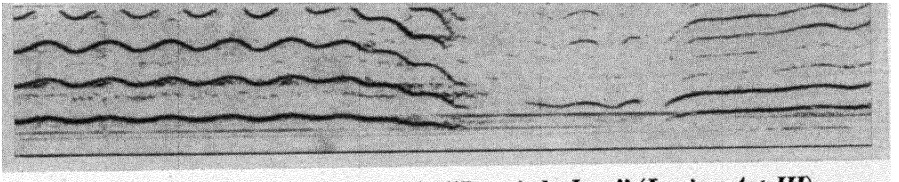


Fig. 27 — Helen Jepson (soprano); "Depuis le Jour" (Louise, Act III). RCA Victor Record No. 14204-A (1-13/16 inches).

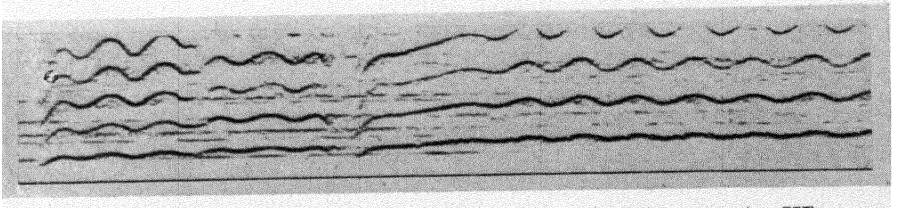


Fig. 28 — Helen Jepson (soprano) "Depuis le Jour" (Louise, Act III). RCA Victor Record No. 14204-A (1-1/8 inches).

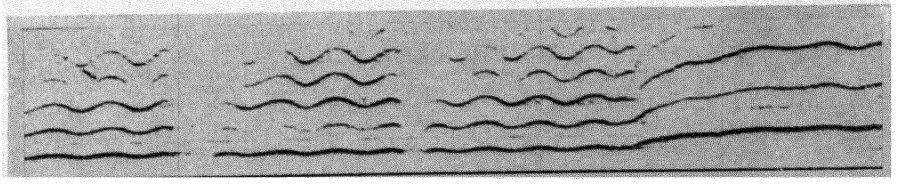


Fig. 29 — Helen Jepson (soprano); "Dupuis le Jour" (Louise, Act III). RCA Victor Record No. 14204-A (3 inches).

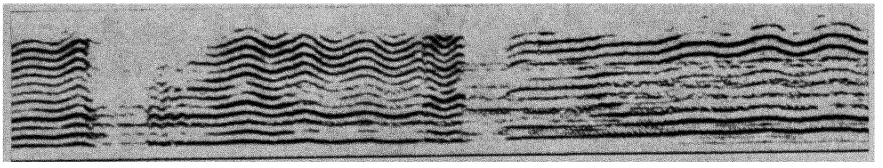


Fig. 30 — Nelson Eddy (baritone); "Will You Remember?" (Maytime). RCA Victor Record No. 4329-A (3/4 inch).

There is some evidence in these illustrations that the vibrato rate may be related loosely to fundamental frequency, the rate dropping as the pitch is increased. Such an association seems to show up more distinctly in voices of low register. Possibly the connection is more directly with effort or volume.

Incidentally, the percentage swing of the vibrato observed in the samples here illustrated ranges from near zero up to a peak in the vicinity of 8.0 per cent, falling off rather rapidly around 10 to 12 per cent. In the rather limited number of voice patterns recorded, the vibrato swing seems to average lower in those of an operatic type than it does in singers of popular music. At the same time, the vibrato rate has averaged higher for operatic singers.

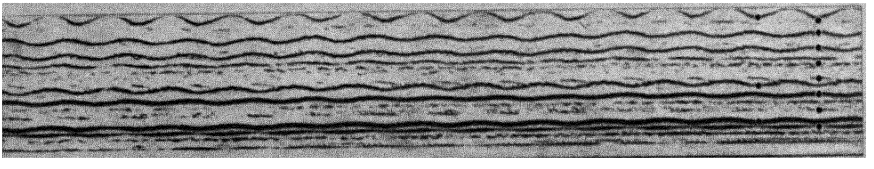


Fig. 31 — Jeannette MacDonald (soprano) and Nelson Eddy (baritone); "Will You Remember?" (Maytime). RCA Victor Record No. 4329-A (2-3/4 inches).

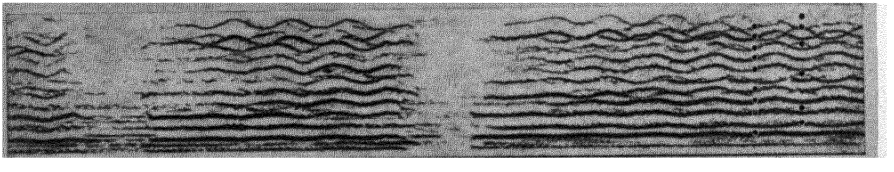


Fig. 32 — Jeannette MacDonald (soprano) and Nelson Eddy (baritone); "Farewell to Dreams". RCA Victor Record No. 4329-B (1-9/16 inches).

Vibrato in duet

Sound spectrograms offer an excellent opportunity to study voices, or voices and instrumental tones in combination. Examples are shown in Figs. 31 and 32, which represent Jeannette MacDonald and Nelson Eddy in a duet. At the right end of both these figures, the overtones of the two voices are identified by two vertical series of dots, "E" representing Eddy and "M" MacDonald. In Fig. 31, the pitch of "E" is F sharp (measured 370 cycles) and that of "M" is approximately A^1 (measured 860 cycles). In Fig. 32, "E" is middle C (measured 258 cycles) and "M" is F sharp (measured 360 cycles). It will be observed that the separate vibratos of Figs. 31 and 32 are in different phase relationship but of almost exactly the same rate.

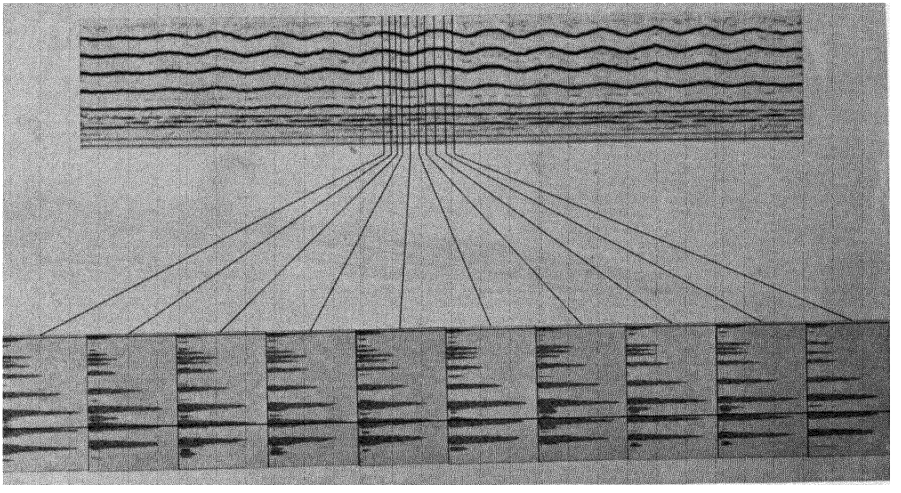


Fig. 33 — Enrico Caruso (tenor); "O Tu Che Seno a Gli Angeli" (Forza del Destino, Act III). RCA Victor Record No. 88207 (3-1/4 inches). Cross sectional analyses are shown beneath the conventional spectrogram.

Cross-section vibrato

A modification of the sound spectrograph makes it possible to produce frequency-loudness, or cross-section patterns at successive points along the time dimensions of a sound sample. Examples of such cross-section patterns are shown in the vertical series of Fig. 33. They picture the loudness dimension more quantitatively than is possible by shades of gray, and for some kinds of investigations they should be particularly revealing. The points in the sound sample from which the cross-section samples were taken

are indicated by the lines to the sound spectrogram at the top. In the cross-section patterns, the amplitude values follow a logarithmic (decibel) scale, and the frequency scale is the same as in the spectrogram but *reversed* in sense. The several peaks that intermingle with the regularly spaced overtones of the voice at the left end of the cross-section patterns are due to instrumental music. A straight line through the fifth harmonic makes the vibrato swing more evident, or it may be seen even more clearly by holding the page at eye level and sighting up the rows of overtones.

Cross-section patterns such as these appear to contain evidence that: (1), the conditions of resonance change over the vibrato cycle; and, (2) that both resonation and generation may determine the final quality of sung sounds. When the overtone amplitudes are plotted against frequency, those components that occur at the borders of resonance produce open ellipses rather than lines, while those at the peaks of resonance approximate simple lines or closed ellipses. If there were an amplitude variation in the originally generated overtones, and if this were out of phase with the vibrato shift, elliptical plots would be obtained for all the overtones regardless of their position in relation to resonance.

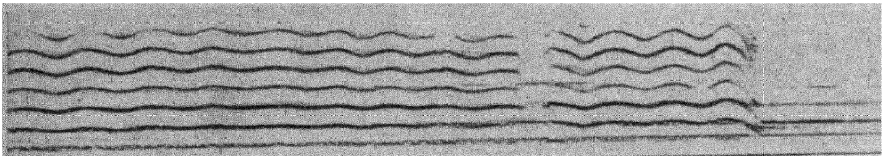


Fig. 34 — Enrico Caruso (tenor); "Vesti la Giubba" (I Pagliacci, Act I). RCA Victor Record No. 6001-A (1-9/16 inches).

Stops

No doubt a great deal could be learned by pattern studies of stops produced by glottal, tongue, or lip closures, and by breath control. Pitch changes may occur with no definite gap in tone production, but the minimum length of a full stop in the samples here illustrated seems to be on the order of 7 or 8 hundredths of a second.

Pitch variations entering and leaving a stop interval assume a wide variety of shapes ranging between essential flatness on both sides and sudden drops or rises. The most common type is evidently one in which the pitch drops ahead of the stop and rises following, although the pre-stop and post-stop conditions may seemingly be dependent upon both the conscious desires of the singer and unconscious influences such as those associated with average pitch change from one side of the stop to the other.

In Fig. 34, to the right of the center, is a stop about 8 hundredths of a second long (disregard the background of instrumental tone that carries through this break), followed by another of substantial length further to the right. In the former, there is a

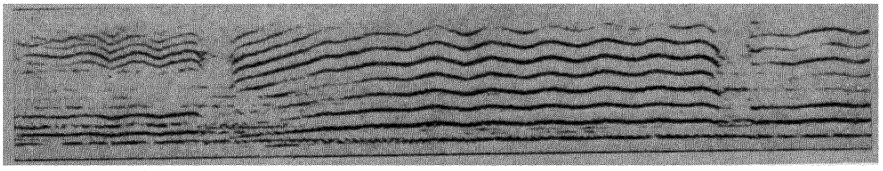


Fig. 35 — Enrico Caruso (tenor); "Je Viens Célébrer la Victoire" (Samson et Dalila, Act I). RCA Victor Record No. 16-5003-B (14/16 inches).

sudden discontinuance of tone on an almost flat pitch; and, following the brief silent interval, there is a small but rapid rise, breaking sharply into the vibrato cycle. At the subsequent long stop, note the quick descent in pitch, timed with the down swing of the vibrato.

At the left of Fig. 35 is another stop preceded by a down swing with the vibrato, but followed by a sudden rise and an up glide in pitch that shows very little vibrato. At the extreme right, entering another brief stop, the regular vibrato breaks with the vibrato swing; and on resumption of tone the pitch has dropped slightly (from 385 to 370 cycles).

Figure 36 is of interest largely in the stop treatment. There is a perceptible rise in pitch before the rapid downward swing entering a stop that is only about 6 hundredths of a second long. Following this, the tone reappears on an almost straight up glide. Here the pitch following the stop seems to have influenced the pitch change preceding the stop interval. Possibly the vocal cord tension changed progressively through the stop period.

A very abrupt stop with a momentary increase in volume followed by a rapid fall of pitch appears in Fig. 37. The suddenness of this stop is fully as apparent to the ear as it is in these patterns. Although it occurs on the down swing of the vibrato the cycle is substantially shortened. The resumption of tone on the extreme right takes a pitch form similar to that following the first stop in Fig. 35 and in other patterns of this chapter.

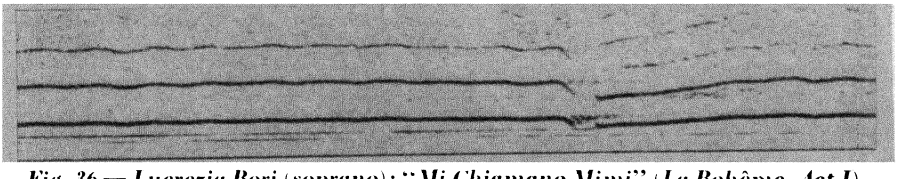


Fig. 36 — Lucrezia Bori (soprano); "Mi Chiamano Mimi" (La Bohême, Act I). RCA Victor Record No. 14206-A (14/16 inch).

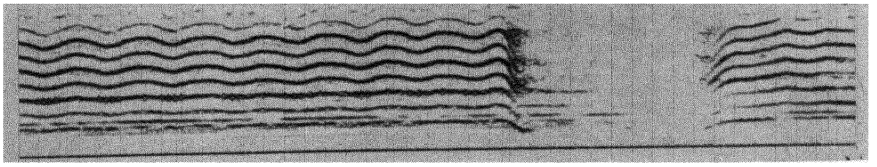


Fig. 37 — John Charles Thomas (baritone); "Di Pronenza il Mar" (La Traviata, Act II). RCA Victor Record No. 14205-B (2-12/16 inches).

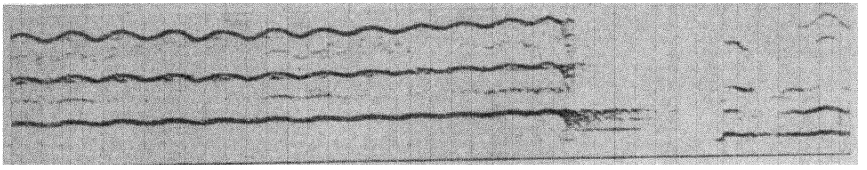


Fig. 38 — Lily Pons (soprano) "Spargi d'amaro Pianto" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-B (2-2/16 inches).

In Fig. 38 is shown an abrupt stop with some interesting characteristics. Notice that a downward blurring occurs near the point where the overtones cease, and that the fundamental carries through for some distance beyond the leading edge of the stop. Effects of this sort will doubtless be understood more clearly as the studies of such patterns are extended. The vibrato rate here is 7.5 cycles, and the swing ranges between 8.3 per cent at the start and about 4.2 per cent just previous to the stop. The average frequency is about 1100 cycles (C^2 sharp).

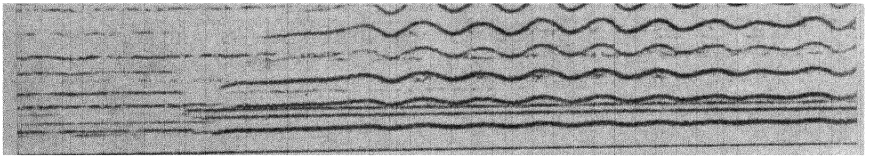


Fig. 39 — Helen Jepson (soprano); "Depuis le Jour" (Louise, Act III). RCA Victor Record No. 14204-A (2-3/4 inches).

The smooth resumption of tone following a prolonged stop, with only very slight evidence of pitch rise at the outset, is illustrated by Fig. 39. Here the vibrato starts almost exactly two cycles after the first small rise in pitch accompanying appearance of voiced tone. (Disregard the instrumental tone in the background.) Many other effects that accompany stops deserve more attention than this brief outline permits. Among the more interesting are pitch variations during brief stops, drops and subsequent rises

when voicing ceases momentarily, and influences that carry through short breaks when there is a rise or fall in pitch during the transition. Also, it will be noticed that the edges of stops often depart from the vertical, due to a change in resonance conditions accompanying closure. Sometimes low-pitched components carry through an otherwise complete closure, doubtless for the same reason that the voicing bar in visible speech carries through the gap introduced by "*b*" and "*d*" sounds.

Glides

Continuous elevation or depression of pitch, here called glides, appear in many of the illustrations so far shown. A decrease in pitch occurs toward the end of Fig. 3 with slight evidence of vibrato. Figures 8 and 9 both show a downward glide at the left with little vibrato until the pitch levels off. Toward the end of Fig. 10 is another downward glide with irregular vibrato superimposed. Fig. 18 shows still another with pronounced vibrato; and in Fig. 23, following the stop, is an additional example with similarly distinct vibrato swings.

Steps

A very rapid increase of pitch in steps is shown by Fig. 40. Unfortunately, instrumental accompaniment confuses the picture; but the voiced intervals may generally be identified by their up-curved shapes resembling the tops of a vibrato, as if even at these rapid rates the tendency toward vibrato persists. A detailed examination of such a passage in the absence of accompaniment would be very interesting, although even in these pictures certain details are recognizable. For example, the sharp attack at the beginning of each stop is evidenced by the flattening and spreading at these points. At the opposite ends, the tone dies quickly but smoothly. The voiced intervals are often only one-twentieth of a second or less in length, and the pitch variation seems to be similar, but characteristically different as the average pitch rises. The variations are apparently much greater, as a percentage, on the lower notes.

Steps in descent, at a more leisurely rate, are shown in Fig. 41. The instrumental notes may be easily identified since their fundamentals appear as the lowest, comparatively straight traces, and successive overtones may be located at multiples of the

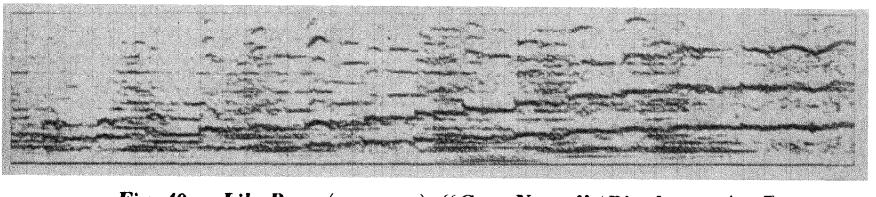


Fig. 40 — Lily Pons (soprano); "Caro Nome" (Rigoletto, Act I). RCA Victor Record No. 14203-A (2-14/16 inches).

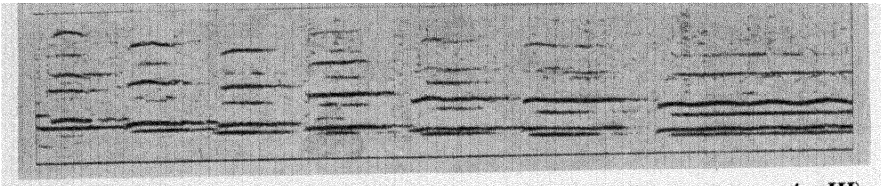


Fig. 41 — Lily Pons (soprano) "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (3-1/8 inches).

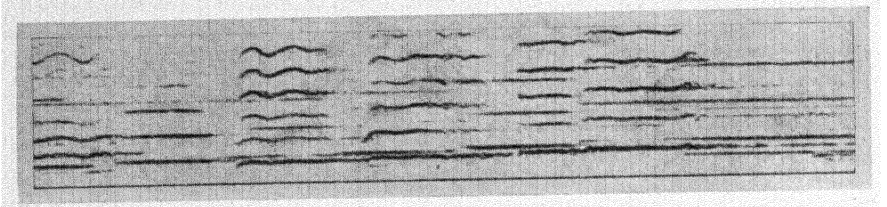


Fig. 42 — Lily Pons (soprano); "Ardon gl'Incensi" (Lucia di Lammermoor, Act III). RCA Victor Record No. 7369-A (6/16 inch).

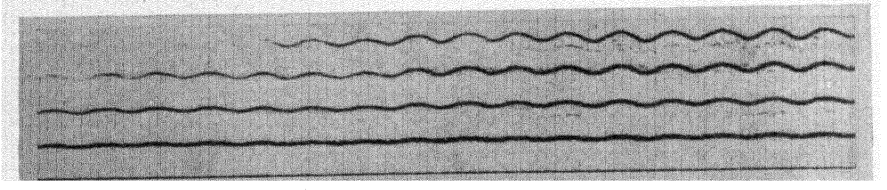


Fig. 43 — Lily Pons (soprano); "Caro Nome" (Rigoletto, Act I). RCA Victor Record No. 14203-A (1-15/16 inches).

fundamental's distance from the base line. In most cases, particularly toward the right, the voice appears about one-thirtieth of a second before the instrumental tone. In every voice attack, there is an upward arc at the outset followed by a leveling off of pitch for the remainder of the step (two- to three-tenths of a second.) When the sustained tone starts at the right, a similar upward arc continues on into a vibrato at about 7.0 cycles a second.

Steps in pitch accompanied by distinct vibrato, and some interesting pitch fluctuations at the breaks, appear in Fig. 42. Note the initial short upward arc carrying on into a vibrato swing in the two middle steps. Also notice the rapid pitch fluctuations at the end of the middle step and between the two steps at the right. The cause of these remains to be determined. They may be either in the voice, or due to beats with superimposed instrumental tone. Incidentally, the average frequencies for the four steps over the middle range of this sample are approximately 516, 590, 640 and 690 cycles.

Crescendo and diminuendo

Figure 43 illustrates a rise in volume, or a crescendo, as indicated by the graduated darkness of recording from left to right. A following decline in volume, or a diminuendo, appears in Fig. 44. The average frequency is the same in these two cases, about 820 cycles (G^1 sharp). In the crescendo, the vibrato starts at about 6.8 cycles and increases to 7.2 cycles toward the right. In the diminuendo, it starts at 7.2 cycles and falls off to 6.8 at low volume.

Effects of resonance upon overtone intensity with frequency swing over the vibrato cycle appear at low volume in the third harmonic of Fig. 43 and in the second of Fig. 44.

Falsetto

In Fig. 45, Tibbett produces a high pitched "*ee*" in "*Figaro*". Previous to the stop ahead of the vowel, the pitch averages about 260 cycles, (middle C). At the highest point during the vowel it reaches 660 cycles (E^1), and immediately following the subsequent stop the pitch drops to about 350 cycles (F), falling at the end of the sample to 220 cycles (A^1). The effect of tongue movement associated with the "*r*" sound of "*garo*" may be clearly seen as an irregularity in the pitch lines about one-fourth of a second beyond the edge of the second stop. The vibrato-like steps in the down glide following the "*r*" sound are also noteworthy.

The prolonged "*Figaro*" in the pattern of Fig. 45 should be compared with rapid repetitions of the word in Fig. 46. The first "*ee*" sound appears immediately following

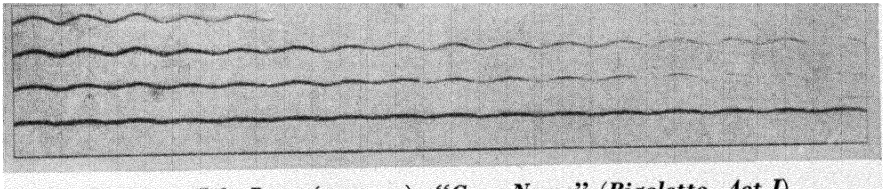


Fig. 44 — Lily Pons (soprano); "Caro Nome" (Rigoletto, Act I). RCA Victor Record No. 14203-A (2 inches).

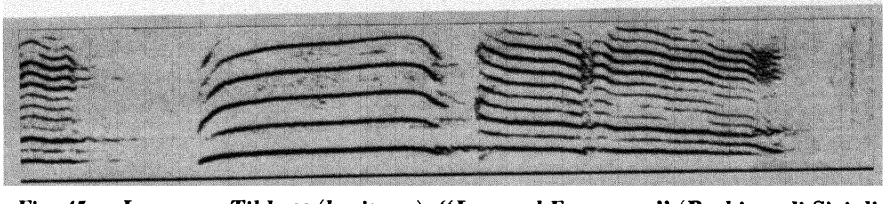


Fig. 45 — Lawrence Tibbett (baritone); "Largo al Factotum" (Barbiere di Siviglia, Act I). RCA Victor Record No. 14202-B (2-1/2 inches).

the first stop as an upward and rounded arc in pitch. The remainder of the word, including the "*r*" sound, will be recognized as a compressed version of the pattern in Fig. 45. In the three following repetitions the arc of the "*ee*" sound is bent to the right, and the time occupied by the complete word of three syllables is less than a quarter of a second. To the extreme right the somewhat prolonged "*o*" sound contains a distinct vibrato.

Both Figs. 45 and 46 may be compared with the unstressed "*Figaro*" of Fig. 19.

Word formation

There are many effects that deserve study in the modulations that accompany sung words. Even when words are sung rapidly, pitch variations, stops, and resonance shifts seem to retain their individualities while considerably compressed in the time dimension. This has already been illustrated by Fig. 46. Another example is in the isolated words by Caruso shown in Fig. 47. Here a distinct vibrato is superimposed upon each

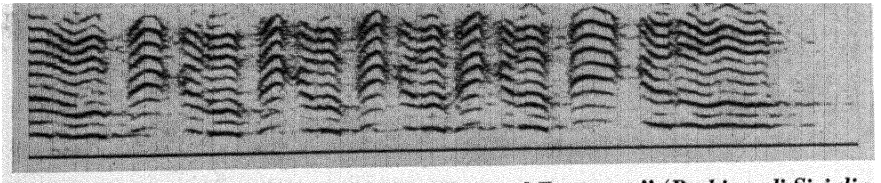


Fig. 46 — Lawrence Tibbett (baritone); "Largo al Factotum" (Barbiere di Siviglia, Act I). RCA Victor Record No. 14202-B (2-9/16 inches).

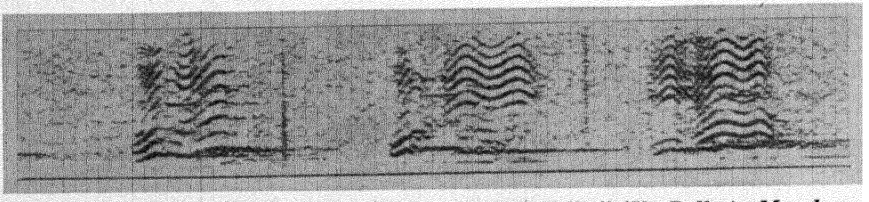


Fig. 47 — Enrico Caruso (tenor); "E Scherzo, od E Follia" (Un Ballo in Maschera, Act I). RCA Victor Record No. 16-5000-B (7/16 inch).

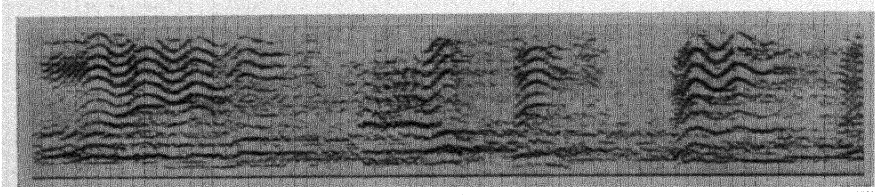


Fig. 48 — Lauritz Melchior (tenor); "Dich Selige Frau Hält Nun Der Freund" (Die Walküre, Act I). RCA Victor Record No. 14204-B (1-14/16 inches).

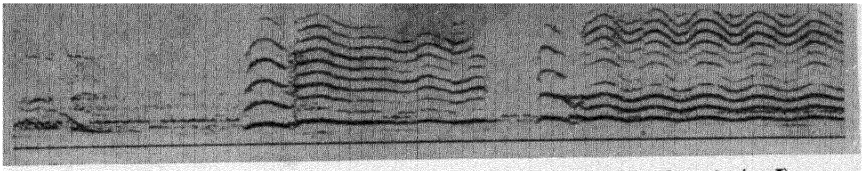


Fig. 49 — Enrico Caruso (tenor); "Vesti la Giubba" (I Pagliacci, Act I). RCA Victor Record No. 6001-A (1-3/4 inches).

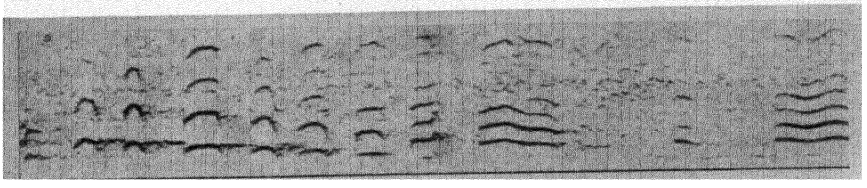


Fig. 50 — Enrico Caruso (tenor), Geraldine Farrar (soprano), Antonio Scotti (baritone), Gina Viafora (soprano); "Addio, Dolce Svegliare" (La Bohême, Act III). RCA Victor Record No. 16-5001-B (1-15/16 inches).

short voiced interval, and the volume rises in step with the vibrato at the beginning of each word, and falls off with it at the end.

Strongly enunciated word sounds with pronounced vibrato are shown in Fig. 48. Note especially the tongue trill near the center of this sample, and the abrupt pitch rise following the stop at the right. These are the words, "*Keim und Spross ents-preisst*", sung very rapidly, and the trill occurs on the word "*Spross*".

Sobs

A sob by Caruso is shown in Fig. 49. The initial "catch" in the voice appears here as a pitch shift somewhat similar to that in a yodel (illustrated later). The pitch change roughly approximates a doubling. There is an upward arc in pitch accompanying the sob that is similar to that of the "*ee*" sound in "*Figaro*" of Fig. 46, except that in Fig. 49 the arc leans to the left, indicating a rapid inflection. At the right, a sustained vibrato appears at a rate of about 7.0 cycles, with a swing of about 11 per cent.

Laughs

In Fig. 50, there is a high-pitched musical laugh by Farrar in which short bursts of tone appear on a descending pitch scale from about 970 cycles (B^1) to 420 cycles (A flat). An upward pitch arc is pronounced on the high tones and flattens out as the tone drops toward the right. Some of the bursts of tone are about as short as ever seem to occur in the process of voicing and unvoicing, i.e., about one-twentieth of a second.

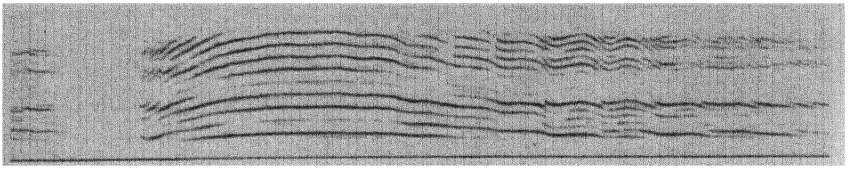


Fig. 51 — Lawrence Tibbett (baritone); "Song of the Flea". RCA Victor Record No. 7779-A (3-1/16 inches).

A prolonged "*ha*" with rising inflection followed by an interesting succession of steps appears in Fig. 51. The series of dark elongated S-shaped strokes that occur at about 1300 cycles give the impression of sudden breaks. Actually, sharp drops in pitch at these points cause overtones to sweep across what seems to be a very sharp resonance at this frequency. The sixth harmonic crosses this resonance ahead of the first "break" (in counting the harmonics notice that some at the bottom of the series are missing), the seventh crosses next, and so on until at the extreme right the fourteenth harmonic appears at this 1300-cycle point, meaning that the voice pitch has dropped from 216 cycles to 1300 divided by 14, or 93 cycles a second! This stepped vibrato has a rate of about 7.0 cycles initially that increases rapidly to something approaching more than twice to original value.

Pitch Wavers

The little "catches" that are familiar in the voices of some popular singers appear as single brief pitch swings in the sound spectrograms. Figures 52 and 53 illustrate the use of this effect by Bing Crosby. The former shows the way he has treated such a break in the word "*town*". Here, following the gap and the plosive "*t*", is a nearly flat pitch interval about 0.1 second long followed by a pitch rise that reaches a peak 0.2 second from the plosive. The upward swing from the average is nearly 20 per cent, and it covers an interval roughly 0.2 second long. The normal vibrato at the extreme left and middle right, following the pitch waver, has a rate of about 5.8 cycles.

In Fig. 53, from another song by Crosby, there appears a similar pitch waver following the plosive "*b*" in "*bonnie*", slightly to the left of the center. Here, also, the waver reaches its peak about 0.2 second after the plosive. In this case, the upward swing is about 25 per cent. Wherever the normal vibrato appears, the rate is about 5.8 cycles. It is interesting to note that these upswings very nearly match the cadence of the vibrato. In several other spectrograms of Crosby's voice, similar pitch wavers always seemed to occur very nearly 0.2 second after a stop.

An example of the same effect used by Ethel Waters is shown in Fig. 54. The upward swing in this pattern is a little over 22 per cent, and the peak of the swing again comes

about 0.2 second after the edge of the preceding stop. The vibrato, which appears to be a separate affair, builds up gradually over the following half-second. The vibrato rate is about 5.2 cycles a second, and the swing is on the order of 16 per cent. Bear in mind that the percentage swing of the vibrato is measured as the spread from the minimum to maximum pitch, while the upswing in these pitch wavers is only that from average to peak!

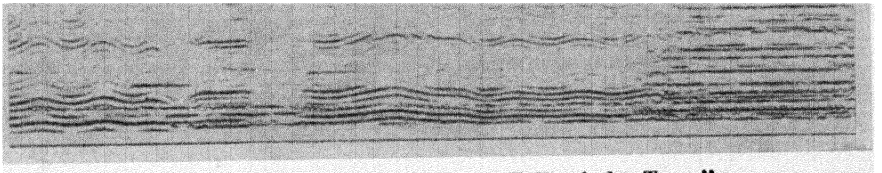


Fig. 52 — Bing Crosby; "It's the Talk of the Town". Decca Record No. 23469-B (1-11/16 inches).

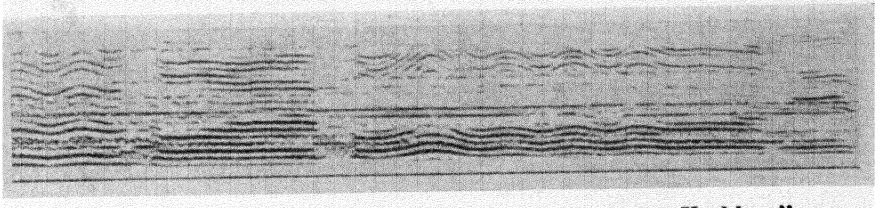


Fig. 53 — Bing Crosby; "I'll Take You Home Again, Kathleen". Decca Record No. 18721-B (14/16 inch).

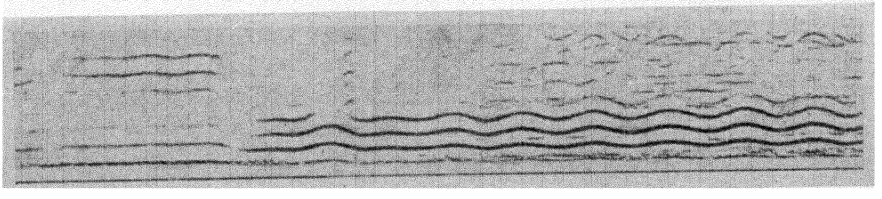


Fig. 54 — Ethel Waters; "Give Me a Heart to Sing To". Decca Record No. 141A (38548) (9/16 inch).

Other well-known voices

A few of the voices among those rated as having strong "popular" appeal for various reasons are illustrated in Figs. 55 to 62 inclusive. These samples are entirely insufficient to show the range of any one singer, but as a group they provide some indication of the variety that exists here.

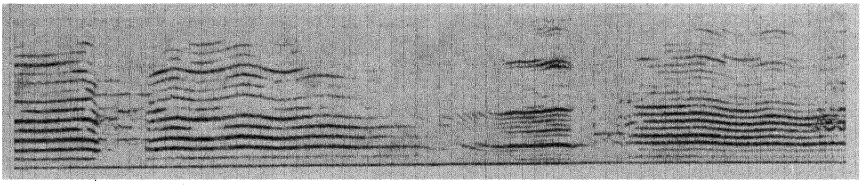


Fig. 55 — Bing Crosby; "It's the Talk of the Town". Decca Record No. 23469-B (1/2 inch).

In Fig. 55, from a Crosby record, the vibrato rate is about 5.0 cycles following the first stop, and it increases to about 6.0 cycles at the extreme right. The swing is around 7 per cent. The words are "*the talk of the town*". Note especially the sudden down sweep of an octave slightly to right of center, and the accentuation of resonances in this low-pitch region.

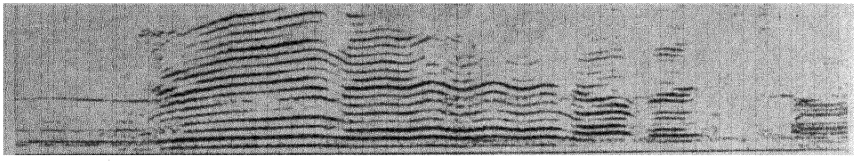


Fig. 56 — Frank Sinatra; "Star Dust". RCA Victor Record No. 27233-A (1/4 inch).

The vibrato rate near the middle of Fig. 56, taken from a Sinatra record, is roughly 6.5 cycles, and the swing about 11 per cent. The words here are, "— *dreamin' of a song*".

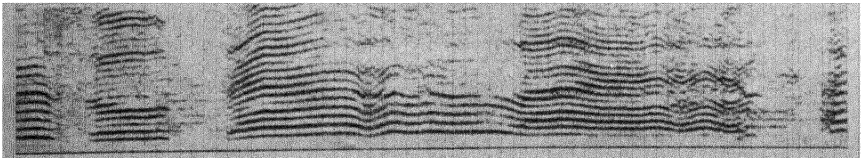


Fig. 57 — Johnny Mercer; "On the Atchison, Topeka and Santa Fe". Capitol Record No. 195 (518-1A) (1/4 inch).

In Fig. 57, from a record by Johnny Mercer, the vibrato is interrupted but in evidence through the words, "— *whistle down the line*".

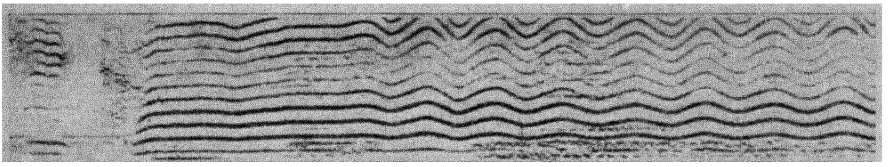


Fig. 58 — Judy Garland; "On the Atchison, Topeka and Santa Fe". Decca Record No. 23436A (72967) (10/16 inch).

A strong vibrato appears in the right half of the Judy Garland patterns of Fig. 58. Its rate is 5.3 per second and the swing is 16 per cent. Following the stop at the left by a little over 0.2 second, there is a slight pitch depression which might be related to the pitch waver discussed earlier. The regular vibrato starts about 0.5 second after the stop.

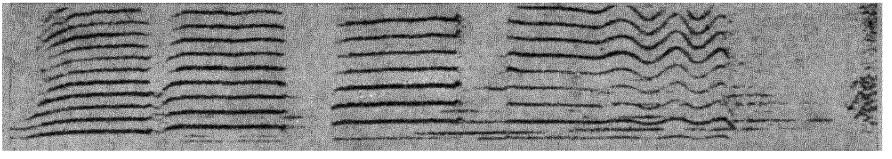


Fig. 59 — Ella Mae Morse; "Ya' Betcha". Capitol Record No. 193 (568-3A) (9/16 inch).

The pattern of Fig. 59, from an Ella Mae Morse record, represents the words, "— *worry 'bout me*". It shows some interesting stops and a sudden initiation of vibrato at the right, following nearly constant pitch.

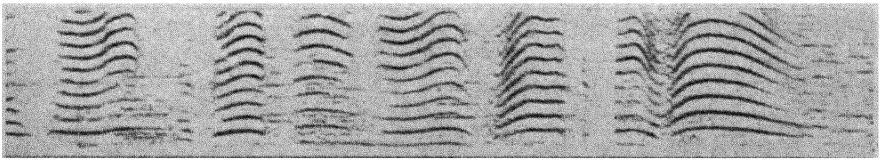


Fig. 60 — Jerry Colonna; "Bell Bottom Trousers". Capitol Record No. 204 (622) (2-3/16 inches).

Staccato effects in Jerry Colonna's singing are clearly illustrated by the patterns of Fig. 60, and there is also interesting variety in the formation of the stops shown here. The sharp edges of plosive stops and the rounded edges of the glottal type are well pictured.

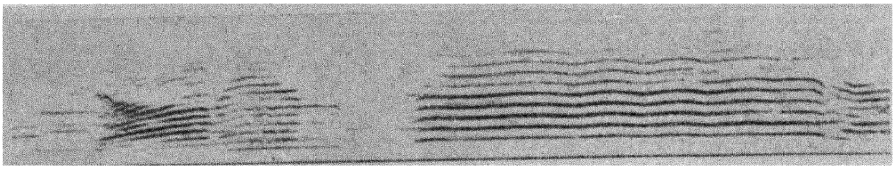


Fig. 61 — Gene Autry; "Don't Live a Lie". Columbia Record No. 36880 (H1427) (1/4 inch).

In Fig. 61, from a Gene Autry record, the vibrato rate at the right is 5.3 cycles and the swing is 4.0 per cent. The words here are "— *down inside*".

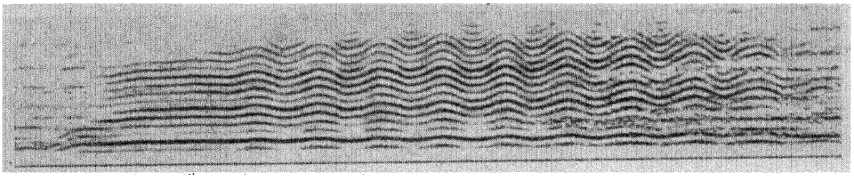


Fig. 62 — Al Jolson; "April Showers". Decca Record No. 23470-B (L3913) (1/2 inch).

Figure 62, taken from a recording by Al Jolson, shows a particularly regular vibrato carried on the sustained vowel of "*way*". The rate is 5.8 cycles and the swing 12 per cent.

Yodeling

The patterns of Figs. 63 and 64, from recordings by Swiss yodelers, illustrate the rapid pitch changes that accompany the yodel "break" and also its association with the vibrato. It is not evident in these patterns that the pitch relationships immediately adjacent to the break are simple. When the break occurs, there is apparently an influence taking the form of a pitch deflection in the direction of the subsequent shift, upward if the following tone is higher pitched and downward if it is lower.

When the pitch change is very abrupt, the action is blurred, as near the left end of Fig. 63. If the change is less sudden, there is appreciably less confusion and even a tendency to assume the appearance of a rapid glide. Examples of the near glide type appear to the left of center in Fig. 63, and toward the right in both Fig. 63 and Fig. 64.

At the right of center in Fig. 63, there is an irregularity in the vibrato that has the appearance of a yodel break, although the average pitch change is comparatively small. In this sample, the components carry through, somewhat as in the very brief partial stops described earlier.

A pattern of a yodel by a non-professional singer is shown in Fig. 65. The breaks are noticeably less regular than in the patterns of the professional, but pitch influences are very clear, in fact much exaggerated. The pitch ratio at the break points is approximately 3 to 5, or 5 to 3. Pitch shifts from low to high are abrupt, and there is evident difficulty in going from high to low pitch. In all the latter cases, considerable confusion marks the transition; and in one instance, toward the extreme right, a complete step is added, the pitch first changing in a ratio of 10 to 9 and then 5 to 3.

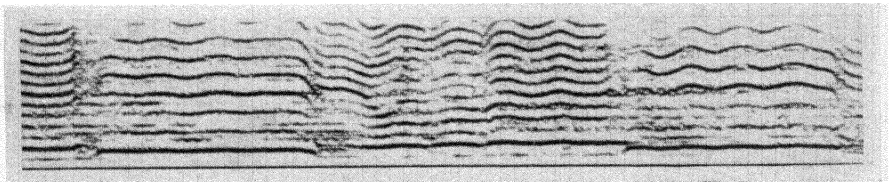


Fig. 63 — Geschw (Jodel); "Alpenklänge" (Ländlerkapelle: Die Vierwaldstättler). Elite Record No. 1857 (1747-11) (1/4 inch).

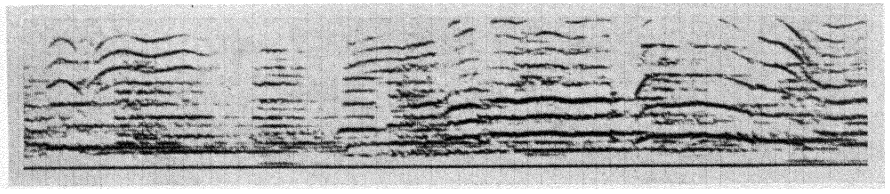


Fig. 64 — Zwyer (Jodel); "Rigi-Jodel" (Landlerkapelle: Die Vierwaldstättler) Elite Record No. 1857 (1750) (3/4 inch).

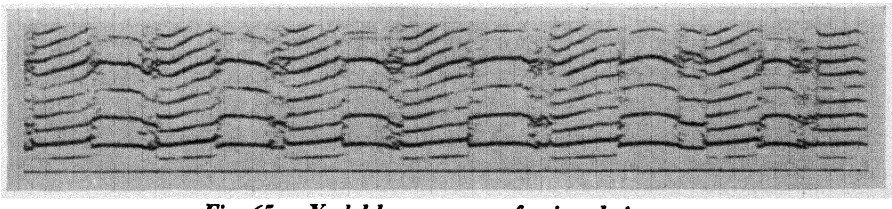


Fig. 65 — Yodel by a non-professional singer.

Whistling

Examples of whistling patterns appear in Figs. 66 to 69 inclusive. Fig. 66 shows a whistled vibrato at a rate of about 6.8 cycles and a swing of 12 to 15 per cent.

In Fig. 67, taken from a record by Lowery, is pictured an abrupt pitch shift between two tones, as contrasted with the smooth sinusoidal variation of Fig. 66.

A rapidly recurrent down sweep in pitch, also from a Lowery record, appears in Fig. 68. These successive sweeps occur at a rate of approximately 40 a second, and are not individually distinguishable by ear.

A whistled imitation of the cardinal is represented by Fig. 69. In comparison with the actual cardinal song, there are some significant similarities and differences. Presumably the ear is less critical than the eye in such a comparison. Whistled patterns are fascinating in many respects and it would seem that, with the aid of spectrograms, a book might be written on this subject alone.

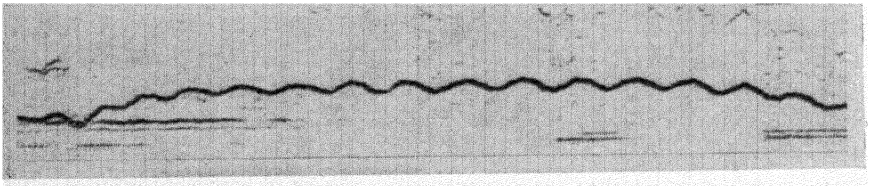


Fig. 66 — Whistling recorded directly from a radio program.

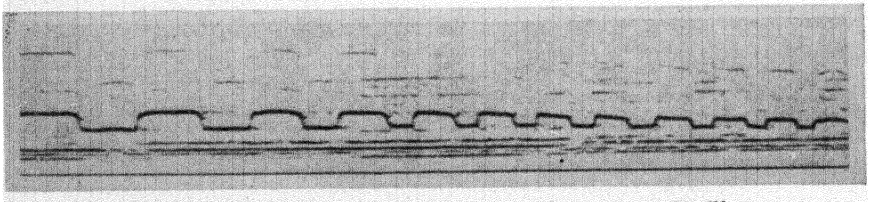


Fig. 67 — Fred Lowery; "Listen to the Mocking Bird". Columbia Record No. 35234 (WB 24940) (1-1/4 inches).

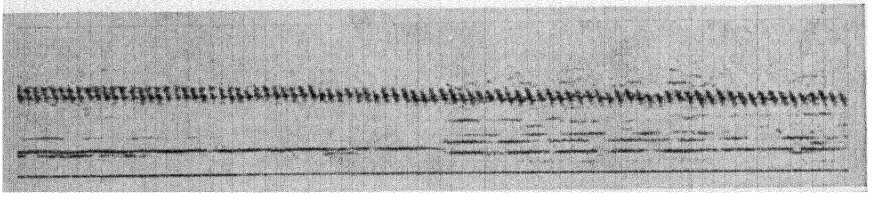


Fig 68 — Fred Lowery; "Listen to the Mocking Bird". Columbia Record No. 35234 (WB 24940) (15/16 inch).

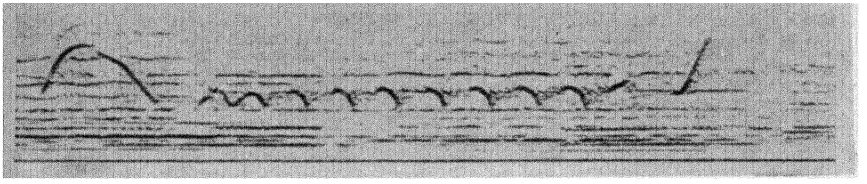


Fig. 69 — Fred Lowery; "Listen to the Mocking Bird". Columbia Record No. 35234 (WB 24940) (2 inches).

Other vocal music interests

A use for sound spectrograms that needs to be touched upon is that concerned with studies of the songs of native groups. Very little attention has been devoted to this possibility, due largely to lack of time, material, and experience in this field; but it would seem that both analysis and discussion of such music could benefit by the use of these sound portrayal techniques.

An example of primitive music is shown in Fig. 70. This was obtained from some recordings turned over to us for analysis by the Laboratory of Comparative Musicology, Department of Anthropology, Northwestern University. The description is indicated in the caption. As the patterns show, it is predominantly the voice of one singer punctuated by regular drumbeats and apparently some handclapping. The complete record includes one minute of recording.

Unfortunately, the arrangement most suitable for occupying the page probably does not show the variations to best advantage. It will be noted that there are five broad vertical divisions divided off by the accentuated drumbeats that appear as vertical spikes. The last division to the right carries over into the first part of each row at the left. Repetitions may be readily recognized. The sounds in the first large interval in the upper left corner are repeated two intervals later. Also, the whole sequence is repeated 44 intervals later, although a few minor variations may be seen in the repeated passages. It is significant that although there are similarities between intervals outside the repetition cycle, these similarities are generally of a much lower degree. No doubt particular singers could be readily identified by their pattern habits.

Pattern requirements

Sound portrayal is in some ways like microphotography: the method of portrayal depends much upon what one wishes to see. If the interest is in scratches on the surface of metal, the microphotograph would take a form entirely different from that used to show crystal structure or molecular arrangement. Similarly, the requirements for picturing the sounds of music will depend upon the things that study suggests need emphasis in these pictures.

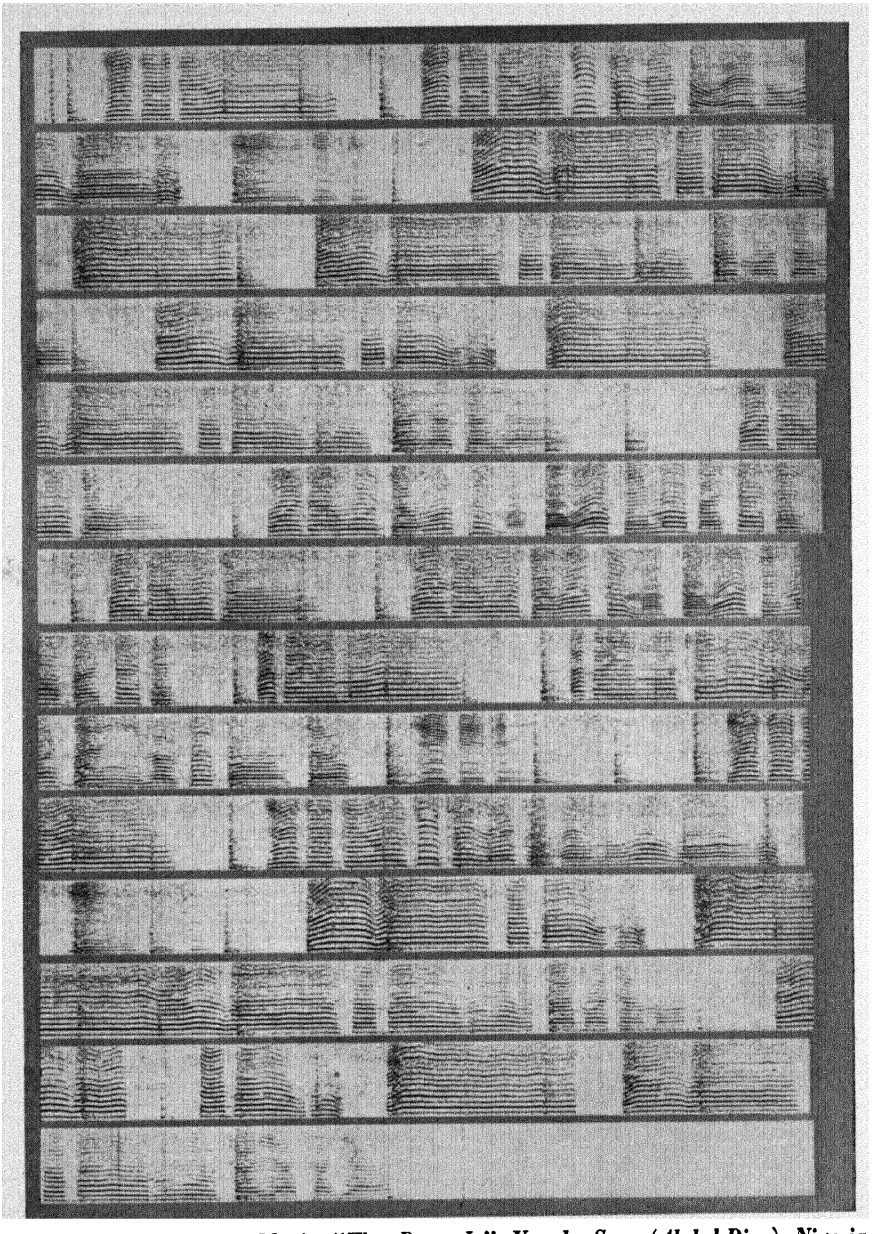


Fig. 70—Primitive Music; "Wan Dawo Jo", Yoruba Song (Abdul Disu), Nigeria. Laboratory of Comparative Musicology, Department of Anthropology, Northwestern University, Side B, Section 3.

As students of vocal music become better acquainted with sound spectrograms, they will doubtless recognize deficiencies in present methods of portrayal as applied to their particular problems, but it is important to bear in mind that these deficiencies are not necessarily inherent. They are matters of instrument design. Ways of picturing sounds to satisfy the needs of investigators in various fields of vocal music can doubtless be developed, but only if the requirements are determined and made available to those concerned with sound spectrograph design and improvement.

At this stage, it is of considerable importance to encourage a broad study of sound portrayal principles as applied to vocal music, rather than to narrow the question down to one of how useful present experimental types of sound pattern instruments may be as aids in the solution of vocal music problems.

Chapter 18
Generalities and Speculations

Visible speech has reached its present stage of development due largely to interests concerned with telephony. One of these interests is in the future possibility of visual telephony for the deaf. Another is in sound portrayal generally, as a technique for use in telephone research and development.

A great deal remains to be done before visual telephony can be realized in a practical way. Translators will have to be considerably improved, simplified, and adapted to the telephone, and there must be training centers where those who are interested in using such instruments may learn to read the patterns. Development work on translators will continue in the Bell Telephone Laboratories, and, it is hoped, elsewhere, but training facilities will need to be established in the educational field, and the ultimate provision of such facilities will necessarily require an extensive preliminary program. Initial steps in such a program have been taken and the more immediate objectives were outlined and discussed in Chapter 13.

Non-speech interests

Non-speech uses of sound portrayal are likely to become more numerous in the future. They will no doubt contribute indirectly to improvements in visible speech through the search for better ways to translate and display such complex wave shapes. Examples of a few of these possibilities follow. Some of them are quite impressive, although comparatively little thought and effort actually have been applied to any such uses thus far.

Animal, bird, and insect sounds

A little experience with the patterns of animal, bird, and insect sounds suggests that there may be fascinating possibilities of analysis, illustration, and discussion that have not been available in the past. For example, permanent patterns of bird songs can be examined in great detail, and there is reason to believe that song differences within one group could be readily identified by visual comparisons. In fact, it would not be surprising if the song habits of individual birds could be recognized by close examination of the patterns. As was explained in the previous chapter, human singers repeat characteristic passages with considerable precision, and it may not be unreasonable to expect individual song habits to be similarly distinguishable among the feathered songsters. The song becomes a signature!

If detailed analysis of song patterns is possible, there would seem to be a wide new field of study open to the ornithologist. Perhaps bird books and periodicals of the future will be filled with song pictures, and serious readers may become well enough acquainted with this sound language to read visible patterns of bird music in the way a musician reads a musical score.

Little attention has been given thus far to the patterns of animal sounds, although here also would seem to be opportunity for extensive study. Unlike the songs of birds, that are usually made up of clear, high notes, with colorful variety in rapid excursions of pitch, the sounds of animals (at least the larger ones) would be expected to contain many overtones and conspicuous resonance shaping. Certainly, these animal sounds would present a type of resonance modulation and pitch inflection different from that of the human voice, although there may be interesting comparisons found between human sounds and those produced by the anthropoid apes. Possibly a study of these could contribute a new chapter to the story of evolution.

In Fig. 1 are shown several bird songs, the sounds of barking dogs, and the crying of an infant. This illustration is one from the original article on sound portrayal.* In the first four patterns (A to D inclusive), the length is approximately 4 seconds and the frequency scale is linear between zero and 7,500 cycles per second. In the last four (E to H inclusive), the length is about 9.4 seconds and the vertical frequency scale is zero to 3,200 cycles per second. Parts of these patterns disclose movements that would doubtless be difficult to describe in terms of what the ear hears. The abrupt downsweeps in the early part of the cardinal's song and the similarly steep up-sweeps in the latter part to the right are illustrations.† So also are the curiously irregular but accurately repeated warbles in the robin's song, some even birdlike in shape. And in the mocking bird's melody, what appears at first to be a smudge of gray toward the left, actually appears, on closer examination, to be rapid repetitions of almost vertical pitch changes. The brown thrasher's pattern reveals a less rapid up-and-down variation in pitch toward the left, and in the first part of the three notes toward the right will be found a "double tone" sound in which two tones without harmonic relation are evident for a short interval. Presumably these must originate in separate parts of the tone-producing mechanism.

Patterns of both audible and inaudible insect sounds should also open up an extensive new field for study. It is conceivable, for example, that insect groups could be classified by their "songs" or noises. A superficial examination of a few recorded mosquito sounds indicates that this might be possible because considerable detail is present, even in these small tunes.

---

* "Visible Patterns of Sound" by Ralph K. Potter; SCIENCE, Vol. 102, pp. 463-470, Nov. 9, 1945.

† A comparison of this actual cardinal song with the whistled imitation in Fig. 69 of Chapter 17 illustrates the order of approximation permitted in such mimicry.

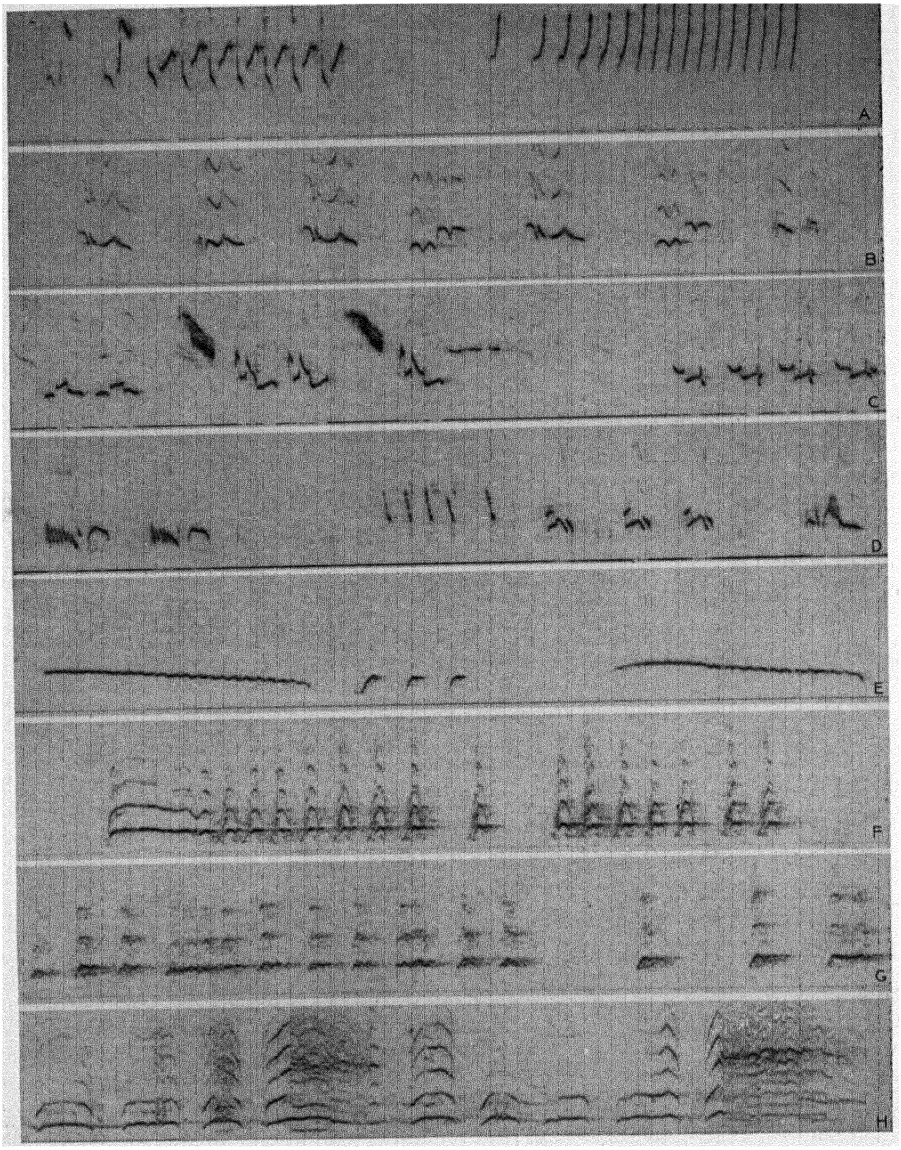


Fig. 1 — Sound patterns of: A, cardinal; B, robin; C, mocking bird; D, brown thrasher; E, screech owl; F, large dog barking; G, small dog yapping; H, baby crying.

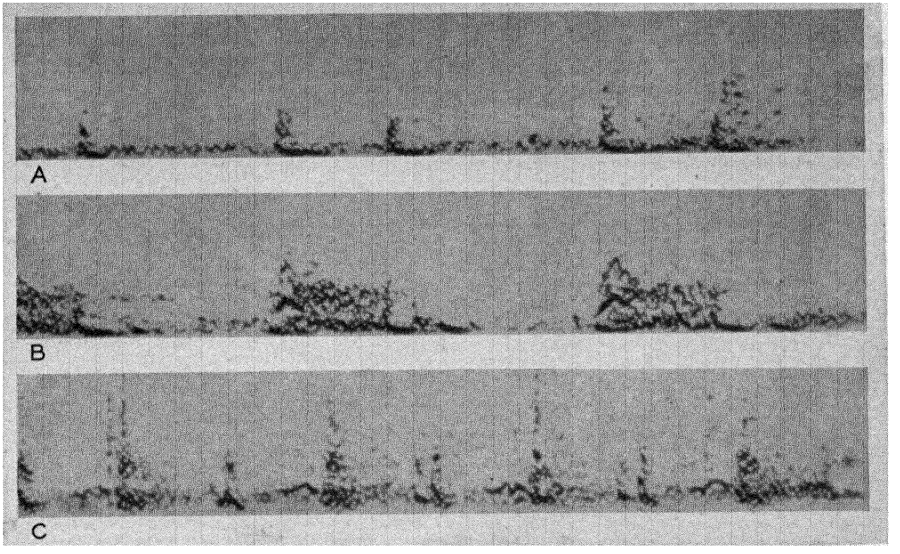


Fig. 2 — Heart-beat spectrograms: (A), normal heart beat, Columbia Record No. 11007-M, band 1; (B), musical systolic murmur, Columbia Record No. 11007-M, band 5; (C), gallop with one sound replaced by murmur, Columbia Record No. 11002-M.

Possible medical interests

One body sound of principal interest to the medical field is the heart beat. Spectrograms of heart-beat sounds may find use in discussion and illustration, if not also in diagnosis. Such patterns can show small variations that it would be difficult, if not impossible, to identify by ear. Most important, perhaps, is the fact that sounds in these pictures may be talked about in terms that are specific and meaningful. They assume a realness and dimensions as significant as those of solid tissue.

Sample patterns of heart beats obtained from available phonograph recordings* are shown in Fig. 2. In these heartbeat spectrograms the time scale is the same as that illustrated in Fig. 1 of Chapter 17, but the vertical frequency or pitch scale is doubled, so that for the same distance from the base line the actual tone is only half as high in pitch as in the speech and vocal music spectrograms. Most of the meaningful action in these pictures is presumably in the low-frequency range, and it is likely that the portrayal could be improved by different scale arrangements. It should be mentioned that in all the non-speech illustrations under discussion, the display of information

* "Stethoscopic Heart Sounds" (George D. Geckeler, M.D.) Columbia Records.

might benefit considerably by the use of different analyzer characteristics and different display aspect ratios.

Pattern A of Fig. 2 is a norma heart beat. Pattern B shows a musical systolic murmur as a brief dark tone trace in the background of noise at the beginning of the two cycles shown. The "gallop" in pattern C appears in the pattern to be just what the name would imply, since the dark sound traces rise and fall at a rapid rate. It is conceivable that minor details, which would not be noticed in a casual inspection of these pictures, could be shown by analysis to contain important information.

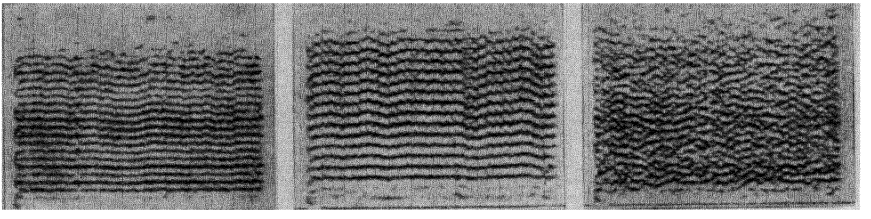


Fig. 3 — Spectrocardiograms: (a), normal heart; (b), abnormal heart; (c), abnormal heart.

An entirely different type of heart-beat spectrogram has been made in which electrical potential variations associated with muscular action in the heart are recorded on a slowly moving magnetic tape that is subsequently speeded up 100 times and played back into the sound spectrograph. Totally inaudible oscillations of a few cycles per second are thus converted to audible tones and treated as sound. Spectrograms of this type are shown in Fig. 3. The time interval in each short sample is 40 seconds and the vertical scale is zero to about 35 cycles a second. When the action is completely regular, as in the normal heart, the "overtones" of the beat rate appear as clean-cut horizontal lines. As irregularities appear, these overtones degenerate into a random spectrum that has the appearance of noise. Many details in these patterns are doubtless significant, but such a form of display is so different from the familiar electrocardiogram that its evaluation would require considerable study.

Other body waves, such as those known popularly as "brain waves" and recorded as encephalograms, may be spread out in sound spectrogram form by means of slow-down or speed-up techniques similar to those described above. Samples of normal and abnormal brain potential variations have been obtained from phonograph records submitted by an outside group for spectrographic recording, but thus far there has been no report concerning results.

Instrumental music

Possibilities of sound portrayal in the case of instrumental music seem attractive, but this vast field remains almost entirely unexplored. One thing appears certain at the outset — for most purposes the linear pitch or frequency scale is inadequate, although for some detailed analyses this scale is probably satisfactory. In Fig. 4 is shown, for example, the action in a violin passage approximately two seconds long from a recording by Jascha Heifetz. The two vertical rows of dots along the left edge identify the overtones in two note groups. Observe the rapid shifts in tone that may in particular be followed in the lowest trace. Notice also the way the third and fifth harmonics beat together forming an uneven trace at the beginning of the record, the abrupt juncture slightly to the left of center, the momentary rise in pitch to the right of center, and the occurrence of what would seem to be a free string oscillation where the vibration is sustained after a shift of the principal tone.

Where interest is in the interrelation of many notes spread over the range of musical movement, limitations of the linear scale become evident. Here the requirements are that the patterns correspond in general with what the ear hears. Middle C on the piano keyboard, at 262 cycles a second, is not far from the base line on the linear scale. Consequently, all action in the lower half of the keyboard range is crowded into a small fraction of the total display range when using this type of display. The answer is probably a logarithmic frequency scale, although such a dimensional treatment produces patterns that are quite unlike the linear type in many respects. In fact, it is necessary to "shift gears" mentally when thinking of one and then the other, much as when using first one language and then another.

In a sound spectrogram with a linear frequency scale, successive overtones of a musical note spread upward and contract downward, accordion-like, with changes in pitch, while on the logarithmic scale, the whole overtone group moves up and down with pitch variation as a fixed pattern. Furthermore, in the logarithmic case, the overtones do not appear at regular intervals as they do in the linear display. They are spaced logarithmically, the lower components being far apart and the upper ones close together. Also, in the logarithmic display, it becomes essential to vary the bandwidth of the analyzing filters with pitch, if the resulting pictures are to be in proper agreement

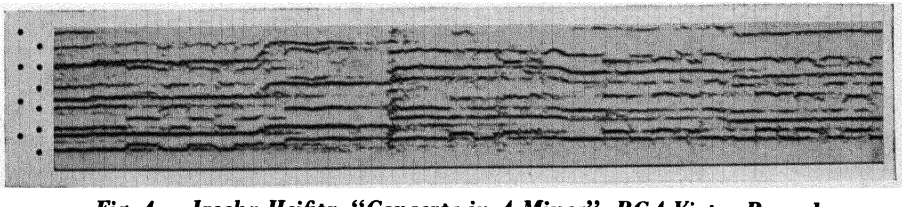


Fig. 4 — Jascha Heifitz, "Concerto in A Minor". RCA Victor Record No. 17096-B (1-5/8 inches).

with what the ear hears in the matter of frequency and time resolution. An approximation of aural equivalence results in a progressive reduction of trace width as a tone is raised in pitch. At low pitch, the trace is wide and heavy, while at high pitch, it becomes comparatively narrow, and finally merges with other overtones where resolution is insufficient to separate successive components.

Figure 1 of Chapter 14 illustrates a very early logarithmic-type pattern of speech. No suitable records of musical tones are at present available in this form, but this illustration will provide some indication of the distinctive characteristics of such a display, since the voiced parts of speech are similar in many respects to tones of musical instruments.

Musical notation

Several persons have asked whether visible patterns of music might not be used as a modernized form of musical notation. Possibly they could, but far too little is known as yet about this visible music to support more than speculative conclusions.

Another question raised is whether visible patterns of music could stir emotions in the way our feelings are stimulated by audible patterns. There has been far too little experience with these patterns to attempt an answer to such a question. It could be that a musician trained to recognize the movement in these pictures of music would feel a thrill upon seeing new and exciting musical patterns, but of course such a reaction suggests aural imagination. Whether there are aesthetic possibilities in the patterns themselves remains to be determined.

Sound patterns might prove very useful in studies of musical instruments, and in their tuning, both individually and in groups. Improved techniques of sound portrayal could perhaps make it possible to establish the characteristics of tone that identify one instrument as outstanding and another as short of this distinction. For such purposes, much better resolution and display than that available in the present experimental equipment would no doubt be required.

Limited experience with visible patterns of instrumental music does seem to suggest that the aspect ratio for such portrayal ought to be quite different from that for speech. Indications are that the musical patterns should be spread out much less in time, perhaps by a factor of two or three or more.

Artistic use of patterns

Patterns of sound take interesting and intricate shapes that have suggested to several observers possibilities in the field of design. It has been proposed that the black and white as well as the colored forms, such as are shown in the frontispiece, might make interesting material as the basis for fabric or wall paper designs. One specific proposal was to use the songs of different birds in different colors and sizes. This also is a field that has received practically no attention except in speculative discussion.

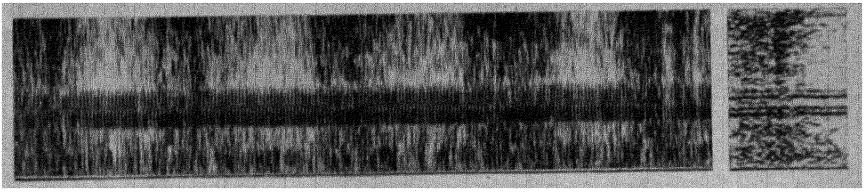


Fig. 5 — Spectrograms of machinery noise made with broad-band and narrow-band filters.

Mechanical vibrations

Mechanical vibrations, either in, below, or above the audible range, may be pictured in their basic frequency-intensity-time relation by variations of the techniques so far described. Although these possibilities also have been explored very superficially, it would seem that there might be many uses for such vibration pictures, in analysis and illustration, and as records for reference. Perhaps spectrograms of audible, super-audible, or subaudible variations will be more revealing than those obtained by present methods of recording. Perhaps also it would be possible for one to see in the visible patterns of machine sounds, things that only the trained ear can at present detect in the sound itself.

In Fig. 5 is shown a sample of sustained machinery noise. At the left, the pattern is of the broad-band, 300-cycle type, and at the right the short sample is the narrow-band, 45-cycle variety. Notice the four tones running through the lower middle section of the pattern, and the regular surges of high-pitch noise-like sound toward the top. Also near the bottom is faint evidence of a partially sustained note.

Patterns of the future

Although there are many potential uses of sound portrayal techniques that should help to further the development of visible speech, the solutions for a number of problems will remain primarily the responsibility of those directly concerned with speech. Contributions of non-speech developments would be mainly in the direction of better ways to take speech sounds to pieces, but finding the best ways to display the pieces to advantage seems certain to remain a problem for solution by the speech group.

Present forms of visible speech are far from perfect, and it is to be expected that speech pattern displays will be improved materially in the future. Such improvements are not likely to produce strange new patterns, but rather more clean-cut and readable versions of those we know at present. In these new forms, important variables would be more conspicuously displayed, and detail of minor importance minimized or deleted. Detail that contains no intelligence, or that cannot be utilized due to the conditions of viewing, is detrimental and obviously should be removed. Components that are useful

ought to be shown to their best advantage, within the limitations of viewing and the characteristics of the display surface. The major problem will be to decide what parts need high-lighting and what ought to be removed, and the treatment will vary depending upon the field of application.

Considerations of use

Watching visible speech on a translator screen is somewhat like watching a rapidly passing landscape. It is hopeless to attempt to see in such a moving scene any substantial amount of detail. In examining small items under these circumstances, the tendency is to concentrate momentarily on single features and visually hop from point to point as the scene goes by. To watch both detail and broad outline simultaneously would be a doubly difficult performance since this requires focusing alternately on one and then the other. The time and effort occupied in following detail would detract from the ability to discern broad outline, and vice versa.

In visible speech, the broad outline contains the intelligence, and the detail contains those variables referred to collectively as quality. If the concern were only with reading speech patterns, quality effects could be erased without loss, and even to advantage. A visible speech translator of hypothetical design, arranged to produce such completely impersonal patterns, would be analogous to that well-known translator combination, the stenographer with a typewriter. This translator would be capable of turning speech into standardized patterns that would look the same as long as the machine could understand what was being said. To complete the analogy, this imaginary machine would need to exercise judgment where the sounds were uncertain, and to request repetition when necessary.

Simplicity in non-standardization

Actually, any machine that is required to convert things into some standardized form becomes complicated. In a mechanical sense, the stenographer with the typewriter represents a very complex apparatus capable of sorting items that depart in many respects from standard forms. It is much easier to construct a machine that does its sorting according to a simple formula. If the rules for translation can be made elementary, the translator will be correspondingly simplified.

If visible speech patterns were to be used only for the purpose of reading,* or revealing intelligibility to the exclusion of quality, the reduction of extraneous material

* Incidentally, there is a curious inconsistency in present attitudes toward spoken and written languages. Personality in speech is apparently important but personality in print is not. The idea of hearing everybody talk in the same standardized voice seems to impress the average person as completely undesirable, but at the same time written words made of symbols that come from a standard mould are evidently quite acceptable. Handwriting is hardly a part of this argument since the personality expressed therein can be related only remotely to that heard and seen in speech.

would suggest a completely impersonalized pattern, but practical considerations in translator design would tend to discourage anything approaching an extreme in this direction. As a result there would probably be some optimum in design falling well short of complete standardization of pattern form. That is, the most simple and compact translators of the future, even for reading alone, are likely to show speech patterns in a form that is considerably influenced by the way in which the speaker talks. A translator that is equivalent to a stenographer with a typewriter will, without much doubt, always be more complicated than one that only writes down what it receives just as it is received.

Quality detail

Some of the most important uses of visible speech will require a display of quality in the speech patterns. In these cases, standardization, to an extent that is accepted in ABC's, would make the patterns worthless. The deaf, for example, not only want to see the intelligence in speech; they want even more to be able to see the quality in order to improve their speech. Perhaps there may be details that are unimportant and that are undesirable for this speech education purpose, but certain others will have to be retained. In learning to talk normally, by comparing one's own speech with that of others, it is not necessary to carry imitation to the extreme of impersonation. Inflection, stress, timing, and resonance control are important, but absolute pitch, vibrato, and other effects concerned more with individuality than general quality, may be unessential.

In studies dealing with phonetics, speech correction, or language, impersonalized visible speech would also be of little or no interest. A minimum requirement in these fields would be speech patterns capable of identifying speech habits, and a maximum would be pictures to show all the variables that can be found. Fine shades of distinction between sounds will necessitate a great deal of detail, which would be permissible in permanent patterns to be analyzed at leisure.

Speculation

Where are improvements in visible speech likely to lead? Are there speculative possibilities that it might be worth while to keep in mind? There is every reason to be practical about the present, but it rarely pays to discourage imagination when looking into the future. Speculation can be valuable as long as it is not confused with fact.

Of the many questions asked about visible speech, the majority are concerned with immediate needs, but a few imaginative individuals are inclined to look ahead:

"*Are 'voice writers' likely to be practical as a substitute for typewriters?*"

"*Will it not be possible to reconvert symbols of visible speech into audible speech?*"

"Is there a possibility that visible speech might be the basis for a new alphabet, which could some day replace the ABC's?"

"Would not a modernized written language, in visible speech form, assist in a solution of the problem of international communications?"

Voice-operated typewriters could doubtless be constructed to write in visible speech, because this would not be beyond the range of reasonable extrapolation from what has been accomplished thus far. Also, we know that symbols of visible speech may talk back since they have already been made to do so in fairly intelligible fashion with comparatively crude experimental equipment. Although there may be use for such a playback, the demand would probably be limited unless there proves to be a more widespread interest in visible speech than that concerned with the specialized fields discussed in previous chapters.

There is a natural temptation to handle the remaining questions, that dare to doubt the permanence of our ABC's, in rather gingerly fashion. Any proposal to abandon these old friends classifies the proposer as something between a non-conformist and a nincompoop. And this is not without reason, for we have had these same ABC's for centuries. A vast amount of thought is stored in their symbolic patterns, and our current existence depends heavily upon their use.

Nevertheless, there is a question behind these questions that deserves bold statement. Are our present ABC's likely to become obsolete in the sense that they will prevent the most effective application of extensive and evergrowing technical developments? Is it possible that, as a result of the tremendous strides of science, this ancient alphabet is becoming a handicap to technical progress in the use of written languages? Will future attempts to apply new and remarkable electronic techniques to the ABC's assume some of the aspects of trying to apply jet propulsion to oxcarts? The answers are not evident, but certainly it is evident that the question is a serious one, deserving thought and study. Nor is it a question that can be settled by either the expert in languages or the expert in electronics alone. It will require the cooperative and specialized consideration of both, along with the generalized consideration of all those seriously interested in reading and writing. Furthermore, the answer may be different tomorrow than it is today, because technical advance will continue on and on while the ABC's seem destined to remain just ABC's.

Apparently, though, it is not out of the question for a new form of writing to come into general use, providing that it shows substantial advantages. There is one case in which radically different forms of writing have replaced the ABC's, and that is where speed stands high in importance. To meet this requirement, new systems of shorthand have been put to use in a comparatively brief time, and without anything akin to a social revolution.

Basic considerations

What might be the basic advantages of visible speech in a comparison with ABC's? It is far too early to answer this question confidently and completely, but the following considerations deserve mention: Visible speech is a form of writing

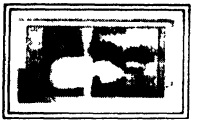
— belonging to the fundamental language of sound;
— following implicitly the primary language of speech;
— in which sounds that are new and strange to writing may be introduced at will and in which *all* sounds, either with or without symbolic meaning, may be inter-mingled freely as in speech;
— wherein the spelling of new words follows painlessly upon their introduction into the spoken language, and in which spelling is automatically simplified land earned in the process of learning to speak;
— that is universal, and in which the sounds of foreign languages appear as they would be heard by the ear;
— that permits the use of voice writing as a substitution for handwriting;
— that allows direct printing of oral material, such as a speech or a news report, so that it would be available for subsequent reading;
— that provides in print the kind of personality that exists in speech;
— and that makes possible the conversion of printed material into intelligible sounds.

Evaluation

Final evaluation of visible speech for use in specialized and general fields will require a great deal of experience and study. There seems no reason to doubt that the development will find widespread application; the question is mainly that of where it can best be applied, and how, and also what the requirements will be for the various uses. The purpose of the present book is primarily that of turning over, to all those who will carry on this work, the major results of the development to date, including a great deal of the past thought on the subject. It is hoped also that the book will provide opportunity to satisfy interests before commercial equipment for making sound patterns will be available, which unfortunately may be a substantial period, because there are numerous practical problems involved in producing designs suitable for manufacture.

While there are many potential interests, the important proving ground for visible speech seems certain to be in the field of deaf education. If and when it is found that deaf youngsters can learn to speak intelligibly and to read improved forms of visible speech with the same facility that they do their ABC's, and to do things with this form of writing that are practically impossible with the ABC's, it will be time to consider seriously questions of more widespread use. What happens during the coming years

in the field of deaf education can have a considerable effect upon the future of visible speech. If it comes into general use as a voice-written language for the deaf it could even start a trend toward modernized writing and printing. The deaf may lead the way!



APPENDICES

BIBLIOGRAPHY

INDEX

Appendix I

Reference Listing of Phonetic Symbols, Key Words, and Webster's Diacritical Markings

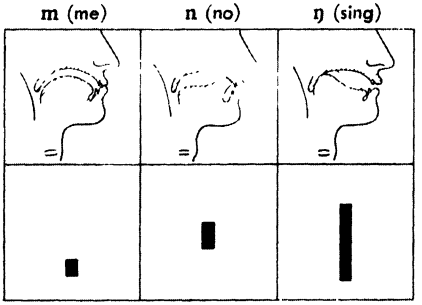
Phonetic Symbols	Key Words	Webster's Diacritical Markings	Phonetic Symbols	Key Words	Webster's Diacritical Markings
i	eve	ē	t	to	t
ɪ	it	ĭ	d	day	d
e	hate	ā, a̱	k	key	k
ɛ	met	ĕ	g	go	g
æ	at	ă	h	he	h
a	ask	a	ɦ	ahead	—
ɑ	father	a, ŏ	f	for	f
ɒ	not	ŏ, o̊	v	vote	v
ɔ	all	ô	θ	thin	th
o	obey	o̱	ð	then	~~th~~
ʊ	foot	o͝o	s	see	s
u	boot	o͞o	z	zoo	z
ɚ	word, bird	û, ẽ	ʃ	she	sh
ɜ	word, bird	—	ʒ	azure	zh
ʌ	up	ŭ	tʃ	church	ch
ə	about	ă, ĕ, ŏ, ŭ, ȧ	dʒ	judge	j
eɪ	say	ā	m	me	m
aɪ	I	ī	n	no	n
ɔɪ	boy	oi	ŋ	sing	ŋ, ng
aʊ	out	ou	w	we	w
oʊ	go	ō	j	you	y
ɪu	new	ū	r	read	r
p	pay	p	l	let	l
b	be	b	hw	when	hw

Appendix II

Summary of the Profiles and Hub Areas of the Non-Vowel Sounds in a Male and a Female Voice

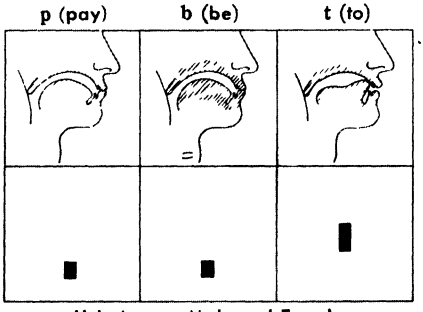
The diagrams on the two following pages have been abstracted from material included in Chapters 8 through 11 to facilitate study of the relationships among hub areas and the generally accepted articulatory positions illustrated in the profiles. Two profiles are shown for some of the sounds that are commonly produced from variable positions. Detailed information concerning articulatory positions and types of modulation may be found in the chapters mentioned above. A somewhat similar summary of the profiles and resonance position of the vowels is presented in Chapter 12.

VOWEL-LIKE NASALS



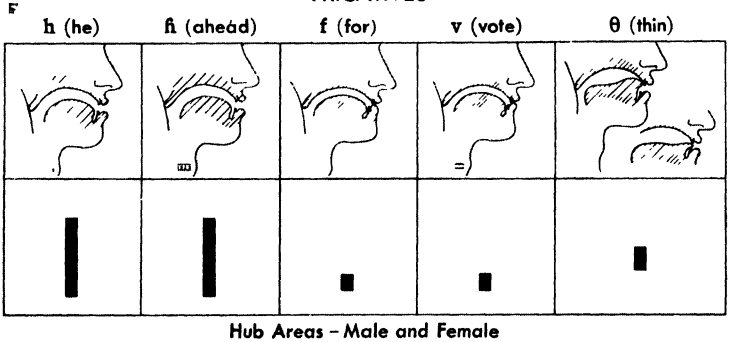
Hub Areas – Male and Female

STOPS



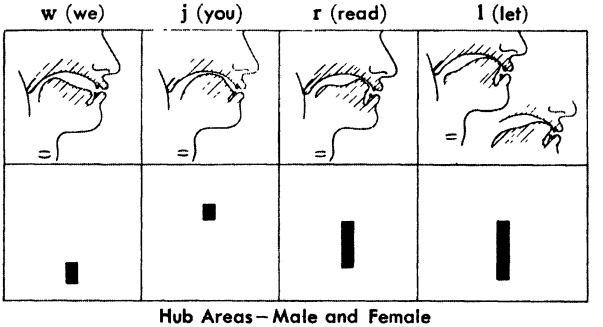
Hub Areas – Male and Female

FRICATIVES



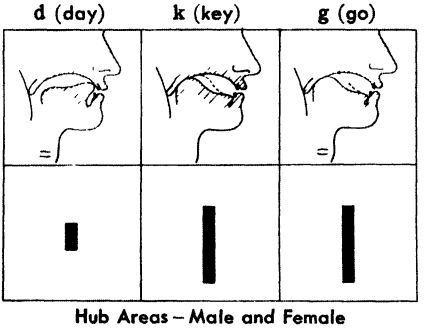
Hub Areas – Male and Female

VOWEL-LIKE GLIDES



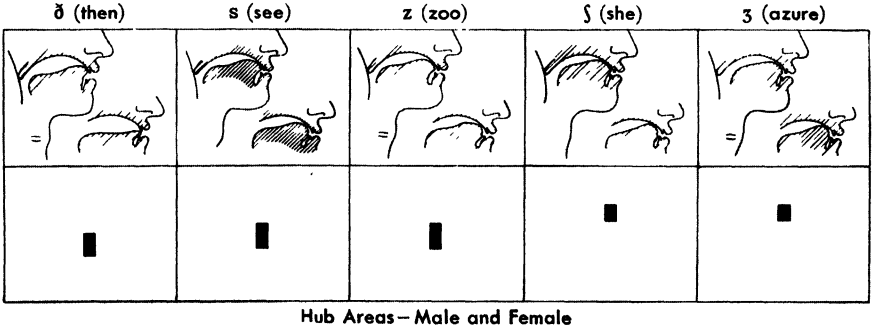
Hub Areas – Male and Female

STOPS



Hub Areas – Male and Female

FRICATIVES



Hub Areas – Male and Female

Bibliography of
Technical Publications on Visible Speech

1. R. K. Potter, *Visible Patterns of Sound*, SCIENCE 102, 465, Nov. 9, 1945. This is the original discussion of the development, with a brief review of its background, its possibilities, and early application studies.

2. R. K. Potter, *Introduction to Technical Discussions of Sound Portrayal*, J. ACOUS. SOC. AM. 18, 4, July 1946. Additional background material.

3. J. C. Steinberg and N. R. French, *The Portrayal of Visible Speech*, J. ACOUS. SOC. AM. 18, 4, July 1946. Discussion of objectives and requirements in the portrayal of speech from the standpoint of legibility. Such portrayal involves an intensity-frequency-time analysis and display of these dimensions so they may be seen individually. Optimum arrangements for this purpose are discussed in relation to the characteristics of the eye and the ear.

4. W. Koenig, H. K. Dunn, and L. Y. Lacy, *The Sound Spectrograph*, J. ACOUS. SOC. AM. 18, 19, July 1946. A technical description of the sound spectrograph, including a discussion of general problems of design and mechanical and electrical arrangements employed in one of the experimental models.

5. R. R. Riesz and L. Schott, *The Visible Speech Cathode-Ray Translator*, J. ACOUS. SOC. AM. 18, 50, July 1946. A technical discussion of a translator employing a moving screen cathode-ray tube, including discussion of tube characteristics and circuit arrangements.

6. Homer Dudley and Otto O. Gruenz, Jr., *Visible Speech Translators with External Phosphors*, J. ACOUS. SOC. AM. 18, 62, July 1946. A technical discussion of translator equipment in which the display screen is in the form of a moving belt coated with a phosphor. Experimental apparatus of this type is illustrated and described.

7. G. A. Kopp and H. C. Green, *Basic Phonetic Principles of Visible Speech*, J. ACOUS. SOC. AM. 18, 74, July 1946. A very brief general outline of the subject intended to show that the basic principles of visible speech are closely integrated with physiological phonetics.

8. J. B. Johnson, *A Cathode-Ray Tube for Viewing Continuous Patterns*, J. OF APP. PHYSICS 17, 91, Nov. 1946. A technical description of the moving screen cathode-ray tube.

Index

ABC'S, the new, 53
ABC's of visible speech, the, 54
Accents, 44
Adam's apple, 34
ALDA, FRANCES, 379, 380
American Indian, Ojibway (language interests), 367
Analysis by repetition, 11
Analysis of sound waves, 9
Animal sound patterns, 410
Aphasia, 342
Aphonia, 329
Applications of visible speech, 287–409
Arabic (language interests), 347
Articulators, 30, 32
Artificial larynx, 319
Arrowhead indicators, 29, 30
Aspect ratio, changes in, 305
Aspirate, 313
Auctioneer's chant, 322
AUTRY, GENE, 404

Baby's cry, patterns of, 412
Back vowels, reading the, 275
Back vowels, —
  hubs of, 69–72
  reading the, 275
Back vowels
  reference patterns
    ɑ (father), 70
    ɒ (not), 70
    ɔ (all), 70, 71
    o (obey), 71
    ʊ (foot), 71
    u (boot), 71, 72
  resonance positions
    ɑ (father), 275
    ɒ (not), 275
    ɔ (all), 276
    o (obey), 276
    ʊ (foot), 277
    u (boot), 277
Base line, 33
    definition of, 34
Basic principles of visible speech, 28–56
Basic speech processes, diagram and explanation of, 30–33
BELL, ALEXANDER GRAHAM, 3
BELL, MELVILLE, 3
Bibliography of technical publications, 430
Bird song patterns, 410, 412
BORI, LUCREZIA, 382, 393
Brain waves, patterns of, 414
Breaks in voice (*also see* Yodelling), 404

Cardinal, imitation of song, 406
CARUSO, ENRICO, 377–380, 391–393, 398, 399
Cavities, by cupped hands, 320
Cavity modulation, 31
Central vowels (*see* Mid vowels)
Chant, tobacco auctioneer's, 322
Chant, primitive music, 408
Characteristics of the hub principle, 33
Chart of progress in training, 27
Children, teaching of, 290
Chinese (language interests), 349
Classification of patterns, 33–38
Classification of sounds, 33–38
Cleft palate, 333
Clues in reading, 38, 39, 63
Cognate sounds, hubs of, 47, 48
COLONNA, JERRY, 403
Color, use of, to show intensity, 14
Combination modulation, 31, 33
Combination sounds (*also see* Combinations, Diphthongs), 56
  hw (when), 247–250
  tʃ (church), 153–157
  dʒ (judge), 153–157
Combinations (*also see* Diphthongs), —
  of fricative sounds, 153–159
  of fricatives and stops, 159–162

Combinations, of fricatives, identifying characteristics, 157–159
of glides and other sounds, 246–250
of glides and fricatives, 257–258
of glides and nasals, 259–260
of glides and stops, 251–256
of nasal sounds, 190–201
of nasal sounds, characteristics, 191
of nasals and fricatives, 195–197
of stop sounds, 104, 109
of sounds, 36, 37, 56, 153, 247
Combined sounds (*also see* Combination sounds, Diphthongs), 36, 37, 56
Comparative review (*see* Review)
Complex waves, analysis of, 8, 9
Consonant influence on vowels, 65
Consonant influence on diphthongs, 75
Consonant pairs, 43, 47, 48
Consonants, hubs of, 42, 43, 47, 48
Conversation by visible speech, 24, 26
Correction of speech, 325, 327
CROSBY, BING, 400, 401, 402
Cross-section patterns, 391
Cupped hands as cavity, 320

Danish (language interests), 351
Deaf interests, 287
Deaf, speech patterns of the, 296
Deafened, contributions of the, 7, 422
Defective speech, spectrograms illustrating, 329–342
Defects in speech, (*see* Speech defects), 329–342
Definitions (*see the item for which the definition is required*)
Development, need for, 7
Diacritical markings, 29, 425
Diagrams (*also see* Patterns, Profile diagrams), —
basic speech processes, 32
complex waves, 8
diphthongs, 270, 280
direct translator, 17
early hand-written symbols, 3
explanation of use of profile, 30
Diagrams, hub areas, 272–282, 428–429
manual symbols, 3, 60, 62
profiles, summary, 272–282, 428–429
(*also see* Profile diagrams)
progress in training, 27
resonance areas, 272–282, 427
simple wave, 8
sound spectrograph, 10
summary, of profiles and hub areas, 427
vowels, 272-282, 428–429
Dialect, effect upon hubs, 44
Diphthongs, —
classification of, 74, 75
consonant influence on, 75
hubs of, 75–78, 280
male and female voices, 280
modulation of, 74, 78
reference patterns of, 75–78, 280
Direct translator, 16
Discrimination index, 21
Dogs barking, patterns of, 412
Drill material, explanation of, 61
Drill phrases and sentences (*at the end of each Lesson Unit in Chapters* 8, 9, 10, *and* 11)
DUDLEY, HOMER W., 317, 430
DUNN, H. K., 430
Duration of fixation, 63
Dysarthria, 340
Dyslalia, 336
Dysphemia, 337
Dysphonia, 330

Ear, as reference in portrayal, 302
EDDY, NELSON, 390
Education in speech, 287
Effect of pitch and voice quality, 5, 6, 45
Encephalic potentials, patterns of, 414
English (language interests), 353
Equipment, experimental, 14
Equipment, limitations of, 288
Esophageal speech, 335
Essential characteristics of the hub principle, 43
Evaluation of visible speech, 289

Evidence of readability, 5
Experimental training, —
beginning of, 5, 22
chart of progress in, 27
class, 22
class schedules, 23
deaf member of class, 25
pictures of class, 23, 24
program, 22–27
progress, 5, 22, 27
results, 6, 22, 27

Fill, 33, 34
(*Also see* Spike fill, Voiced fill)
Fill, explanation of, 33, 34, 112–113, 119–120, 127–128, 135–136, 143, 144, 149–153
Filters, use in analysis of sound waves, 11, 12
Fixation span, 63
Foreign language interests, 343–375
American Indian (Ojibway), 367
Arabic, 347
Chinese, 349
Danish, 351
English, 353
French, 355
Gaelic, 357
German, 359
Icelandic, 361
Italian, 363
Norwegian, 365
Ojibway (American Indian), 367
Russian, 369
Spanish, 371
Swedish, 373
Yiddish, 375
Foreign language patterns, recording procedure, 343
FRENCH, N. R., 430
French (language interests), 355
Frequency resolution, 14
Frequency scale, 14
Frequency scale, linear versus logarithmic, 415

Fricative sounds, 34, 35
(*Also see* Fricatives, voiced and voiceless)
Fricative sounds, —
combinations of, 153–159
comparative review, 149–153
hubs of, 110–165
modulation of, 110–165
recognition of, 110
Fricatives voiced and voiceless, 33, 110–165
h (he) and **ĥ** (ahead), 111–118
f (for) and **v** (vote), 118–129
θ (thin) and ð (then), 126–133
s (see) and **z** (zoo), 134–142
ʃ (she) and ʒ (azure), 142–149
Frictional modulation, 31
Front vowels, —
reference patterns
i (eve), 66
ɪ (it), 67
e (hate), 67
ɛ (met), 68
æ (at), 68
a (ask), 68, 69
resonance positions
i (eve), 272
ɪ (it), 272
e (hate), 273
ɛ (met), 273
æ (at), 274
a (ask), 274
Front vowels, hubs of, 66–70
Front vowels, resonance positions, 272–274, 281, 282, 427

Gaelic (language interests), 357
Gap, explanation of, 33, 34, 82, 83, 88, 89, 94–96
Gap, voiced, 33, 34, 82, 83, 88, 89, 94–96
GARLAND, JUDY, 403
German (language interests), 359
GESCHW, 405
Glides, —
combinations of, 246–250
comparative review of, 244–246
hubs of, 202–269

Glides, modulation of, 202–269
  patterns of, —
    w (we), 203–211
    j (you), 211–218
    r (read), 218–231
    l (let), 231–244
  recognition of, 202
Glottal stop, 80
GREEN, HARRIET C., 300, 430
GRUENZ, OTTO O., JR., 430

Hand-written symbols, 3, 60–64
Heart beats, patterns of, 413
HEIFITZ, JASCHA, 415
HEMPEL, FRIEDA, 380
HOMER, LOUISE, 379
Hub, —
  areas, summary diagram of, 427
  definition of, 40
  essential characteristics of, 33
  explanation of, 39, 51, 271
  key to how sounds change each other, 39
Hubs (*also see* Sounds, *and the various classes of sounds*)
Hubs, —
  effect of pitch, 45
  effect of dialect, 44
  fixed and variable nature of, 49, 50, 51
  in diphthongs, 46, 65, 74
  in relation to profiles, 272–280, 427
  of back vowels, 69–72, 76
  of central vowels, 73
  of consonants, 42
  of consonant pairs, 47
  of different speakers, 44, 61, 63
  of fricative sounds
    h (he) and **ɦ** (ahead), 114
    f (for) and **v** (vote), 121, 122
    θ (thin) and **ð** (then), 128–130
    s (see) and **z** (zoo), 136–137
    ʃ (she) and **ʒ** (azure), 144–146
  of front vowels, 65, 69
  of glides
    w (we), 205–207
    j (you), 213–215
Hubs, of glides,
    r (read), 221–222
    l (let), 234–237
  of middle vowels, 73
  of nasal sounds
    m (me), 167–174
    n (no), 174–183
    ŋ (sing), 183–188
  of non-vowel sounds, 281
  of stop sounds
    p (pay) and **b** (be), 83–84
    t (to) and **d** (day), 89–90
    k (key) and **g** (go), 96–97
  of vowel-like sounds, 47
  relation to reading, 43, 44
  variable, 50
  visible and hidden nature of, 40–43

Icelandic (language interests), 361
Identifying characteristics (*also see* Sounds, *and the various classes of sounds*)
  of patterns, 33–38
  of stop sounds, 106
Improvement of patterns, 21, 410–422
Inaudible vibrations, patterns of, 117
Indicators, explanation of, 29, 30
Inflection, 33
Inflection, —
  effects of, 30, 45
  exaggerated, 322
  indication of, 20
  patterns of, 295
Influence, —
  of consonants on diphthongs, 75
  of consonants on vowels, 65
  of sounds, 38–51
Insect sound patterns, 410
International phonetic alphabet, use of, 29
  reference listing, 425
Italian (language interests), 363

JEPSON, HELEN, 389, 394
JOHNSON, J. B., 430
JOLSON, AL, 404

Key to how sounds change each other, 39
Key words, reference list of, 425
Key words, use of, 29, 60
KOENIG, W., 430
KOPP, GEORGE A., 300, 430

LACY, L. Y., 430
Larynx, 34
Larynx, artificial, 319
Learning rates for visible speech, 26, 27, 61
Legibility of patterns, 21
Lesson units, —
  description of, 56
  design of, 59–64
  purpose of, 59, 61, 64
Linear versus logarithmic frequency scale, 415
Logarithmic frequency scale, 14, 415
LOWERY, FRED, 406, 407

MCCORMACK, JOHN, 383
MACDONALD, JEANETTE, 390
Machinery noise, patterns of, 417
Manual symbols, —
  design of, 64
  explanation of, 63, 64
  illustration of, 60, 62
  of visible speech, 3, 62–64
  refinement of, 64
  relation to IPA type, 60, 62
MARTINELLI, GIOVANNI, 381, 382
Mechanism of speech, 30, 32
MELCHIOR, LAURITZ, 383, 398
MERCER, JOHNNY, 402
Michigan State Normal College, School of Special Education, 300
Michigan, University of, 300
Mid vowels, —
  reference patterns
    ʌ (up), 72, 73
    ə (about), 72, 73
    ɚ (word, bird), 72, 74
    ɜ (word, bird), 72, 74
Mid vowels, resonance positions
    ʌ (up), 278
    ə (about), 278
    ɚ (word, bird), 279
    ɜ (word, bird), 279
Mid vowels, hubs of, 72–74
  resonance positions of, 278–279, 281, 282
Modulation, cavity, 31
  combination, 31, 33
  definition of, 30
  frictional, 31
Modulation of back vowels, 69–72
  of diphthongs, 74–78
  of fricative sounds, 110–165
    h (he) and ɦ (ahead), 111–113
    f (for) and v (vote), 119–120
    θ (thin) and ð (then), 126–128
    s (see) and z (zoo), 134–136
    ʃ (she) and ʒ (azure), 142–143
  of front vowels, 66–69
  of glides, 202, 203–269
    w (we), 203–204
    j (you), 211–212
    r (read), 211–220
    l (let), 232
  of mid vowels, 72–74
  of nasal sounds, 166–167
    m (me), 167–168
    n (no), 175 176
    ŋ (sing), 183–184
  of stop sounds, 79–109
    p (pay) and b (be), 81
    t (to) and d (day), 87, 88
    k (key) and g (go), 94
  of vowels, 65, 66
  start-stop, 30
  types of, 30
  vocal cord, 31
MORSE, ELLA MAE, 403
Movements, transitional, 39
Music, instrumental, 415
  primitive, 408
  vocal, 376–409
Musical notation, sound patterns for, 416

Narrow-band filters, 11, 12, 13
Nasal sounds, 166–201
  m (me), 167–174
  n (no), 174–183
  ŋ (sing), 183–188
Nasal sounds, —
  combinations of, 190–201
  comparative review of, 188–190
  hubs of, 166–201
  modulation of, 166, 167
  recognition of, 166
Nasality, 313
Natural phonetic symbols, 28, 29
New ABC's, the, 53, 56–59
Nigeria, primitive music of, 408
Non-linearities in voicing, 311
Non-vowel hubs, summary of, 281, 427
Non-vowel sounds, hubs of, 281, 427
Northwestern University, 408
Norwegian (language interests), 365

Objectives of book, 7, 21
ONÉGIN, SIGRID, 380, 381
Oscillograph traces, 4, 8, 315

PAGET, SIR RICHARD, 308, 311
Palate, cleft, 333
Parrot speech, 323
Patterns (*also see listings of patterns under the various kinds of sounds*)
Patterns, —
  animal sounds, 410, 412
  baby's cry, 410, 412
  bird songs, 410, 412
  classification of, 33–38
  effects of pitch, 45
  effects of voice quality, 5, 6
  for speech education, 293
  identifying characteristics of, 33–38
  improvement problems, 21, 410–422
  insect sounds, 410–412
  manual type (*see* Manual symbols)
  non-essential detail, 61, 63
  of unstressed sounds, 51–53
Patterns, of stressed sounds, 51–53
  possibilities, 302, 415
  readability, 5
  relation to hearing, 14
  representation of, 64
  speech education, 293
  visual association of, 291
Personality in speech, 418
Phonetic interests, 302
Phonetic research possibilities, 307
Phonetic symbols, —
  International Phonetic Alphabet, 29, 425
  manual symbols of visible speech, 60, 62, 63
  natural form, 28, 29
  visible speech natural, 28, 56
Phonetic transcription, problems of, 321
Pitch, effect on hubs, 45
Pitch indication, 20, 45
Pitch patterns, 295
Pitch range, 31
Pitch range, wide, 304
Pitch scale, logarithmic, 304
Pitch waves, 400
Plosive sounds (*also see* Stop sounds), 31
Pointer system, 29
PONS, LILY, 385, 386, 387, 388, 394, 396, 397
POTTER, RALPH K., 300, 430
Primitive music, 407
Principles of visible speech, 28
Profile diagrams
  explanation of, 30, 61
  hubs in relation to, 272–280, 427
  of various sounds
    p (pay), 33
    b (be), 34
    t (to), 87
    d (day), 87
    k (key), 94
    g (go), 94
    h (he), 111
    ɦ (ahead), 111
    f (for), 35, 118
    v (vote), 35, 118

Profile diagrams, of various sounds,
θ (thin), 126
ð (then), 126
s (see), 134
z (zoo), 134
ʃ (she), 142
ʒ (azure), 142
tʃ (church), 37
m (me), 167
n (no), 174
ŋ (sing), 183
w (we), 203
j (you), 211
r (read), 218
l (let), 231
i (eve), 36, 272, 282
ɪ (it), 272, 282
e (hate), 273, 282
ɛ (met), 273, 282
æ (at), 274, 282
a (ask), 274, 282
ɑ (father), 275, 282
ɒ (not), 275, 282
ɔ (all), 276, 282
o (obey), 276, 282
ʊ (foot), 277, 282
u (boot), 277, 282
ʌ (up), 278, 282
ɝ (word, bird), 279, 282
ɜ (word, bird), 279, 282
eɪ (say), 280
aɪ (I), 37, 280
ɪu (new), 280
ɔɪ (boy), 280
ɑʊ (out), 280
oʊ (go), 280
Profile pictures (*also see* Profile diagrams), 30
Progress chart, 27
Pronunciation, differences in, 44
Pronunciation, regional, 45
Pronunciation versus spelling, 29
Proposed program, —
adult group, 293
Proposed program, hard-of-hearing, the, 293
intermediate group, 292
nursery group, 290
teachers, 289
Puff, 31

Quality, effect of voice, 5, 6, 45

Rate of speaking, transitional sounds, 38, 42
Readability of patterns, 5, 21
Reading (*also see the listings under the various classifications of sounds in this Index; also see the Table of Contents, Chapters 8 to 12, for list of reading various sounds*)
Reading, —
clues in, 63
eye movements in, 63
fixation span and duration in, 63
hubs in relation to patterns in, 43, 44
in relation to lesson units, 59
limitations of text book, 59
patterns, in relation to hubs when, 43, 44
procedure in learning, 59
proficiency in learning, 59, 61
progress chart, 27
significance of total patterns, 39, 40, 43, 63
visualization in, 61
word-sentence approach to, 63
Reading visible speech, 39, 43, 59
Reading visible speech, clues in, 63
Reading visible speech, fixation span and duration, 63
Recognition of fricative sounds, 110
Recognition of glides, 202
Recognition of nasal sounds, 166
Recognition of stop sounds, 79
Reference patterns (*also see the listings under the various kinds of sounds*)
Reference patterns, —
for back vowels, 69–72

Reference patterns, for diphthongs, 74–78
for front vowels, 66–69
for mid vowels, 72–74
Resolution, frequency, 14
Resonance bars, 33
Resonance bars, —
explanation of, 39, 202–203
numbering of, 36, 166
relation to articulation, 308
Resonance, by cupped hands, 320
Resonance cavities, 28
Resonance positions in relation to profile diagrams, 272–280, 427
Resonance positions of non-vowel sounds, 427
Resonance positions of vowel sounds, 272–280
Resonance sources, 309
Resonances, in relation to bars, 65
Review, comparative, —
of fricative sounds, 149–153
of glides, 244–246
of nasal sounds, 188–190
of stop sounds, 101–104
Rhythm, 312
RIESZ, R. R., 430
Russian (language interests), 369

Scale, frequency, 14, 415
SCHOTT, L., 430
Schwa, 72, 73, 270
SCOTTI, ANTONIO, 379
SCRIPTURE, EDMUND WILLER, 315
SEMBRICH, MARCELLA, 379
SEVERINA, GINA, 379
Simple wave, diagram of a, 8
SINATRA, FRANK, 402
Sound omissions and substitutions, 325
Sound patterns (*also see listings under the various classifications of sounds*)
Sound patterns, —
artistic designs in, 416
classification of, 33–38
improvements in, 21, 410–422
of animals, 410, 412
Sound patterns, of birds, 410, 412
of insects, 410
of machinery noises, 417
Sound portrayal, non-speech interests, 410
Sound spectrograph, 8–10, 15
Sound substitutions and omissions, 325
Sounds, classification of, 33–38
Sounds, combinations of (*also see* Combination sounds, Combinations, Diphthongs), 36, 37, 56, 153, 247
Sounds, glide types
Sounds, identifying characteristics of speech,—
p (pay), 33, 81–87
b (be), 34, 81–87
t (to), 87–93
d (day), 87–93
k (key), 94–100
g (go), 94–100
h (he), 111–1118
ɦ (ahead), 111–118
f (for), 35, 118–126
v (vote), 35, 118–126
θ (thin), 126–133
ð (then), 126–133
s (see), 134–142
z (zoo), 134–142
ʃ (she), 142–149
ʒ (azure), 142–149
tʃ (church), 37, 153–157
dʒ (judge), 153–157
m (me), 167–174
n (no), 174–183
ŋ (sing), 183–188
w (we), 203–211
j (you), 211–218
r (read), 218–231
l (let), 231–244
hw (when), 247–250
Sounds influence each other, key to how, 38
Sounds, nasal, 166–201
Sounds, of audience, 323
Sounds, —
stressed, 51–53
transitional, 38, 42
unstressed, 51–53

Sounds, voiced fricative, 35, 54, 110–165
Sounds, voiceless fricative, 34, 35, 54, 110, 111–165
Sounds, voiced stop, 34, 54, 79–109
Sounds, voiceless stop, 33, 34, 54, 79–109
Sounds, vowel and vowel-like, 36, 55, 166–269
Span of fixation, 63
Spanish (language interests), 371
Special Education, Michigan State Normal School for, 300
Spectrocardiograms, 414
Spectrograms, —
  dimensions of, 12
  illustrating speech defects, 329–342
  logarithmic versus linear scale in, 14, 415
  use of color for intensity in, 14
Spectrograph, 4, 8
Spectrograph, —
  characteristics of, 14
  diagram of, 10
  picture of, 15
Speculations concerning visible speech, 410–422
Speech, —
  artificial, 319
  by parrot, 323
  by ventriloquist, 321
Speech Clinic, University of Michigan, 300
Speech correction interests, 325
Speech correction, —
  records of progress, 327
  remedial therapy, 327
  research, 327
  training clinicians, 327
Speech defects, —
  aphasia, 342
  aphonia, 329
  cleft palate, 333
  dysarthria, 340
  dyslalia, 336
  dysphemia, 337
  dysphonia, 330
Speech education, 287
Speech education, patterns for, 294
Speech, esophageal, 335
Speech mechanism, 30
Speech mechanism, diagram of, 32
Speech patterns, 12
Speech patterns, intelligence in, 13
Speech patterns, whispered, 13
Speech, personality in, 418
Speech processes, diagram and explanations of, 30–32
Speech, quality of, 313
Speech, rapid, 322
Speech, rhythm in, 312
Speech sound classification, 33
Speech sounds, identifying characteristics of (*see* Sounds, identifying characteristics of speech)
Speed of reading, 63
Spelling variations, effect on patterns, 61
Spelling variations of
  p (pay) and **b** (be), 81
  t (to) and **d** (day), 88
  k (key) and **g** (go), 94
  h (he) and **ɦ** (ahead), 112
  f (for) and **v** (vote), 119
  θ (thin) and **ð** (then), 127
  s (see) and **z** (zoo), 134
  ʃ (she) and **ʒ** (azure), 143
  tʃ (church) and **dʒ** (judge), 153
  m (me), 168
  n (no), 175
  ŋ (sing), 184
  w (we), 204
  j (you), 212
  r (read), 220
  l (let), 232
  hw (when), 247
Spelling versus pronunciation, 29, 61
Spike fill, 33, 34, 82–83, 88–89, 94–96
Spike, voiced (*see* Voiced spike)
Start-stop modulation, 30
Steinberg, J. C., 430
Stops in vocal music, 392
Stop sounds, 31, 33, 34, —
  combinations of, 104–109
  comparative review of, 101–104
  drill phrases and sentences, 107–109

Stop sounds, hubs of, 79–109
identifying characteristics of, 106
(*also see* Sounds, identifying characteristics of)
modulation of, 79–109
recognition of, 79
words including, 106–107
Stop sounds, voiceless and voiced, 79–109, —
p (pay) and **b** (be), 81–87
t (to) and **d** (day), 87–93
k (key) and **g** (go), 94–100
Stress, 33, 38, 51
Stressed sound patterns, 51–53
Suggestions for using the lesson chapters, 59
Swedish (language interests), 373
Symbols, early hand-written, 3
Symbols, manual type, 62, 63, 64
Symbols, manual visible speech, 60, 63
Symbols, natural phonetic form, 28, 29

Teaching requirements, 22
Terminology, in relation to patterns, 33, 38
TETRAZZINI, LUISA, 380
TIBBETT, LAWRENCE, 384, 397, 398, 400
Timing, 38
Tobacco auctioneer's chant, 322
Tongue twisters, 317
Training (*also see* Experimental training)
Training, chart of progress in, 27
Training program, 22–27
Training requirements, 59
Training, visible speech and established methods, 291
Transitional influences, 39
Transitional movements, 39
Transitional sounds, rate of speaking, 38, 42
Translation, principles of, 287
Translation, technical problems, 324
Translator, 16–21
Translator, —
belt type, 17
cathode-ray type, 17, 18
characteristics of, 16
Translator, classroom, 20
development problems, 418
display requirements, 25, 26, 27
large display, 20
original model, 5, 16
pattern improvement, 20
pitch indication, 20
requirements for, 10
schematic diagram of, 17
small experimental type, 19
Trill, 320, 385
Twisters, tongue, 317
Typewriters, voice-operated, 419

University of Michigan, Speech Clinic, 300
Unstressed sound patterns, 51–53
Unstressed sounds, 51
Usefulness, 7

Variations in spelling (*see* Spelling variations)
Ventriloquist's speech, 321
Vibrato, —
accompanying enunciation, 389
in duet, 391
in relation to stops, 393
rate variations, 387
repetition of, 387
shape variations in patterns of, 382
wide swing, 385
Violin music, 415
Visible speech, —
basic advantages of, 421
basic principles of, 28
conversation by, 21
conversation rates in, 26
demonstrations of, 299
development through assistance of the deaf, 301
evaluation of, 3, 421
future ABC's through, 420
future development of, 410
future possibilities of, 419
name of, 3

Visible speech, possibilities with X-ray, 311
principles of, 28, 287, 315
progress in learning, 25, 27
quality detail in, 419
reading, 39, 418
reconversion to sounds of, 419
relation to spelling, 29, 61
relative to speed production, 307
research project, 5
speculations, 410
Visible speech versus shorthand, 420
Vocabulary, basic and supplemental, 26, 61
Vocal cord modulation, 31
Vocal cord vibration, change in rate of, 80
Vocal cords, partial paralysis of, 332
Vocal music, —
abrupt stops, 394
crescendo, 397
data on illustrations, 377
diminuendo, 347
falsetto, 397
glides, 395
interests, 376
laughs, 399
pattern requirements, 407
pitch wavers, 400
primitive, 407, 408
sobs, 399
steps, 395
stops, 392
visible versus audible effects, 376
word formation, 398
yodel, 404
Vocal quality, 33, 313
Vocoder speech, 317
Voice bar, 33, 34, 166
Voice quality, effect of, 5, 6, 45
Voiced and voiceless fricatives, 110–165
Voiced fill, explanation of, and examples, 35, 112–113, 119–120, 127–128, 135–136, 143–144, 149–153
Voiced fricative sounds, 35, 54, 110–165
Voiced gap, 34, 82, 83, 88, 89, 94–96
Voiced spike, 34, 82, 83, 88, 89, 94–96
Voiceless and voiced stop sounds, 34, 54, 79–109
(*Also see* Stop sounds, voiceless and voiced)
Voiceless fricative sounds, 34, 35, 54, 110–165
(*Also see* Fricative sounds)
Voicing bar, 33, 34, 166
Vowel and vowel-like sounds, 36, 55, 166–269
Vowel-like sounds, hubs of, 47
(*Also see* Hubs)
Vowel sounds, 36
Vowel sounds, hubs of, 272–280
(*Also see* Hubs)
Vowels, 65, 66, 69, 72
(*Also see* Back vowels, Front vowels, Mid vowels)
Vowels, —
classification of, 65
consonant influence on, 65
hubs of, 40, 41, 66, 74, 271–282
modulation of, 65, 66

WATERS, ETHEL, 400, 401
Wave analysis, 7, 8
Wave records (by SCRIPTURE), 315
Webster's diacritical markings, use of, 29, 425
Whisper, 13, 313
Whispered speech, 13, 313
Whistling, 312, 406
Wide-band filters, 11, 12, 13
Word structure, 39

Yiddish (language interests), 375
Yodel, by non-professional singer, 405
Yodelling, 404
Yoruba song, 408

ZWYER, 405